



**Universidade Federal de Uberlândia**  
**Faculdade de Matemática - FAMAT**  
**Coordenação Dos Cursos de Bacharelado e Licenciatura**  
**em Matemática**

**ALGUMAS OBSERVAÇÕES QUANTO AOS**  
**PROCESSOS DE AVALIAÇÃO EM MATEMÁTICA**

**TIAGO LEITÃO SOBREIRA**

Uberlândia-MG

2018

**TIAGO LEITÃO SOBREIRA**

**ALGUMAS OBSERVAÇÕES QUANTO AOS  
PROCESSOS DE AVALIAÇÃO EM MATEMÁTICA**

Trabalho de Conclusão do Curso  
apresentado à Coordenação do Curso de  
Licenciatura em Matemática como requisito  
parcial para obtenção do grau de Licenciado  
em Matemática.

Orientador: Walter dos Santos Motta Junior

Uberlândia-MG

2018



Universidade Federal de Uberlândia

Faculdade de Matemática

Coordenação do Curso de Licenciatura em Matemática

A banca examinadora, conforme abaixo assinado, certifica a adequação deste trabalho de conclusão de curso para obtenção do grau de Licenciado em Matemática.

Uberlândia, \_\_\_\_\_ de \_\_\_\_\_ de 20 \_\_\_\_\_

BANCA EXAMINADORA

---

Nome completo do orientador

---

Nome completo do segundo membro da banca de defesa

---

Nome completo do terceiro membro da banca de defesa

Uberlândia-MG

2018

# Agradecimentos

Agradeço a família por estar presente, por cuidar de mim, por me fazer companhia, por auxiliar dentro do possível, apoiar sempre, pela educação que me proporcionou, ensinamentos que vão ficar para toda a vida e claro, por me fazerem entender a importância de estudar e me dedicar a outras pessoas.

Agradeço aos amigos, pois auxiliaram muito no meu desenvolvimento no curso de diversas formas diferentes, além de estarem ali para distrair e divertir quando precisava.

Agradeço aos professores, por todos os ensinamentos que puderam me passar, oportunidades de estudo que me deram, incentivos quando o próximo objetivo parecia algo impossível, ajuda em projetos, trabalhos e dúvidas de todo tipo e principalmente pelos exemplos que irei seguir para já começar a profissão com alguns passos dados.

Agradeço aos alunos, pelos pensamentos que me fizeram ter, pelas perguntas que não soube responder, pelo carinho e atenção que tiveram comigo e por me darem cada vez mais motivos para seguir em frente sempre melhorando.

Agradeço a mim por ter aceitado mudar tanto e aceitar os conselhos recebidos, mesmo que poucos, por não ter desistido e enfim concluído o curso.

## RESUMO

O objetivo inicial deste trabalho de conclusão de curso é analisar como itens são classificados segundo a Matriz de Referência do ENEM, norteadora dos processos de elaboração de itens desse exame, sendo este o mais abrangente programa de avaliação em massa do Brasil na atualidade. Posteriormente, objetivamos analisar como a Teoria da Resposta ao Item, presente nas avaliações das respostas dos itens do ENEM, se estrutura conceitualmente. Para exemplificar a sua utilização associamos essa ferramenta com questões do exame de acesso ao PROFMAT 2018. A finalidade é traçar aspectos que orientem os professores elaboradores quanto ao desenvolvimento de processos avaliativos de diferentes naturezas.

**Palavras Chave:** Teoria de Resposta ao Item. Matriz de referência. ENEM. PROFMAT.

## **ABSTRACT**

The initial objective of this course completion work is analyze how question items are classified according to ENEM Reference Matrix, which guides the process of elaboration for this exam questions, that is actually the most embracing program of mass evaluation in Brazil. Later, we had as objective analyze the conceptually structure of Item Response Theory, being on ENEM items answers evaluations. To exemplify its utilization we associate this tool with questions from the access exam to PROFMAT 2018. The goal is to map aspects that guide elaborators teachers for developing evaluative process of different natures.

**Key-words:** Item Response Theory. Reference Matrix. ENEM. PROFMAT.

# SUMÁRIO

1 INTRODUÇÃO.....	7
2 O ENEM E A MATRIZ DE REFERÊNCIA EM MATEMÁTICA.....	10
3 A TEORIA DE RESPOSTA AO ITEM: ASPECTOS GERAIS.....	21
3.1 O Modelo Logístico Unidimensional de 3 Parâmetros.....	22
3.1.1 A Escala de Proficiência.....	23
3.1.2 A Curva Característica do Item (CCI).....	25
3.1.3 O Parâmetro de Dificuldade (b).....	26
3.1.4 O Parâmetro de Discriminação (a).....	29
3.1.5 O Parâmetro de Acerto ao Acaso (c).....	33
3.2 Análise Pedagógica do Item.....	34
4 UMA APLICAÇÃO PARCIAL DA TEORIA DE RESPOSTA AO ITEM NO EXAME NACIONAL DE ACESSO (ENA) DO PROFMAT – UFU / 2018 .....	36
5 CONSIDERAÇÕES FINAIS.....	43
6 REFERÊNCIAS.....	44

# INTRODUÇÃO

Nas últimas duas décadas, os processos de avaliação em larga escala foram difundidos e utilizados de diferentes maneiras. Dessa forma, foi necessária a criação de processos para medir a qualidade dessas avaliações.

Dentre os métodos designados para avaliação, foi criada nos anos 50 a Teoria de Resposta ao Item – TRI, por Frederic Lord<sup>2</sup>, cuja ideia era estabelecer uma medida única e comparável de avaliação, mesmo quando as populações são submetidas a avaliações diferentes. Lord utilizava apenas modelos dicotômicos do tipo certo ou errado e, em 1970, Samejima<sup>3</sup> generalizou a teoria e introduziu o modelo politômico (com diversas alternativas de resposta). Com a dificuldade computacional inicial ultrapassada, esse processo ganhou mais notoriedade e passou a ser aplicado em diversos países.

No Brasil, em 2009, o ENEM (Exame Nacional do Ensino Médio) adotou a TRI. A partir dessa data, podemos tirar conclusões mais interessantes e fundamentadas sobre o desempenho dos estudantes brasileiros da educação básica, comparando os resultados ano a ano, mesmo quando submetidos a populações diferentes.

1 Frederic M. Lord (1912 - 2000), nascido em Hanover foi um psicometrista. Ele foi fonte de grande parte da pesquisa sobre a teoria de resposta ao item, incluindo dois livros importantes: *Mental Test Scores* (1968, com Melvin Novick, e dois capítulos de Allen Birnbaum), e *Applications of Item Response Theory to Practical Testing Problems* (1980).

2 Fumiko Samejima (1930), nascida em Tokio, Dra. Fumiko Samejima é reconhecida como uma pioneira em psicometria moderna. Suas contribuições foram diversas, mas seu trabalho sobre a teoria traço latente, também conhecido como teoria de resposta ao item (TRI), tem sido descrito como um "ponto de referência ... na teoria teste moderno" (*Avaliação Educacional: Problemas e Práticas*, 1991). Ela é reconhecida uma fundadora da TRI politômico (Ackerman, 1998) e os seus métodos têm sido amplamente aplicado por mais de 30 anos.



O modelo mais usual e conhecido de se obter os resultados de uma prova é aquele que leva em consideração apenas os escores brutos dos indivíduos que realizaram o teste. Ou seja, o resultado da prova é obtido comparando a quantidade de itens respondidos corretamente por cada pessoa. Esse modelo, conhecido como Teoria Clássica dos Testes - TCT, é amplamente utilizado no processo de avaliação do conhecimento de um indivíduo, para o qual pretende-se atribuir um grau de aprovação ou reprovação em relação ao domínio de determinados conteúdos; ou é utilizado no processo de seleção, no qual além de se poder atribuir um grau de aprovação ou reprovação, é necessário classificar cada indivíduo em relação aos demais participantes do teste.

A Teoria Clássica dos Testes apresenta algumas limitações como modelo estatístico: não permite comparar indivíduos que tenham respondido a testes distintos e não permite fazer o acompanhamento de um mesmo sujeito durante as várias etapas do processo de construção do conhecimento. Isso decorre do fato de que na TCT os resultados obtidos dependem do conjunto de itens selecionados para a construção da referida prova, tornando o resultado encontrado um caso particular.

No campo das avaliações educacionais, o processo de aferição do conhecimento de um indivíduo dependerá diretamente deste conjunto de itens selecionados para a confecção do teste, é como se o conhecimento deste aluno variasse a cada prova diferente que ele faça. Dessa forma, uma pessoa que realize dois testes diferentes no mesmo dia pode ter o resultado muito bom em um dos testes e, no outro, um resultado muito ruim. Como forma de sanar essas limitações da TCT, cresceu a utilização de técnicas oriundas da TRI, que é um conjunto de modelos estatísticos utilizados para medir traços latentes de uma pessoa, isto é, características que variam de pessoa para pessoa e não podem ser aferidas ou observadas de forma direta, como altura, peso, idade etc. Para estimar um traço latente, deve-se utilizar variáveis secundárias que estão diretamente relacionadas a esse traço latente. Por exemplo, se desejarmos estimar a idade de uma pessoa sem cometer a indelicadeza de lhe perguntar diretamente qual é a sua idade, podemos perguntar se esta pessoa acompanhou os desdobramentos da Segunda Grande Guerra, se ela viu o Golpe Militar de 64 ou se participou do Movimento das Diretas Já.

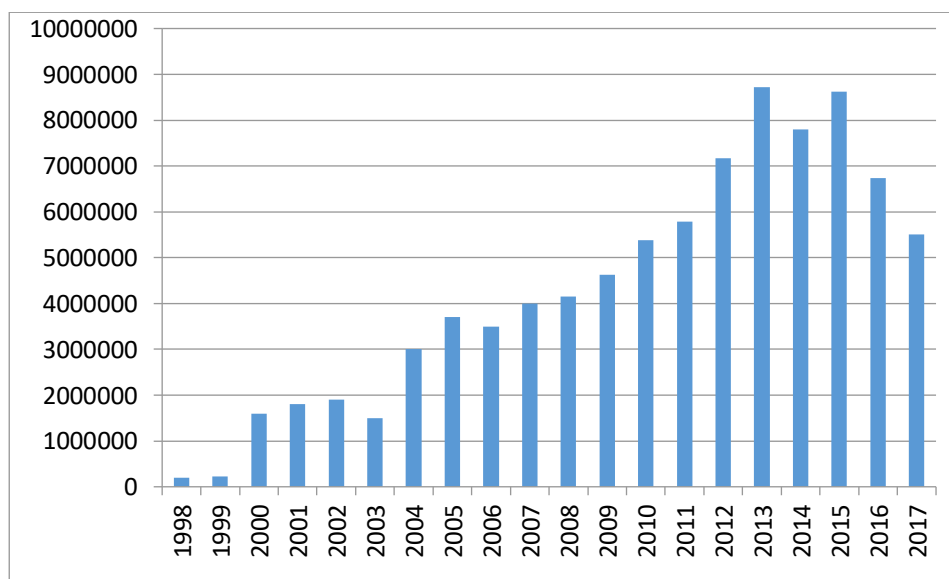
O presente trabalho propõe-se a apresentar inicialmente a matriz de referência utilizada pelo ENEM, principal instrumento em que a TRI atua, e elementos técnicos / qualitativos que fazem da TRI uma teoria capaz de sanar algumas deficiências presentes na TCT.

Em seguida, utilizando a TRI, faremos uma análise parcial de questões da prova de seleção do Mestrado Profissional em Rede Nacional em Matemática (PROFMAT) do ano de 2018.

## 2 O ENEM E A MATRIZ DE REFERÊNCIA EM MATEMÁTICA

Originalmente o Exame Nacional do Ensino Médio (ENEM), criado em 1998, constituía-se num exame individual, não obrigatório, oferecido aos estudantes do ensino médio, com o objetivo de auxiliar o governo na elaboração de políticas educacionais de melhoria da educação do país, além de permitir ao aluno fazer uma auto avaliação do aprendizado adquirido por ele durante a educação básica. Destaca-se inicialmente algumas políticas educacionais motivadas pelo ENEM, como a unificação do currículo do ensino médio e uma espécie de concorrência entre escolas públicas para obtenção de maiores notas no exame, favorecendo a obtenção de maiores recursos financeiros.

A prova do ENEM era composta, até 2008, por 63 questões de múltipla escolha estruturadas em uma matriz de 5 competências, que correspondiam a domínios específicos da estrutura mental, expressas por 21 habilidades, tendo um total de 3 questões para cada habilidade e uma redação, aplicados em um único dia. O quadro que segue, construído a partir de informações extraídas de boletins do INEP, mostra a evolução das participações de estudantes no exame ao longo dos últimos anos.



De acordo com os documentos oficiais do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) relativos ao exame, não havia o intuito de fazer com que o participante memorizasse fórmulas ou procedimentos algorítmicos, isso se evidencia de acordo com o seguinte trecho:

A mobilização de conhecimentos requerida pelo ENEM manifesta-se por meio da estrutura de competências e habilidades do participante, que o possibilita ler (perceber) o mundo que o cerca, simbolicamente representado pelas situações-problema, interpretá-lo (decodificando-o, atribuindo-lhe sentido) e, sentindo-se “provocado”, agir, ainda que em pensamento (atribui valores, julga, escolhe, decide, entre outras operações mentais).

Destaca-se no quadro anterior o aumento ocorrido na participação dos alunos em 2004. Nesse ano foi criado o programa “Universidade para Todos”, que concedia bolsa de estudos integrais ou parciais em instituições particulares de ensino superior. A partir deste momento, o ENEM passou a ter outra dimensão, algo além do que o inicialmente planejado, fato este que motivou o início de estudos almejando uma reestruturação no exame. Esta nova dimensão, fez com que a procura pelo ENEM aumentasse, aqueles que almejavam ingressar em uma instituição de ensino superior, iriam agora tentar conseguir uma bolsa e o ENEM se constituía num parâmetro avaliativo nesse sentido. Este aumento, fez com que instituições de ensino superior passassem a usar o ENEM em processos seletivos como forma de ingresso, em primeira ou segunda etapa de seus vestibulares, ou adicionando pontos na classificação obtida no vestibular. Destaca-se um trecho retirado de boletins do INEP nesse sentido:

A utilização dos resultados do ENEM nos processos de seleção das instituições de ensino superior foi a primeira modalidade social de uso do exame e, desde sua implantação, constitui-se no mais forte atrativo aos que a ele se submetem. Com o advento do Programa Universidade Para Todos (PROUNI), essa utilização social amplia-se cada vez mais com resultados efetivos no sentido de proporcionar o ingresso de jovens no ensino superior. Até então, a maioria desses jovens desistia de continuar os estudos tendo em vista a escassez de vagas em instituições públicas e à falta de condições de pagar uma faculdade particular.

A proposta de reestruturação do ENEM em 2009 tinha como objetivo central, democratizar as oportunidades às vagas, permitindo aos participantes, concorrer pelo ingresso em instituições

federais de ensino superior em qualquer região do país. Para isso, o Sistema de Seleção Unificada (Sisu) foi implementado.

De acordo com o documento intitulado Proposta: unificação dos processos seletivos das Instituições Federais de Ensino Superior a partir da reestruturação do Exame Nacional do Ensino Médio (ENEM), encaminhado pelo presidente do Inep/MEC à Associação Nacional dos Dirigentes das Instituições Federais de Ensino Superior (Andifes) em 30 de março de 2009, o que se deseja com a reformulação do exame é democratizar as oportunidades de concorrência às vagas federais de ensino superior por meio da unificação da seleção às vagas das IFES, utilizando uma única prova, e racionalizar a disputa por essas vagas, de forma a democratizar a participação nos processos de seleção para vagas em diferentes regiões do país. (RABELO, 2013)

Salientamos que existem estudos avaliando a migração de alunos nos estados visando o ingresso em instituições de ensino. Nesse contexto, uma possível justificativa para esta total liberdade de escolha por instituições em todo o país, prevista no ENEM, foi a comparação dos dados da Pnad/IBGE/2007, mostrando que nos Estados Unidos a porcentagem de alunos que mudaram de estado para ingressar em uma instituição de ensino superior era próxima de 20%, enquanto aqui no Brasil estava próxima de 0,04% na época. É fácil afirmar que esta ação teve êxito, pois já que em 2013, pelos dados aferidos, a porcentagem de alunos brasileiros que migraram de estado para cursar faculdade estava próxima de 13%. Outro motivo para as mudanças vividas recentemente no ENEM vinham de encontro à facilitação daqueles estudantes que viajavam por todo o país fazendo provas de cada instituição, estes agora fariam uma única prova e escolheria em qual iria ingressar, diminuindo muito seus custos e possíveis problemas estruturais.

Em 2009 a mudança na formatação do ENEM foi definida pela Portaria Inep/MEC nº 109, de 27 de maio de 2009, estabelecendo a sistemática para a realização do exame e realinhando as diretrizes e procedimentos, passando de avaliação do desempenho escolar e acadêmico dos participantes, para aferição do desenvolvimento das competências e habilidades fundamentais ao exercício da cidadania. Nesse dispositivo legal, são estabelecidos os seguintes objetivos para o exame:

- I. oferecer uma referência para que cada cidadão possa proceder à sua auto-avaliação com vistas às suas escolhas futuras, tanto em relação ao mundo do trabalho quanto em relação à continuidade de estudos;
- II. estruturar uma avaliação ao final da educação básica que sirva como modalidade alternativa ou complementar aos processos de seleção nos diferentes setores do mundo do trabalho;
- III. estruturar uma avaliação ao final da educação básica que sirva como modalidade alternativa ou complementar aos exames de acesso aos cursos profissionalizantes, pós-médios e à Educação Superior;
- IV. possibilitar a participação e criar condições de acesso a programas governamentais;
- V. promover a certificação de jovens e adultos no nível de conclusão do ensino médio nos termos do artigo 38, §§ 1o e 2o da LDB;
- VI. promover avaliação do desempenho acadêmico das escolas de ensino médio, de forma que cada unidade escolar receba o resultado global;
- VII. promover avaliação do desempenho acadêmico dos estudantes ingressantes nas Instituições de Educação Superior.

Nesta nova versão, o ENEM possibilita também que jovens e adultos consigam uma certificação de ensino médio completo caso atinjam uma nota mínima, o que antigamente era feito pelo Exame Nacional para Certificação de Jovens e Adultos (ENCCEJA) do ensino médio.

Um aspecto muito importante incorporado nesse ENEM revitalizado diz respeito ao sistema de análise das respostas apresentadas pelos alunos. É adotada a **Teoria de Resposta ao Item (TRI)**,

Para implantar essas mudanças, foi necessário recorrer às técnicas oriundas da TRI (Teoria de Resposta ao Item). A utilização dessa teoria no ENEM abriu a possibilidade de se construir uma série histórica do desempenho dos estudantes e dos egressos do Exames Nacionais de Avaliação Educacional ensino médio brasileiro, como é feito com os resultados do SAEB e da Prova Brasil. (RABELO, 2013)

Além da adoção da TRI, uma nova Matriz de Referência foi elaborada, contrastando principalmente no fato desta nova matriz abranger 30 habilidades distintas e distribuí-las em 45 questões, não garantindo assim, uma igualdade de cobrança de cada uma.

As competências de área foram submetidas ao tratamento cognitivo das competências do sujeito do conhecimento e permitiram a definição de habilidades específicas, que estabelecem as ações ou operações que descrevem desempenhos a serem avaliados na prova. Nessa concepção, as referências de cada área descrevem as interações mais abrangentes ou complexas (nas competências) e as mais específicas (nas habilidades) entre as ações dos participantes, que são os sujeitos do conhecimento, com os objetos de conhecimento, selecionados e organizados a partir das Orientações Curriculares Nacionais para o Ensino Médio (OCNEM). Essa estrutura aproxima o novo ENEM dessas Orientações, sem abandonar o modelo de avaliação centrado no desenvolvimento de competências. (RABELO, 2013)

No total, a nova prova possui 180 questões divididas em 45 questões por área do conhecimento, sendo estas:

- i. Linguagens, Códigos e suas Tecnologias (incluindo redação);
- ii. Ciências Humanas e suas Tecnologias;
- iii. Ciências da Natureza e suas Tecnologias;
- iv. Matemática e suas Tecnologias.

Decidiu-se ainda alterar o método de seleção das questões, que anteriormente eram elaboradas por empresas contratadas pelo Inep e agora passaram a ser extraídas de um banco de itens calibrados, elaborados sob a supervisão do próprio INEP, em um ambiente informatizado, mas sob rígidas regras de segurança. Os itens são pré-testados e escolhidos para compor as provas a partir de análises psicométricas oriundas da aplicação da teoria clássica dos testes e da teoria de resposta ao item. A função teórico-metodológico está fortemente presente na concepção de sua matriz de referência.

A construção desse banco de itens teve início em 2009, com a criação do novo modelo do ENEM. Esses itens são submetidos a um pré-teste para sua calibração. Este processo é descrito como:

(...) foram realizadas chamadas públicas para contratação de elaboradores e revisores de itens para esses processos. Os candidatos se inscreveram a partir de

editais específicos e, após uma pré-seleção via análise curricular, foram submetidos a capacitações focadas em técnicas de elaboração de itens. Esse trabalho tem revelado que uma quantidade significativa de professores desconhece as sutilezas inerentes ao processo de elaboração de itens para avaliação de larga escala, já que as licenciaturas não abordam esse tema quando da formação docente. Além disso, não existe literatura especializada que contemple as especificidades dos processos brasileiros, que se coadunam com os pressupostos presentes nas diretrizes e matrizes que norteiam as avaliações. (RABELO, 2013)

Vamos apresentar agora a Matriz de Referência atual em Matemática, definida por eixos cognitivos, competências de áreas e habilidades, a saber:

1. Eixos cognitivos (comuns a todas as áreas de conhecimento):

- i. Dominar linguagens (DL): dominar a norma culta da Língua Portuguesa e fazer uso das linguagens matemática, artística e científica e das línguas espanhola e inglesa.
- ii. Compreender fenômenos (CF): construir e aplicar conceitos das várias áreas do conhecimento para a compreensão de fenômenos naturais, de processos historicogeográficos, da produção tecnológica e das manifestações artísticas.
- iii. Enfrentar situações-problema (SP): selecionar, organizar, relacionar, interpretar dados e informações representados de diferentes formas, para tomar decisões e enfrentar situações-problema.
- iv. Construir argumentação (CA): relacionar informações, representadas em diferentes formas, e conhecimentos disponíveis em situações concretas, para construir argumentação consistente.



- v. Elaborar propostas (EP): recorrer aos conhecimentos desenvolvidos na escola para elaboração de propostas de intervenção solidária na realidade, respeitando os valores humanos e considerando a diversidade sociocultural.

## 2. Competências de área e habilidades

- i. Competência de área 1 (C1) - Construir significados para os números naturais, inteiros, racionais e reais.

H1 Reconhecer, no contexto social, diferentes significados e representações dos números e operações - naturais, inteiros, racionais ou reais.

H2 Identificar padrões numéricos ou princípios de contagem.

H3 Resolver situação-problema envolvendo conhecimentos numéricos.

H4 Avaliar a razoabilidade de um resultado numérico na construção de argumentos sobre afirmações quantitativas.

H5 Avaliar propostas de intervenção na realidade utilizando conhecimentos numéricos.

- ii. Competência de área 2 (C2) - Utilizar o conhecimento geométrico para realizar a leitura e a representação da realidade e agir sobre ela.

H6 Interpretar a localização e a movimentação de pessoas/objetos no espaço tridimensional e sua representação no espaço bidimensional.

H7 Identificar características de figuras planas ou espaciais.

H8 Resolver situação-problema que envolva conhecimentos geométricos de espaço e forma.

H9 Utilizar conhecimentos geométricos de espaço e forma na seleção de argumentos propostos como solução de problemas do cotidiano.

- iii. Competência de área 3 (C3) - Construir noções de grandezas e medidas para a compreensão da realidade e a solução de problemas do cotidiano.

H10 Identificar relações entre grandezas e unidades de medida.

H11 Utilizar a noção de escalas na leitura de representação de situação do cotidiano.

H12 Resolver situação-problema que envolva medidas de grandezas.

H13 Avaliar o resultado de uma medição na construção de um argumento consistente.

H14 Avaliar proposta de intervenção na realidade utilizando conhecimentos geométricos relacionados a grandezas e medidas.

- iv. Competência de área 4 (C4) - Construir noções de variação de grandezas para a compreensão da realidade e a solução de problemas do cotidiano.

H15 Identificar a relação de dependência entre grandezas.

H16 Resolver situação-problema envolvendo a variação de grandezas, direta ou inversamente proporcionais.

H17 Analisar informações envolvendo a variação de grandezas como recurso para a construção de argumentação.

H18 Avaliar propostas de intervenção na realidade envolvendo variação de grandezas.

- v. Competência de área 5 (C5) - Modelar e resolver problemas que envolvem variáveis socioeconômicas ou técnico-científicas, usando representações algébricas.

H19 Identificar representações algébricas que expressem a relação entre grandezas.

H20 Interpretar gráfico cartesiano que represente relações entre grandezas.

H21 Resolver situação-problema cuja modelagem envolva conhecimentos algébricos.

H22 Utilizar conhecimentos algébricos/ geométricos como recurso para a construção de argumentação.

H23 Avaliar propostas de intervenção na realidade utilizando conhecimentos algébricos.

- vi. Competência de área 6 (C6) - Interpretar informações de natureza científica e social obtidas da leitura de gráficos e tabelas, realizando previsão de tendência, extrapolação, interpolação e interpretação.

H24 Utilizar informações expressas em gráficos ou tabelas para fazer inferências.

H25 Resolver problema com dados apresentados em tabelas ou gráficos.

H26 Analisar informações expressas em gráficos ou tabelas como recurso para a construção de argumentos.

- vii. Competência de área 7 (C7) - Compreender o caráter aleatório e não determinístico dos fenômenos naturais e sociais e utilizar instrumentos adequados para medidas, determinação de amostras e cálculos de probabilidade para interpretar informações de variáveis apresentadas em uma distribuição estatística.

H27 Calcular medidas de tendência central ou de dispersão de um conjunto de dados expressos em uma tabela de frequências de dados agrupados (não em classes) ou em gráficos.

H28 Resolver situação-problema que envolva conhecimentos de estatística e probabilidade.

H29 Utilizar conhecimentos de estatística e probabilidade como recurso para a construção de argumentação.

H30 Avaliar propostas de intervenção na realidade utilizando conhecimentos de estatística e probabilidade.

Para exemplificar como os itens elaborados são classificados, segundo a matriz acima citada, vamos considerar algumas questões presentes no ENEM 2015 e descrever tais características presentes em cada um deles:

**MATEMÁTICA E SUAS TECNOLOGIAS**

**Questões de 136 a 180**

**QUESTÃO 136**

No contexto da matemática recreativa, utilizando diversos materiais didáticos para motivar seus alunos, uma professora organizou um jogo com um tipo de baralho modificado. No início do jogo, virá-se uma carta do baralho na mesa e cada jogador recebe em mãos nove cartas. Deseja-se formar pares de cartas, sendo a primeira carta a da mesa e a segunda, uma carta na mão do jogador, que tenha um valor equivalente àquele descrito na carta da mesa. O objetivo do jogo é verificar qual jogador consegue o maior número de pares. Iniciado o jogo, a carta virada na mesa e as cartas da mão de um jogador são como no esquema:

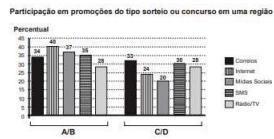


Segundo as regras do jogo, quantas cartas da mão desse jogador podem formar um par com a carta da mesa?

- A 9
- B 7
- C 5
- D 4
- E 3

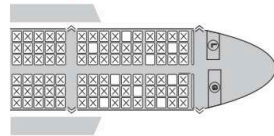
**QUESTÃO 137**

Uma pesquisa de mercado foi realizada entre os consumidores das classes sociais A, B, C e D que costumam participar de promoções tipo sorteio ou concurso. Os dados comparativos, expressos no gráfico, revelam a participação desses consumidores em cinco categorias: via Correios (juntando embalagens ou recortando códigos de barra), via internet (cadastrando-se no site da empresa/marca promotora), via mídias sociais (redes sociais), via SMS (mensagem por celular) ou via rádio/TV.



**QUESTÃO 144**

Uma família composta por sete pessoas adultas, após decidir o itinerário de sua viagem, consultou o site de uma empresa aérea e constatou que o voo para a data escolhida estava quase lotado. Na figura, disponibilizada pelo site, as poltronas ocupadas estão marcadas com X e as únicas poltronas disponíveis são as mostradas em branco.



O número de formas distintas de se acomodar a família nesse voo é calculado por

- A  $\frac{9!}{2!}$
- B  $\frac{9!}{7! \times 2!}$
- C 7!
- D  $\frac{5!}{2!} \times 4!$
- E  $\frac{5!}{4!} \times \frac{4!}{3!}$

Uma empresa vai lançar uma promoção utilizando apenas uma categoria nas classes A e B (A/B) e uma categoria nas classes C e D (C/D).

De acordo com o resultado da pesquisa, para atingir o maior número de consumidores das classes A/B e C/D, a empresa deve realizar a promoção, respectivamente, via

- A Correios e SMS.
- B internet e Correios.
- C internet e internet.
- D internet e mídias sociais.
- E rádio/TV e rádio/TV.

**QUESTÃO 138**

Uma fábrica de sorvetes utiliza embalagens plásticas no formato de paralelepípedo retangular reto. Internamente, a embalagem tem 10 cm de altura e base de 20 cm por 10 cm. No processo de confecção do sorvete, uma mistura é colocada na embalagem no estado líquido e, quando levada ao congelador, tem seu volume aumentado em 25%, ficando com consistência cremosa.

Inicialmente é colocada na embalagem uma mistura sabor chocolate com volume de 1 000 cm<sup>3</sup> e, após essa mistura ficar cremosa, será adicionada uma mistura sabor morango, de modo que, ao final do processo de congelamento, a embalagem fique completamente preenchida com sorvete, sem transbordar.

O volume máximo, em cm<sup>3</sup>, da mistura sabor morango que deverá ser colocado na embalagem é

- A 450.
- B 500.
- C 600.
- D 750.
- E 1 000.

**QUESTÃO 139**

Em uma central de atendimento, com pessoas a receberem senhas numeradas de 1 até 100. Uma das senhas é sorteada ao acaso.

Qual é a probabilidade de a senha sorteada ser um número de 1 a 20?

- A  $\frac{1}{100}$
- B  $\frac{19}{100}$
- C  $\frac{20}{100}$
- D  $\frac{21}{100}$
- E  $\frac{80}{100}$

**QUESTÃO 140**

Alguns medicamentos para felinos são administrados com base na superfície corporal do animal. Foi receitado a um felino pesando 3,0 kg um medicamento na dosagem diária de 250 mg por metro quadrado de superfície corporal.

O quadro apresenta a relação entre a massa do felino, em quilogramas, e a área de sua superfície corporal, em metros quadrados.

Relação entre a massa de um felino e a área de sua superfície corporal

Massa (kg)	Área (m <sup>2</sup> )
1,0	0,100
2,0	0,159
3,0	0,208
4,0	0,252
5,0	0,292

Para economizar em suas contas mensais de água, uma família de 10 pessoas deseja construir um reservatório para armazenar a água captada das chuvas, que tenha capacidade suficiente para abastecer a família por 20 dias. Cada pessoa da família consome, diariamente, 0,08 m<sup>3</sup> de água.

- A 0,624.
- B 52,0.
- C 156,0.
- D 750,0.
- E 1 201,9.

**QUESTÃO 141**

Para que os objetivos da família sejam atingidos, a capacidade mínima, em litros, do reservatório a ser constituído deve ser

- A 16.
- B 800.
- C 1 600.
- D 8 000.
- E 16 000.

**QUESTÃO 142**

Uma competição esportiva envolveu 20 equipes com 10 atletas cada. Uma denúncia à organização dizia que um dos atletas havia utilizado substância proibida. Os organizadores, então, decidiram fazer um exame antidoping. Foram propostos três modos diferentes para escolher os atletas que vão realizá-lo:

Modo I: sortear três atletas dentre todos os participantes;

Modo II: sortear primeiro uma das equipes e, desta, sortear três atletas;

Modo III: sortear primeiro três equipes e, então, sortear um atleta de cada uma dessas três equipes.

Considere que todos os atletas têm igual probabilidade de serem sorteados e que P(I), P(II) e P(III) sejam as probabilidades de o atleta que utilizou a substância proibida seja um dos escolhidos para o exame no caso do sorteio ser feito pelo modo I, II ou III.

Comparando-se essas probabilidades, obtêm-se

- A P(I) < P(II) < P(III)
- B P(II) < P(I) < P(III)
- C P(I) < P(I) < P(III)
- D P(I) = P(II) < P(III)
- E P(I) = P(II) = P(III)

**QUESTÃO 143**

Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), produtos sazonais são aqueles que apresentam ciclos bem definidos de produção, consumo e preço. Resumidamente, existem épocas do ano em que a sua disponibilidade nos mercados varejistas ora é escassa, com preços elevados, ora é abundante, com preços mais baixos, o que ocorre no mês de produção máxima da safra.

A partir de uma série histórica, observou-se que o preço P, em reais, do quilograma de um certo produto sazonal pode ser descrito pela função  $P(x) = 8 + 5\cos\left(\frac{\pi x - \pi}{8}\right)$ , onde x representa o mês do ano, sendo x = 1 associado ao mês de janeiro, x = 2 ao mês de fevereiro, e assim sucessivamente, até x = 12 associado ao mês de dezembro.

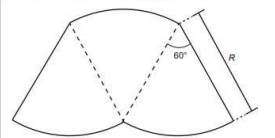
Disponível em: www.ibge.gov.br. Acesso em: 2 ago. 2012 (adaptado).

Na safra, o mês de produção máxima desse produto é

- A janeiro.
- B abril.
- C junho.
- D julho.
- E outubro.

**QUESTÃO 145**

O proprietário de um parque aquático deseja construir uma piscina em suas dependências. A figura representa a vista superior dessa piscina, que é formada por três setores circulares idênticos, com ângulo central igual a 60°. O raio R deve ser um número natural.



O parque aquático já conta com uma piscina em formato retangular com dimensões 50 m x 24 m.

O proprietário quer que a área ocupada pela nova piscina seja menor que a ocupada pela piscina já existente.

Considere 3,0 como aproximação para π.

O maior valor possível para R, em metros, deverá ser

- A 16.
- B 28.
- C 29.
- D 31.
- E 49.

**QUESTÃO 146**

Alguns exames médicos requerem uma ingestão de água maior do que a habitual. Por recomendação médica, antes do horário do exame, uma paciente deveria ingerir 1 copo de água de 150 mililitros a cada meia hora, durante as 10 horas que antecederiam um exame. A paciente foi a um supermercado comprar água e verificou que havia garrafas dos seguintes tipos:

- Garrafa I: 0,15 litro
- Garrafa II: 0,30 litro
- Garrafa III: 0,75 litro
- Garrafa IV: 1,50 litro
- Garrafa V: 3,00 litro

A paciente decidiu comprar duas garrafas do mesmo tipo, procurando atender à recomendação médica e, ainda, de modo a consumir todo o líquido das duas garrafas antes do exame.

Qual o tipo de garrafa escolhida pela paciente?

- A I
- B II
- C III
- D IV
- E V

**QUESTÃO 147**

Um concurso é composto por cinco etapas. Cada etapa vale 100 pontos. A pontuação final de cada candidato é a média de suas notas nas cinco etapas. A classificação obedece à ordem decrescente das pontuações finais. O critério de desempate baseia-se na maior pontuação na quinta etapa.

Candidato	Média nas quatro primeiras etapas	Pontuação na quinta etapa
A	90	60
B	85	85
C	80	95
D	60	90
E	60	100

A ordem de classificação final desse concurso é

- A A, B, C, E, D.
- B A, C, E, D.
- C B, E, A, D.
- D C, B, E, D, A.
- E E, C, D, B, A.

**QUESTÃO 148**

O índice pluviométrico é utilizado para mensurar a precipitação da água da chuva, em milímetros, em determinado período de tempo. Seu cálculo é feito de acordo com o nível de água da chuva acumulada em 1 m<sup>2</sup>, ou seja, se o índice for de 10 mm, significa que a altura do nível de água acumulada em um tanque aberto, em formato de um cubo com 1 m<sup>2</sup> de área de base, é de 10 mm. Em uma região, após um forte temporal, verificou-se que a quantidade de chuva acumulada em uma lata de formato cilíndrico, com raio 300 mm e altura 1 200 mm, era de um terço da sua capacidade.

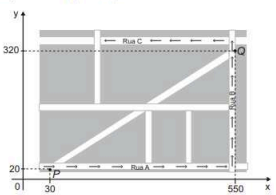
Utilize 3,0 como aproximação para π.

O índice pluviométrico da região, durante o período do temporal, em milímetros, é de

- A 10,8.
- B 12,0.
- C 32,4.
- D 108,0.
- E 324,0.

**QUESTÃO 149**

Devido ao aumento do fluxo de passageiros, uma empresa de transporte coletivo urbano está fazendo estudos para a implantação de um novo ponto de parada em uma determinada rota. A figura mostra o percurso, indicado pelas setas, realizado por um ônibus nessa rota e a localização de dois de seus atuais pontos de parada, representados por P e Q.



Os estudos indicam que o novo ponto T deverá ser instalado, nesse percurso, entre as paradas já existentes P e Q, de modo que as distâncias percorridas pelo ônibus entre os pontos P e T e entre os pontos T e Q sejam iguais.

De acordo com os dados, as coordenadas do novo ponto de parada são

- A (290; 20).
- B (410; 0).
- C (410; 20).
- D (440; 0).
- E (440; 20).

**QUESTÃO 150**

Deseja-se comprar lentes para óculos. As lentes devem ter espessuras mais próximas possíveis da medida 3 mm. No estoque de uma loja, há lentes de espessuras: 3,10 mm; 3,021 mm; 2,96 mm; 2,099 mm e 3,07 mm.

Se as lentes forem adquiridas nessa loja, a espessura escolhida será, em milímetros, de

- A 2,099.
- B 2,96.
- C 3,021.
- D 3,07.
- E 3,10.

Número da Questão	Competência	Habilidade	Objeto de Conhecimento Avaliado
136	<b>C1</b>	H1	Números Racionais
137	<b>C6</b>	H24	Análise de Dados
138	<b>C2</b>	H8	Volume
139	<b>C7</b>	H28	Noções de Probabilidade
140	<b>C6</b>	H26	Análise de Dados
141	<b>C3</b>	H14	Conversão Entre Sistemas de Medidas
142	<b>C7</b>	H29	Noções de Probabilidade
143	<b>C5</b>	H22	Funções Trigonométricas
144	<b>C1</b>	H2	Princípios de Contagem
145	<b>C2</b>	H9	Áreas
146	<b>C3</b>	H14	Unidades de Medidas de Grandezas
147	<b>C7</b>	H29	Média
148	<b>C5</b>	H23	Equações
149	<b>C2</b>	H9	Geometria Plana
150	<b>C1</b>	H1	Números Racionais

## 3 A TEORIA DE RESPOSTA AO ITEM: ASPECTOS GERAIS

Vejamos alguns aspectos que definem a TRI. Essa ferramenta se caracteriza como um conjunto de modelos matemáticos que procuram representar a probabilidade de um indivíduo dar certa resposta a um item, como função dos parâmetros do item e da habilidade (ou habilidades) do respondente. Essa relação é sempre expressa de tal forma que quanto maior a habilidade, maior a probabilidade de acerto no item. É fundamental tratarmos de alguns fatores presentes nessa caracterização. O primeiro a ser tratado é a habilidade ( $\theta$ ) ou proficiência do respondente, esta se refere ao nível de aptidão de um indivíduo para responder corretamente um conjunto de itens. Essa habilidade é o traço latente que queremos medir desse indivíduo. O segundo fator aborda os parâmetros do item, são eles:

- i. parâmetro de discriminação (a) – consiste na aptidão do item em distinguir indivíduos com habilidades diferentes;
- ii. parâmetro de dificuldade (b) – trata-se da habilidade mínima que um respondente precisa para ter uma grande probabilidade de dar a resposta correta;
- iii. parâmetro de acerto ao acaso (c) – é a probabilidade de um respondente com baixa proficiência responder corretamente um item.

Sendo assim, a TRI procura medir variáveis não observáveis (traço latente) que influenciam as respostas dadas aos itens, utilizando a aferição das variáveis observáveis (respostas aos itens). Ou seja, estabelece uma relação entre a habilidade do respondente e os parâmetros do item com a probabilidade de acerto no item, de tal forma que, quanto maior a proficiência do indivíduo, maior é a sua probabilidade de responder corretamente o item.

Como a TRI é um conjunto de modelos estatísticos, vale salientar a importância de conhecer os principais modelos para saber escolher o mais adequado a ser utilizado. No Brasil, o modelo predominantemente utilizado é o logístico unidimensional de 3 parâmetros. À vista disso, esse trabalho limitar-se-á na abordagem das principais características desse modelo,

utilizando-o no próximo capítulo para analisar parcialmente a prova de acesso ao PROFMAT UFU (2018).

Destacamos a importância em conhecermos as variáveis que diferenciam os modelos entre si. A escolha do modelo a ser empregado deve levar em consideração os três aspectos a seguir:

- Natureza do item: refere-se à forma como os itens são corrigidos. Podem ser itens dicotômicos (sim ou não) ou dicotomizados (corrigidos como certo ou errado), ou itens não dicotômicos;
- Quantidade de populações envolvidas: trata-se do número de populações que responderão os testes. Podemos ter uma população ou mais de uma;
- Número de traços latentes que estão sendo medidos: versa sobre a quantidade de habilidades que o teste medirá. Pode ser uma habilidade (modelos unidimensionais) ou mais de uma.

### 3.1 O Modelo Logístico Unidimensional de 3 Parâmetros (ML3)

Com o intuito de facilitar a compreensão deste modelo vamos apresentar a definição do Modelo Logístico de 3 Parâmetros para então explicar como cada conceito influencia na caracterização dada. Conforme encontrado em Andrade, Tavares e Valle (2000):

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

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

$U_{ij}$  é uma variável dicotômica que assume os valores 1, quando o indivíduo  $j$  responde corretamente o item  $i$ , ou 0 quando o indivíduo  $j$  não responde corretamente ao item  $i$ .

$\theta_j$  representa a habilidade (traço latente) do  $j$ -ésimo indivíduo.

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

$b_i$  é o parâmetro de dificuldade (ou de posição) do item  $i$ , medido na mesma escala da habilidade.

$a_i$  é o parâmetro de discriminação (ou de inclinação) do item  $i$ , com valor proporcional à inclinação da Curva Característica do Item — CCI no ponto  $b_i$ .

$c_i$  é o parâmetro do item que representa a probabilidade de indivíduos com baixa habilidade responderem corretamente o item  $i$  (muitas vezes referido como a probabilidade de acerto casual).

$D$  é um fator de escala, constante e igual a 1. Utiliza-se o valor 1,7 quando deseja-se que a função logística forneça resultados semelhantes ao da função ogiva normal.

Os valores dos parâmetros  $a$ ,  $b$  e  $c$ , são calculados através de pré-testagens (calibragem dos itens), utilizando o método da máxima verossimilhança. Este modelo leva em consideração os três parâmetros de um item para determinar a proficiência de um indivíduo após responder um conjunto de itens. Porém, é importante saber o que realmente quer dizer uma pessoa ter uma determinada proficiência ou habilidade em Matemática. Isso deve ser entendido como uma pessoa com competência (capacidade) para entender e praticar conhecimentos da área de Matemática.

### 3.1.1 A Escala de Proficiência

O processo de construção de uma escala de habilidade permite que a mesma assumam valores reais. Dessa forma, faz-se necessário para a construção da escala escolher uma origem, representada pelo valor médio das proficiências dos indivíduos que responderam os testes, e uma unidade de medida, representada pelo desvio-padrão das habilidades dos respondentes do teste. Geralmente, utiliza-se uma escala com média igual a 0 e desvio-padrão 1, representada por escala (0, 1) ou uma escala com média igual a 500 e desvio-padrão 100, representada por escala (500, 100), sendo esta última escala a utilizada no Exame Nacional do ensino Médio (ENEM). Quanto ao parâmetro de discriminação, devemos dividi-lo por 100 quando passamos da escala (0,1) para a escala (500, 100).

É importante ressaltar que independente da escala adotada os resultados encontrados serão os mesmos. E, mais ainda, que a interpretação feita sob o olhar das duas escalas é a mesma. Por exemplo, um indivíduo com habilidade 2 na escala (0, 1) tem proficiência de 2 desvios-padrão



acima da média. Correspondendo na escala (500, 100) à habilidade 700, pois também representa 2 desvios-padrão acima da média. Uma fórmula para fazer a mudança de escala da habilidade de um indivíduo é:

$$x = \frac{y - 500}{100} ,$$

onde x representa a proficiência na escala (0, 1) e y representa a proficiência na escala (500, 100).

Analisando uma situação hipotética, será possível compreender mais facilmente o que significa uma pessoa ter uma certa habilidade e acentuaremos mais ainda as principais diferenças entre a Teoria Clássica e a Teoria de Resposta ao Item.

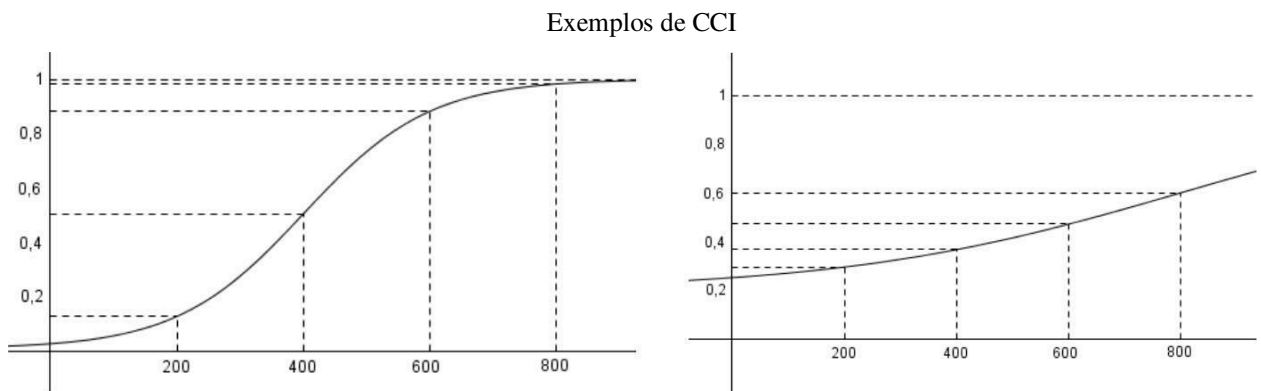
Imaginemos agora que um grupo responda uma prova muito especial, pois esta prova utiliza a TCT e a TRI para medir o conhecimento dos seus respondentes. Após a correção e a divulgação dos resultados, verifica-se que foram atribuídas a cada candidato duas notas: uma baseada no escore bruto, sendo determinada pelo percentual de acerto (TCT) e outra baseada nos três parâmetros (discriminação e dificuldade do item e acerto ao acaso) descritos anteriormente (TRI).

Observando as notas de um respondente hipotético, observamos que obteve nota 5, ou seja, acertou 50% dos itens da prova e teve uma proficiência de 600 pontos. Em relação à primeira nota, só podemos afirmar que ele acertou a metade dos itens da prova, não conseguindo explicitar se foram itens fáceis, medianos ou difíceis, também não se obtêm uma faixa de acerto desses itens, pois não se consideram os itens acertados no “chute”, não sabemos quais itens diferem os indivíduos com uma grande habilidade daqueles com pouca habilidade. Analisando a segunda nota, conclui-se que esse respondente tem uma probabilidade alta de acertar os itens com parâmetro de dificuldade até 600 pontos. Não significa que ele tenha acertado todos, mas sim que o mesmo tenha um grande índice de acertos em relação aos itens abaixo dessa faixa de dificuldade (pode-se pensar numa faixa do tamanho do desvio padrão e centrada na proficiência, isto é, variando de 550 até 650 pontos). Além disso, é possível classificar os itens nessas faixas de habilidades, determinando, assim, em quais conteúdos ele tem um domínio maior.

Vale destacar que a proficiência e o Parâmetro de Dificuldade estão sempre na mesma escala, facilitando, portanto, a interpretação gráfica, por meio da qual pode-se analisar simultaneamente ambos os conceitos.

### 3.1.2 A Curva Característica do Item (CCI)

O modelo matemático que define a TRI é uma função probabilidade. Portanto, sua imagem estará sempre no intervalo  $[0, 1]$ . De acordo com Rabelo (2013), “o número  $P(U_{ij} = 1|\theta_j)$  pode ser identificado com a proporção de respostas corretas ao item  $i$  no grupo de indivíduos com habilidade  $\theta_j$ ”. Essa relação é descrita através de uma curva sigmoide (curva em forma de “S”), onde o eixo horizontal representa a escala de habilidade e o eixo vertical a probabilidade do indivíduo com uma habilidade  $\theta_j$  dar a resposta correta ao item  $i$ . Podemos destacar duas assíntotas horizontais e notam-se, com uma certa precisão, os três parâmetros do item. Essa curva recebe o nome de Curva Característica do Item (CCI) e permite-nos observar as principais informações a respeito de um item de forma rápida e clara.



Fonte: Gomes, (2014)

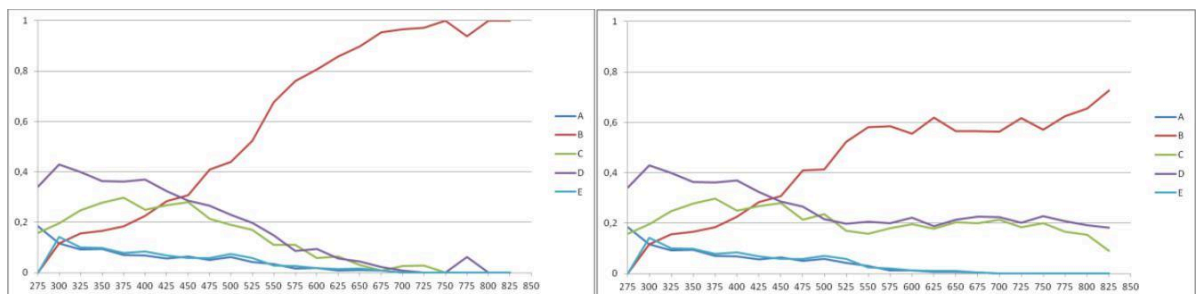
No gráfico à esquerda, podemos observar uma CCI com os valores dos parâmetros dentro do esperado para termos um item cumprindo bem a sua função avaliadora. O gráfico da direita mostra a CCI de um item com baixa dificuldade e pouco poder de discriminação.

Na teoria clássica, temos a Análise Gráfica do Item – AGI, que corresponde a um gráfico de linhas onde é representada a proporção de respostas dadas em cada opção de um determinado item em relação ao escore bruto total dos indivíduos que fizeram o teste. Na prática, o gráfico

mostra no eixo horizontal, as faixas de notas em ordem crescente obtidas pelos respondentes do teste, e apresenta uma linha para representar cada uma das alternativas da questão. É esperado que os percentuais de escolhas das linhas correspondentes às alternativas erradas vão diminuindo conforme o escore total vá aumentando e o percentual de escolha da linha correspondente a alternativa correta aumente conforme o escore total aumente.

Ao analisarmos um item considerando sua AGI, podemos verificar pedagogicamente o porquê uma opção errada atrai indivíduos com bom desempenho, proporcionando, assim, um melhor entendimento sobre o processo de ensino-aprendizagem do estudante.

Exemplos de AGI



Fonte: Gomes, (2014)

O gráfico a esquerda apresenta a AGI de um item considerado bom, pois conforme aumenta a proficiência dos respondentes a frequência de escolhas das alternativas incorretas diminuiu. No gráfico a direita observamos que alguns itens incorretos são atrativos para indivíduos com uma proficiência alta, mostrando que o item apresenta alguma falha na sua construção.

### 3.1.3 O Parâmetro de Dificuldade (b)

O parâmetro de dificuldade sempre estará na mesma escala da habilidade, o que sugere uma relação intrínseca entre os dois conceitos. Uma forma de se pensar no parâmetro “b”, é defini-lo como sendo o nível de habilidade mínima necessária para um indivíduo ter uma probabilidade alta para acertar o referido item. Essa probabilidade alta é igual a  $(1+c)/2$ , onde  $c$  é o parâmetro de acerto ao acaso do item. Se desconsideramos o acerto ao acaso ( $c = 0$ ), a dificuldade do item fica definida como a proficiência mínima que um respondente precisa para ter a probabilidade de 50% para dar a resposta correta.

Agora vamos apresentar uma função probabilidade (que assumirá valores no intervalo  $[0, 1]$ ) e que também seja uma função logística (lembre-se do nome deste modelo). Essa função é:

$$F(x) = \frac{1}{1+e^{-x}}$$

Devemos refletir neste momento sobre quais adaptações devem ser feitas para introduzirmos o parâmetro de dificuldade na função acima. 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, o que varia é o nível de proficiência, que determinará a probabilidade de acerto da questão. Portanto, teremos uma variação crescente da diferença entre a proficiência e a dificuldade do item. Por conseguinte, substituiremos o incremento “ $x$ ” da função  $F(x)$ . O que acarretará na função:

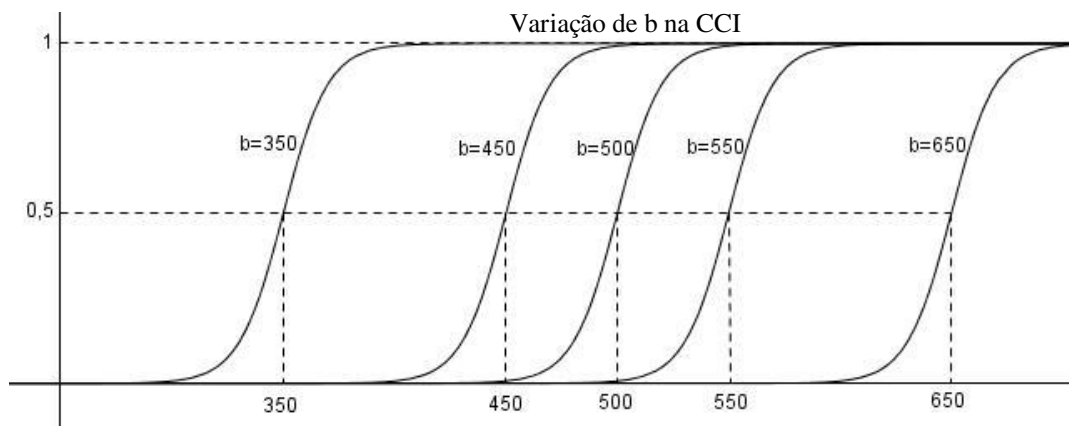
$$P(U = 1|\theta) = \frac{1}{1+e^{-(\theta-b)}}$$

Nesse momento, é importante recordarmos que uma prova é composta por diversos itens e é feita por muitos alunos com habilidades distintas. Logo, é necessário acrescentarmos variáveis para representar os diferentes itens ( $i$ ) e os diferentes alunos ( $j$ ), transformando nossa equação em:

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

onde  $U_{ij}$  representa a resposta do aluno  $j$  ao item  $i$ . Pode assumir valor 1, quando o indivíduo  $j$  responde corretamente o item  $i$ , ou valor 0, quando indivíduo  $j$  não responde corretamente o item  $i$ ;  $\theta_j$  representa a habilidade do aluno  $j$ ;  $b_i$  é a dificuldade do item  $i$ .

A seguir temos alguns exemplos de itens com diferentes níveis de dificuldade.



Fonte: Gomes, (2014)

Um ponto muito relevante acerca da dificuldade do item remete ao fato de como saber se um item é considerado muito fácil, fácil, mediano, difícil ou muito difícil. Ressaltamos também a importância de se conhecer a distribuição dos itens de um teste em relação aos níveis de dificuldades citados anteriormente. Destaca-se que o nível de dificuldade ideal para os itens de um teste depende da sua finalidade. Em avaliação educacional, recomenda-se uma distribuição de níveis de dificuldade de itens no teste dentro de uma curva normal: 10% dos itens em cada uma das duas faixas extremas, 20% em cada uma das faixas seguintes e 40% na faixa média. Segundo Rabelo (2013), a tabela a seguir mostra a distribuição e a classificação dos itens adotados por grande parte dos autores para testes de avaliação educacional, utilizando a escala (500, 100). Vale ressaltar que, geralmente, as provas de Matemática apresentam questões com um alto grau de dificuldade, comprometendo assim a utilização da tabela.

Tabela: classificação da questão de acordo com b (dificuldade do item)

CLASSIFICAÇÃO	DISTRIBUIÇÃO ESPERADA	DIFICULDADE DO ITEM
Muito fáceis	10%	Até 372
Fáceis	20%	De 373 a 448
Medianos	40%	De 449 a 551
Difíceis	20%	De 552 a 627
Muito Difíceis	10%	628 em diante

Fonte: Rabelo, (2013,p. 134)

Na verdade, a dificuldade de um item varia nos reais, porém, em aproximadamente 99,7% dos casos, esses valores estão no intervalo (200, 800). Dessa forma, questões com o parâmetro “b” fora desse intervalo, sugerem que estas devem ser descartadas por apresentarem algum erro na sua construção. Fazendo uma analogia, para se medir a dificuldade de um item na TCT, basta calcular a proporção de acertos do item. Ou seja, é a razão entre o número de indivíduos que responderam corretamente o item pelo total de indivíduos que foram submetidos ao item.

### 3.1.4 O Parâmetro de Discriminação (a)

Primeiramente, para entendermos melhor esse parâmetro, devemos ter clareza a respeito do que venha ser a discriminação de um item. É plausível pensarmos que numa prova feita com vários respondentes com habilidades distintas, alguns itens serão considerados fáceis pelos indivíduos com uma proficiência alta, mas podem ser considerados difíceis pelos indivíduos com uma baixa proficiência. A discriminação é exatamente essa característica do item capaz de diferenciar indivíduos com habilidades distintas. Dessa forma, o parâmetro de discriminação, é o responsável por aumentar, ou não, a diferença entre as probabilidades de alunos com habilidades distintas responderem corretamente o item.

Ao observarmos a CCI, o parâmetro de discriminação é encontrado no ponto de inflexão da curva, isto é, o ponto onde ela muda a sua concavidade, tendo um valor proporcional à inclinação desta curva. Espera-se que um item tenha o valor do parâmetro “a” positivo, pois, caso contrário, a CCI indicaria que, quanto maior a proficiência de um indivíduo, menor a sua probabilidade de responder corretamente o item. E, mais ainda, espera-se também que o valor do “a” não seja muito próximo do zero, uma vez que isso definiria que indivíduos com proficiências muito distintas tenham, praticamente, a mesma probabilidade de acertarem o item. Os itens que não diferenciam indivíduos com habilidades diferentes são ditos como tendo um baixo poder de discriminação. De acordo com Rabelo (2013), para esclarecer quando um item é ou não discriminativo, alguns autores utilizam a classificação abaixo como referência. Porém, é senso comum que itens com valor de discriminação acima de 0,0070 tendem a discriminar melhor, entretanto, se essa discriminação tiver um valor muito elevado, acima de 0,0170, os itens podem não funcionar também.

Tabela : classificação de questão de acordo com a (discriminação)

VALORES DO PARÂMETRO "a"	DISCRIMINAÇÃO
$a = 0$	Nenhuma
$0 < a \leq 0,0035$	Muito baixa
$0,0035 < a \leq 0,0065$	Baixa
$0,0065 < a \leq 0,0135$	Moderada
$0,0135 < a \leq 0,0170$	Alta
$a > 0,0170$	Muito alta

Fonte: Rabelo, (2013)

Quando inserimos na equação que segue

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

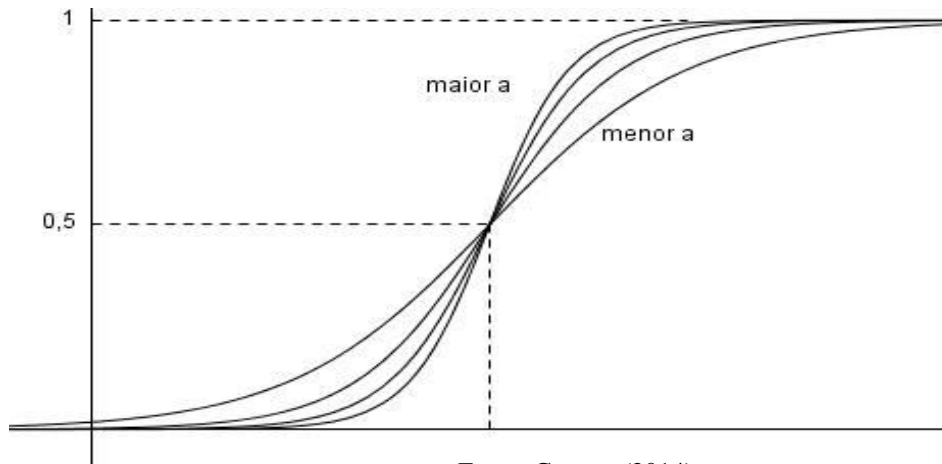
o parâmetro de discriminação, ela passa ter a seguinte forma:

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

com  $a_i$  representando a discriminação do item  $i$ .

O gráfico a seguir mostra as variações na CCI em decorrência de alterações feitas nos valores do parâmetro "a".

Figura: variação de a na CCI.



Fonte: Gomes, (2014)

Na TCT, a discriminação de um item é feita da seguinte forma: separamos os respondentes do teste em três grupos, conforme o escore total no teste, chamados de grupo inferior (corresponde aos 27% de menor desempenho), grupo superior (corresponde aos 27% de maior desempenho) e grupo intermediário (corresponde aos 46% restantes). Almeja-se que a proporção de acertos do grupo superior ( $P_{sup}$ ) seja maior que a proporção de acertos do grupo intermediário ( $P_{int}$ ) e que a proporção de acertos do grupo intermediário seja maior que a proporção de acertos do grupo inferior ( $P_{inf}$ ). Sendo assim, a discriminação do item é definida como a diferença entre a proporção de acertos do grupo superior e a proporção de acertos do grupo inferior:

$$\text{Disc} = P_{sup} - P_{inf}$$

Conforme Rabelo (2013) utiliza-se a classificação a seguir para determinar o quanto um item é discriminativo. Pode-se dizer que, quanto maior a diferença entre  $P_{sup}$  e  $P_{inf}$ , maior será o poder de discriminação do item.



Tabela : classificação dos itens de acordo com a discriminação na TCT

VALORES DE DISCRIMINAÇÃO	CLASSIFICAÇÃO
$Disc < 20$	Item deficiente, que deve ser rejeitado
$20 \leq Disc < 30$	Item marginal, sujeito a reelaboração
$30 \leq Disc < 40$	Item bom, mas sujeito a aprimoramento
$Disc \geq 40$	Item bom

Fonte: Rabelo (2013, p. 136).

Ainda na TCT temos outra medida que também é muito utilizada para verificar o poder de discriminação de um item, o coeficiente de correlação ponto-bisserial. Ele compara o escore total dos indivíduos no teste com o escore total dos indivíduos que acertaram um item em particular, isto é, o número de acertos do item na prova. Variando no intervalo  $[-1, 1]$ , espera-se que o ponto-bisserial apresente coeficiente maior que 0,30, pois valores negativos ou próximos de zero indicam que indivíduos com um bom desempenho total no teste estão dando respostas incorretas. Isso mostra que o item não está cumprindo o seu papel de distinguir alunos com uma grande habilidade dos alunos com uma baixa habilidade:

$$\rho_{pb} = \frac{S_p - S}{\sigma} \sqrt{\frac{p}{1-p}},$$

em que,  $S_p$  é a nota média no teste para os indivíduos que acertaram o item;  $S$  é a nota média no teste para todos os indivíduos;  $\sigma$  é o desvio-padrão das notas obtidas no teste por todos os indivíduos;  $p$  é a proporção de acertos no item.

Esse coeficiente pode ser calculado também para cada uma das opções de respostas da questão, determinando, assim, uma medida de correlação entre o escore total no teste e o escore dos respondentes que marcaram cada uma das diferentes alternativas de respostas de um único item. Para isso, basta substituir, na equação anterior, a nota média no teste para os indivíduos que acertaram o item pela nota média no teste dos indivíduos que marcaram o item.

Observando a equação anterior, é fácil notar que a expressão  $S_p - S$  determinará se o ponto-bisserial será positivo ou não. Dessa forma, anseia-se que a opção correta de resposta tenha o coeficiente ponto-bisserial positivo (média dos indivíduos que acertaram o item maior que a média de todos os indivíduos que responderam a prova) e as demais opções tenham coeficientes negativos (acontecendo o inverso do caso anterior). Se a alternativa correta tiver o coeficiente negativo e ou alguma alternativa errada tiver o coeficiente positivo, isso pode ser um indicativo de algum problema na construção das respostas desse item ou na construção do próprio item, mostrando, assim, que o referido item não discrimina bem e que deve ser descartado de um teste de avaliação educacional.

### 3.1.5 O Parâmetro de Acerto ao Acaso (c)

Este parâmetro representa a probabilidade de um indivíduo com baixa proficiência acertar casualmente um item, utilizando o popular “chute”. É razoável pensarmos que num teste com itens de múltipla escolha, cada uma das  $p$  alternativas tem probabilidade de  $\frac{1}{p}$  de ser escolhida, logo, se o acerto ao acaso for superior a este valor, significa que a resposta correta atrai alunos com pouca habilidade por ser diferente das demais opções de resposta de alguma forma. Itens onde esse fenômeno acontece, em geral, foram mal elaborados ou tiveram suas alternativas de resposta mal elaboradas.

Na grande maioria dos testes de avaliação educacional, os itens são criados com 4 ou 5 opções de resposta, definindo o parâmetro “c” ideal de, no máximo, 0,25 ou 0,20, respectivamente.

Introduzindo o último parâmetro na equação

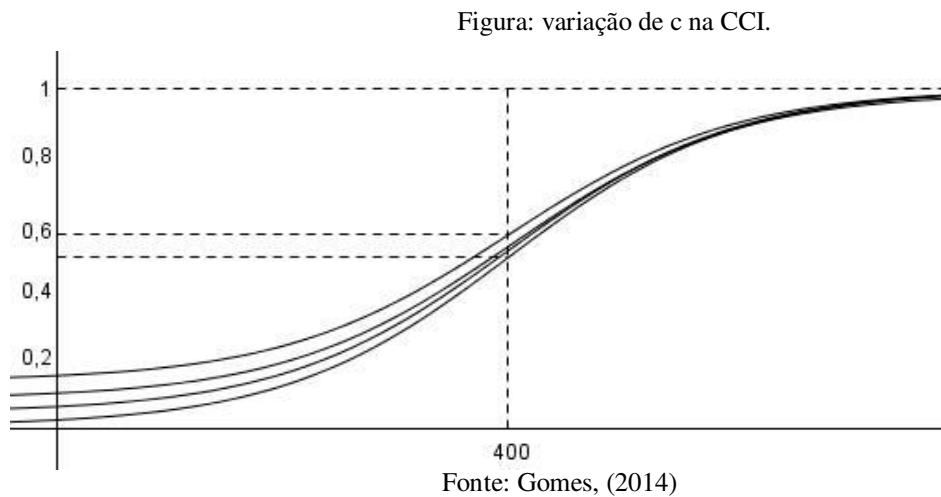
$$P(U_{ij} = 1 | \theta_j) = \frac{1}{1 + e^{-a_i(\theta_j - b_i)}}$$

chegaremos à definição completa do modelo logístico de três parâmetros, conforme essa equação

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

onde  $c_i$  representa a probabilidade de acerto ao acaso do item  $i$ ,  $D$  é um fator de escala igual a 1 quando utilizamos a métrica logística e igual a 1,7 quando utilizamos a métrica normal.

Observando a CCI, vemos que a assíntota horizontal inferior da curva, intercepta o eixo vertical exatamente no parâmetro “ $c$ ”.



### 3.2 Análise Pedagógica do Item

Até o momento, procuramos explorar os principais conceitos que envolvem a TRI e as diferenças em relação aos conceitos análogos na teoria clássica. Entretanto, em termos de análise pedagógica dos itens, as duas teorias se completam, permitindo um diagnóstico muito mais amplo sobre o nível de proficiência dos alunos e o processo de construção dos itens. Salientamos que algumas características e classificações dos parâmetros da TRI e medidas da TCT são consideradas ideais para itens que servirão para processos de avaliações educacionais, entretanto a prova de acesso ao PROFMAT não se propõe a essa finalidade e sim a aferir os conhecimentos matemáticos necessários para os futuros alunos cursarem com êxito as disciplinas inerentes ao PROFMAT.

Neste trabalho, cada uma das seis primeiras questões do exame nacional de acesso do ano de 2018 será pedagogicamente avaliada. Será exibida uma tabela onde serão apresentados dados estatísticos da questão (número de respondentes e porcentagem de acertos), dados da TCT (frequência de marcação de cada alternativa e os seus respectivos bisseriais) e os parâmetros da TRI ( $a$ ,  $b$  e  $c$ ). Serão expostos ainda dois gráficos, a CCI e a AGI. É importante destacar que os gráficos da CCI serão construídos utilizando o programa BILOG que utiliza a escala de proficiência (0,1). Entretanto, a análise dos itens utilizando a TRI será feita na escala (500,

100), sem nenhum prejuízo na interpretação, pois em ambas as escalas os gráficos são semelhantes.

A seguir, temos um exemplo de um tipo de tabela que pode ser utilizada e um breve resumo sobre as informações nela contidas.

Gabarito:					
Assunto:					
TCT	Total:		Acertos:		Percentual de acertos:
Opções	A	B	C	D	E
Frequência					
Bisserial					
TRI	a		b		c

- **Gabarito:** indica a resposta correta do item;
- **Assunto:** apresenta o principal conteúdo abordado na questão;
- **Total:** é a quantidade de respondentes ao teste;
- **Acertos:** é a quantidade de respondentes que acertaram o teste;
- **Percentual de Acertos:** indica a dificuldade do item medido pela TCT;
- **Frequência:** é a quantidade de indivíduos que marcaram essa alternativa com resposta correta;
- **Bisserial:** estabelece uma relação entre a nota média de todos os indivíduos no teste e a nota média dos indivíduos que marcaram essa alternativa como resposta correta. Espera-se valor positivo para a opção correta e valores negativos para as demais opções;
- **a:** indica o poder de discriminação do item pela TRI;
- **b:** é o grau de dificuldade do item pela TRI;
- **c:** é a probabilidade de acerto ao acaso.

## **4 UMA APLICAÇÃO DA TRI NO EXAME NACIONAL DE ACESSO (ENA) DO PROFMAT UFU 2018**

O exame nacional de acesso ao PROFMAT UFU 2018(<https://profmat-sbm.org.br/>) foi realizado por 125 candidatos e composto por 30 questões de múltipla escolha, teve duração máxima de 4 horas e avaliou os seguintes itens:

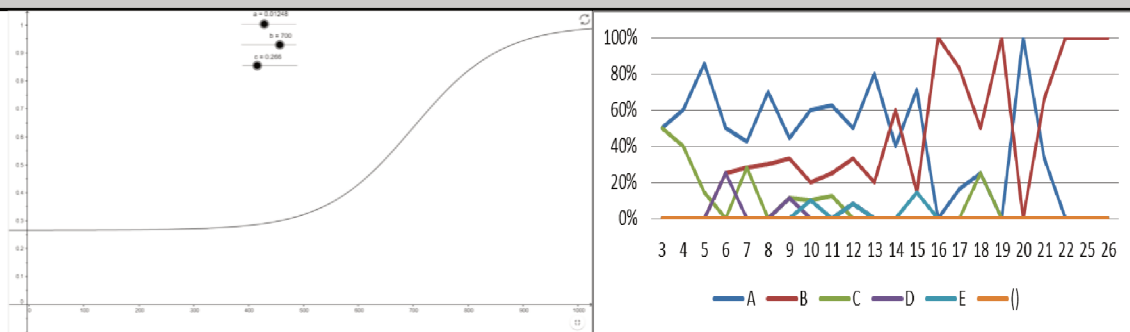
- a) Construção de significados para os números;
- b) O conhecimento geométrico e a realidade;
- c) Grandezas e medidas e resolução de problemas do cotidiano;
- d) Variações de grandezas;
- e) Resolução de problemas algébricos;
- f) Organização de dados e tratamento da informação.

Os parâmetros a, b e c da TRI relativos ao exame nacional de acesso foram calculados utilizando o programa BILOG, diretamente sobre os resultados obtidos da prova, sem fazer a pré-testagem dos itens. Os valores dos parâmetros e as estatísticas da prova foram cedidos pelo orientador deste trabalho. Vale destacar que uma amostra de 125 indivíduos é pequena, sendo possível que os dados obtidos sofram algumas tendenciosidades por “poucos” indivíduos com resultados próximos aos extremos de cada parâmetro.

[01] Se  $\ell$  é o lado e  $A$  é a área de um triângulo equilátero, então é correto afirmar que:

- (A)  $A$  e  $\ell$  são grandezas diretamente proporcionais.
- (B)  $A$  e  $\ell^2$  são grandezas diretamente proporcionais.
- (C)  $A$  e  $\ell$  são grandezas inversamente proporcionais.
- (D)  $A^2$  e  $\ell$  são grandezas inversamente proporcionais.
- (E)  $A^2$  e  $\ell^2$  são grandezas inversamente proporcionais.

Questão	Gabarito	Total	Acertos	Dificuldade	I <sub>D</sub>	Bisserial
1	B	125	45	0,36	0,4848	0,5549
Acerto: 27% maiores notas			Acerto: 27% menores notas			
0,6363			0,1515			
Opções	A	B	C	D	E	()
Frequência	0,512	0,36	0,08	0,024	0,024	0
Bisserial	-0,33	0,5549	-0,49	-0,23	0,053	0
TRI	a	1,2479	b	700	c	0,2659

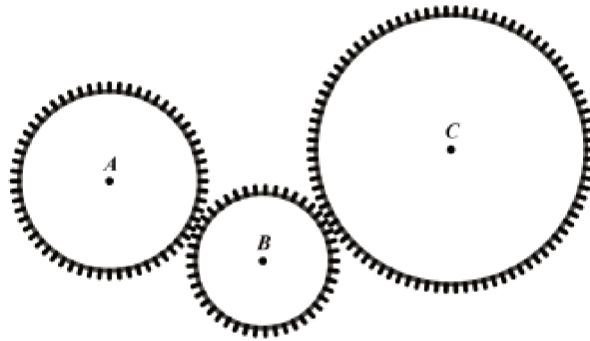


Utilizando os dados obtidos pela AGI, é possível afirmar que este exercício apresenta uma dificuldade moderada e discrimina bem os respondentes com alto desempenho dos de baixo desempenho. Apesar disto, esta análise deixa claro ainda que a alternativa A possui algum tipo de atrativo, que fica claro no gráfico, além da alternativa E possuir bisserial positivo, indicando que indivíduos de bom desempenho a escolheram.

Pela análise da CCI, o item apresenta uma dificuldade muito alta e é moderada quanto à discriminação. É possível inferir ainda que a atratividade da alternativa correta esta dentro do esperado.

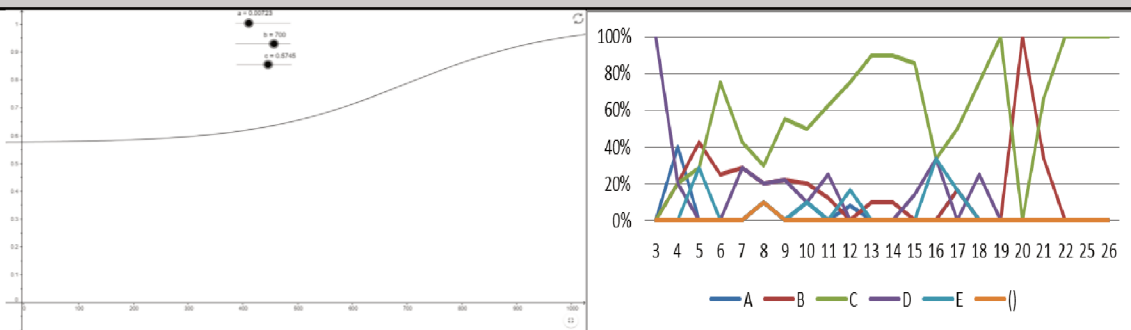
Nesta breve análise identificamos uma pequena divergência quanto aos resultados obtidos de cada um, principalmente na dificuldade do item.

[02] Na figura, estão representadas as engrenagens A, B e C, que possuem, respectivamente, 60, 45 e 90 dentes cada uma. Quantas voltas completas dará a engrenagem A se a engrenagem C der 4 voltas completas e mais  $\frac{2}{3}$  de volta?



(A) 5    (B) 6    (C) 7    (D) 8    (E) 9

Questão	Gabarito	Total	Acertos	Dificuldade	I <sub>D</sub>	Bisserial
2	C	125	76	0,608	0,3636	0,416
Acerto: 27% maiores notas			Acerto: 27% menores notas			
0,7272			0,3636			
Opções	A	B	C	D	E	()
Frequência	0,048	0,152	0,608	0,12	0,064	0,008
Bisserial	-0,25	-0,233	0,416	-0,3	-0,114	-0,179
TRI	a	0,7234	b	700	c	0,5744



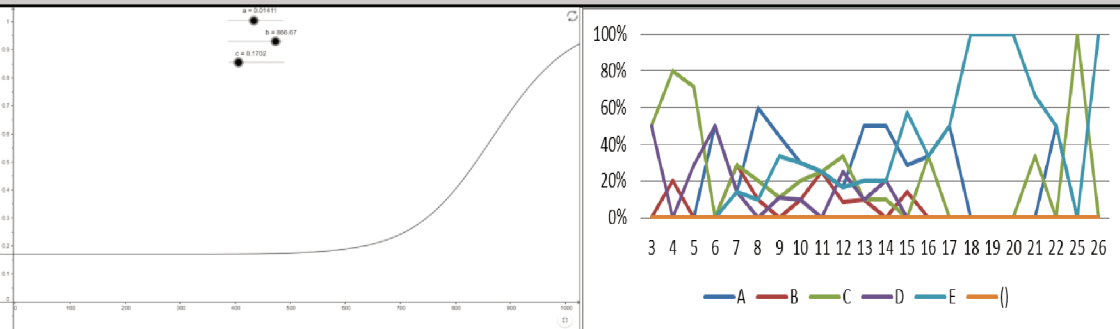
A AGI denota que este item, apesar de bom, está sujeito a ser melhorado para futuras avaliações. A dificuldade é moderada e a capacidade de separar os alunos é boa.

A CCI indica que este item discrimina moderadamente, além de contrastar denotando uma dificuldade muito alta. Por fim, é possível entender ainda que este item teve muitos acertos ao acaso (“chutes”), ou seja, a alternativa correta se destaca entre todas.

[03] Qual é a soma dos valores de  $k \in \mathbb{R}$  para os quais a equação  $\frac{k}{x^2 - 4} + \frac{1}{x + 2} = 0$  não possui solução real?

- (A) 0 (B) 1 (C) 2 (D) 3 (E) 4

Questão	Gabarito	Total	Acertos	Dificuldade	I <sub>D</sub>	Bisserial
3	E	125	35	0,28	0,5151	0,5403
Acerto: 27% maiores notas			Acerto: 27% menores notas			
0,5757			0,0606			
Opções	A	B	C	D	E	()
Frequência	0,304	0,08	0,224	0,112	0,28	0
Bisserial	0,045	-0,21	-0,36	-0,31	0,5403	0
TRI	a	1,4106	b	866,67	c	0,1702



O item é considerado difícil pela AGI, apesar deste discriminar bem os avaliados, existe um possível elemento chamativo nas alternativas A e C.

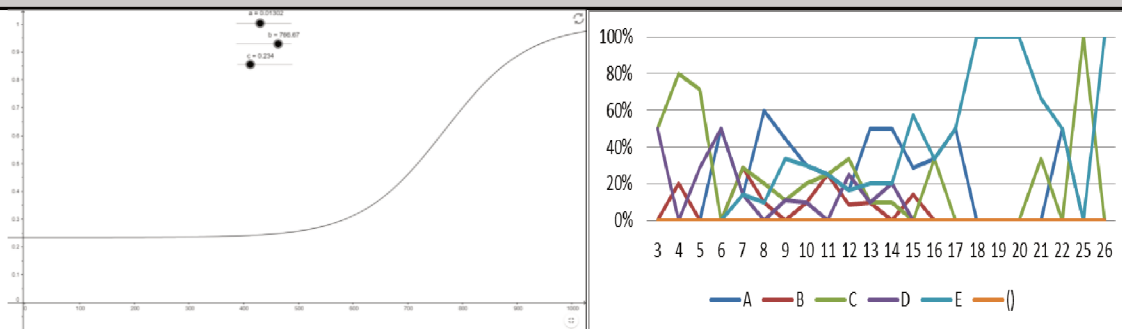
As informações obtidas da CCI dão o indício de uma possível alternativa mais atrativa que a correta, já que o acerto ao acaso da alternativa E, apesar de estar dentro do esperado, é relativamente baixo. Conclui-se também que o item apresenta um índice de discriminação moderado



[04] Quantas raízes reais possui a equação  $3 + \sqrt{3 + x^4} = x^2$  ?

(A) 0 (B) 1 (C) 2 (D) 3 (E) 4

Questão	Gabarito	Total	Acertos	Dificuldade	I <sub>D</sub>	Bisserial
4	A	125	33	0,264	0,1515	0,25
Acerto: 27% maiores notas			Acerto: 27% menores notas			
0,3333			0,1818			
Opções	A	B	C	D	E	()
Frequência	0,264	0,072	0,528	0,032	0,104	0
Bisserial	0,25	-0,089	0,043	-0,29	-0,364	0
TRI	a	1,3021	b	766,67	c	0,234



Esta questão é, segundo a AGI, difícil e deficiente, pois não apresenta acertos condizentes com as habilidades dos respondentes e possivelmente será descartada para futuros usos. Outro problema perceptível é a atratividade da alternativa A.

Confirmando as informações anteriores, os dados da CCI indicam uma discriminação baixa da questão e um elevado nível de dificuldade.

[05] Considere as asserções abaixo e a relação proposta entre elas.

I. A soma e o produto das raízes da equação  $2x^2 + 5x + 2 = 0$  são  $-5$  e  $2$ , respectivamente.

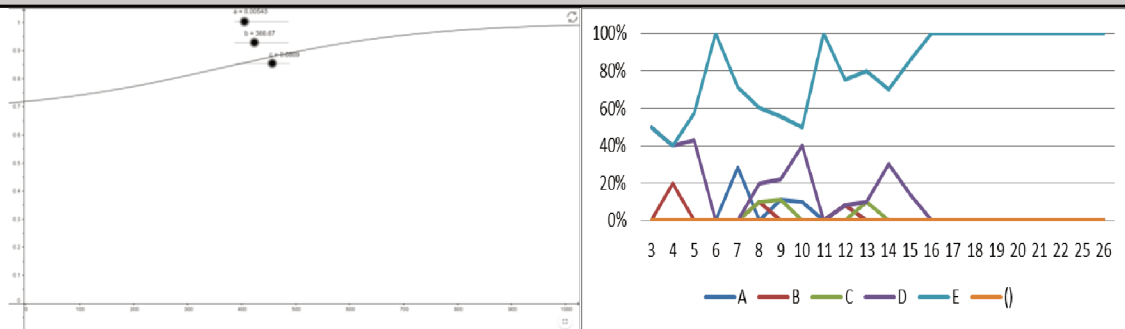
**PORQUE**

II. Se  $s$  e  $p$  são, respectivamente, a soma e o produto das raízes da equação  $ax^2 + bx + c = 0$ , com  $a, b, c \in \mathbb{R}$  e  $a \neq 0$ , então  $s = -b$  e  $p = c$ .

A respeito dessas asserções, assinale a opção correta.

- (A) As asserções I e II são proposições verdadeiras, e a II é uma justificativa da I.
- (B) As asserções I e II são proposições verdadeiras, mas a II não é uma justificativa correta da I.
- (C) A asserção I é uma proposição verdadeira, e a II é uma proposição falsa.
- (D) A asserção I é uma proposição falsa, e a II é uma proposição verdadeira.
- (E) As asserções I e II são proposições falsas.

Questão	Gabarito	Total	Acertos	Dificuldade	I <sub>D</sub>	Bisserial
5	E	125	94	0,752	0,3333	0,4233
Acerto: 27% maiores notas			Acerto: 27% menores notas			
0,9393			0,6060			
Opções	A	B	C	D	E	()
Frequência	0,04	0,024	0,024	0,16	0,752	0
Bisserial	-0,26	-0,321	-0,15	-0,35	0,4233	0
TRI	a	0,5426	b	366,67	c	0,6808

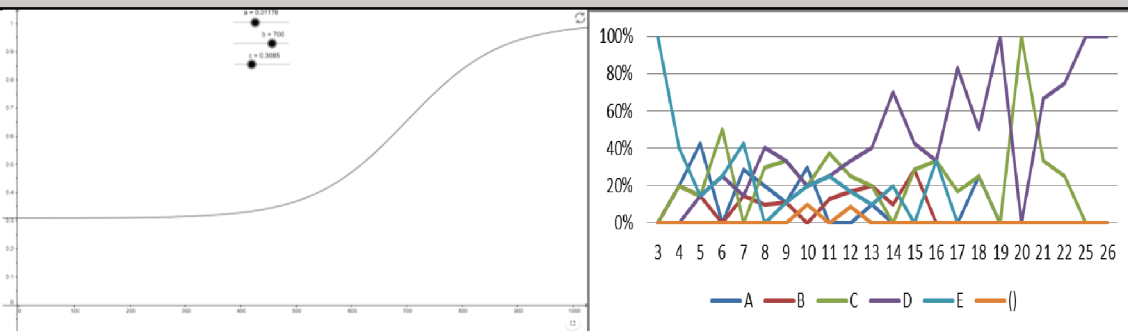


Ambas as análises indicam que esta questão é fácil ou muito fácil. Apesar de ser um bom item, ele precisa de algum tipo de melhoria, provavelmente relacionada ao destaque que existe na alternativa E, que possui um índice de acerto ao acaso realmente alto, ou seja, é fácil acertar a questão mesmo não tendo o conhecimento necessário para respondê-la.

[06] Que número inteiro pode ser escrito como  $\sqrt{19 + 6\sqrt{10}} - \sqrt{19 - 6\sqrt{10}}$  ?

(A) 3    (B) 4    (C) 5    (D) 6    (E) 7

Questão	Gabarito	Total	Acertos	Dificuldade	I <sub>D</sub>	Bisserial
6	D	125	48	0,384	0,4242	0,5328
Acerto: 27% maiores notas			Acerto: 27% menores notas			
0,6060			0,1818			
Opções	A	B	C	D	E	()
Frequência	0,112	0,104	0,224	0,384	0,16	0,016
Bisserial	-0,5743	-0,1932	0,0207	0,5328	-0,5407	-0,0798
TRI	a	0,9715	b	666,6667	c	0,3085



É claro uma divergência da AGI com si mesma quanto a esta questão, pois indica que é um bom item, de dificuldade moderada e que discrimina bem, apesar de deixar claro pelo bisserial que a alternativa C é de alguma forma atrativa. Já a CCI informa que a chance de um respondente marcar corretamente chutando é alta. Indica ainda que é muito difícil.

## 5 CONSIDERAÇÕES FINAIS

O estudo da Teoria de Resposta ao Item mostrou-se de grande importância, pois além de apresentar os conceitos que fizeram esta teoria ser adotada nos principais exames de avaliação educacional, proporcionou um maior conhecimento sobre a Teoria Clássica dos Testes, muitas vezes empregada pelos professores em suas avaliações, mas sem ter suas principais medidas utilizadas como forma de mensurar o resultado da avaliação como um todo. Vale destacar que o professor que conhecer os principais índices da TRI e da TCT, torna-se capaz de elaborar avaliações mais adequadas a finalidade a que se destina.

A análise parcial do exame nacional de acesso ao PROFMAT 2018 permitiu na prática, acompanhar como os resultados de um grande processo de seleção pode colaborar para melhor prepararmos os futuros professores e ajudar na formação continuada dos docentes já formados. Assim, mostra-se possível elaborarmos um plano de ação almejando trabalhar os conteúdos e habilidades que os candidatos mais apresentam dificuldades. Em particular, no caso de uma seleção de professores, essas ações podem melhorar, consideravelmente, a qualidade das aulas para os alunos desses professores, contribuindo assim, com uma melhor formação para os principais sujeitos envolvidos no processo ensino-aprendizagem.

## 6 REFERÊNCIAS

ANDRADE, D. F. ; VALLE, Raquel da Cunha; TAVARES, Heliton Ribeiro. *Introdução a teoria da resposta ao ítem: conceitos e aplicações*. SINAPE, 2000. Disponível em: <<http://www.avaliaeducacional.com.br/referencias/arquivos/LivroTRI%20-%20Dalton.pdf>>. Acesso em: Out. 2013.

ARAUJO, E. A. C. ; ANDRADE, D. F. ; BORTOLOTTI, Silvana Ligia Vincenzi . *Teoria da Resposta ao Item*. Revista da Escola de Enfermagem da USP (Impresso), v. 43, p. 1000-1008, 2009. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0080-62342009000500003&lang=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0080-62342009000500003&lang=pt)>. Acesso em: Out. 2013.

BARNHART, Ryan. *Fumiko Samejima (n. 1930)*. Disponível em: <<http://www.apadivisions.org/division-35/about/heritage/fumiko-samejima-biography.aspx>>. Acesso em: Fev. 2014.

GOMES, S. L. *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 2012*. Rio de Janeiro: IMPA, 2014.

INEP. *Entenda a sua nota no Enem – Guia do Participante*. Disponível em: <[http://download.inep.gov.br/educacao\\_basica/enem/guia\\_participante/2012/guia\\_do\\_participante\\_notas\\_2012.pdf](http://download.inep.gov.br/educacao_basica/enem/guia_participante/2012/guia_do_participante_notas_2012.pdf)>. Acesso em: Abr. 2014.

INEP. *Procedimento de cálculo das notas do Enem*. Disponível em: <[http://download.inep.gov.br/educacao\\_basica/enem/nota\\_tecnica/2011/nota\\_tecnica\\_procedimento\\_de\\_calculo\\_das\\_notas\\_enem\\_2.pdf](http://download.inep.gov.br/educacao_basica/enem/nota_tecnica/2011/nota_tecnica_procedimento_de_calculo_das_notas_enem_2.pdf)>. Acesso em: Out. 2013.

KLEIN, Ruben. *Alguns aspectos da teoria de resposta ao item relativos à estimação das proficiências*. Ensaio (Fundação Cesgranrio. Impresso), v. 21, p. 35-56, 2013. Disponível em: <[http://www.scielo.br/pdf/ensaio/v21n78/aop\\_0213.pdf](http://www.scielo.br/pdf/ensaio/v21n78/aop_0213.pdf)>. Acesso em: Out. 2013

MACHADO, Ledo Vaccaro. *Avaliação em larga escala e proficiência matemática*. Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Ensino de Matemática, UFRJ, 2010. Disponível em: <<http://www.pg.im.ufrj.br/pemat/26%20Ledo%20Vaccaro.pdf>>. Acesso em: Nov. 2013.

RABELO, Mauro. *Avaliação educacional: fundamentos, metodologia e aplicações no contexto brasileiro*. Rio de Janeiro: SBM, 2013.