



UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
Programa de Pós-Graduação em Matemática
Mestrado Profissional - PROFMAT/CCT/UFCG



Teoria de Resposta ao Item como Modelo de Avaliação de Aprendizagem de Matemática no Ensino Médio

Priscila Cordeiro de Sousa

Trabalho de Conclusão de Curso

Orientador: Prof. Dr. Luiz Antônio da Silva Medeiros

Orientador: Prof. Dr. Gilberto da Silva Matos

Campina Grande - PB

Outubro/2018

S725t

Sousa, Priscila Cordeiro de.

Teoria de resposta o item como modelo de avaliação de aprendizagem de matemática no ensino médio / Priscila Cordeiro de Sousa. - Campina Grande, 2018.

85 f. : il. color.

Dissertação (Mestrado Profissional em Matemática) - Universidade Federal de Campina Grande, Centro de Ciências e Tecnologia, 2018.

"Orientação: Prof. Dr. Luiz Antônio da Silva Medeiros, Prof. Dr. Gilberto da Silva Matos".

Referências.

1. Psicometria. 2. Teoria de Resposta ao Item (TRI). 3. Avaliação da Proficiência dos Alunos. 4. Teoria Clássica dos Testes (TCT). 5. Ferramentas de Avaliação da Aprendizagem. I. Medeiros, Luiz Antônio da Silva. II. Matos, Gilberto da Silva. III. Título.

CDU 37.091.279.5:51(043)

FICHA CATALOGRÁFICA ELABORADA PELO BIBLIOTECÁRIO GUSTAVO DINIZ DO NASCIMENTO CRB - 15/515

Priscila Cordeiro de Sousa

Teoria de Resposta ao Item como Modelo de Avaliação de Aprendizagem de Matemática no Ensino Médio

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática, junto ao Programa de Pós-graduação Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, da Universidade Federal de Campina Grande - UFCG.

Orientador: Prof. Dr. Luiz Antônio da Silva Medeiros

Orientador: Prof. Dr. Gilberto da Silva Matos

Campina Grande - PB

Outubro/2018

Teoria de Resposta ao Item como modelo de avaliação de aprendizagem de Matemática no Ensino Médio

por

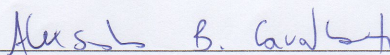
Priscila Cordeiro de Sousa

Trabalho de Conclusão de Curso apresentado ao Corpo Docente do Programa de Pós-Graduação em Matemática - CCT - UFCG, modalidade Mestrado Profissional, como requisito parcial para obtenção do título de Mestre.

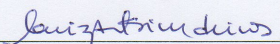
Aprovado por:



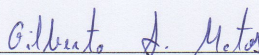
Prof. Dr. Cícero Carlos Felix de Oliveira - IFCE



Prof. Dr. Alessandro Bezerra Cavalcanti - UFCG



Prof. Dr. Luiz Antônio da Silva Medeiros - UFCG
Orientador



Prof. Dr. Gilberto da Silva Matos - UFCG
Co-Orientador

Universidade Federal de Campina Grande
Centro de Ciências e Tecnologia
Unidade Acadêmica de Matemática
Curso de Mestrado Profissional em Matemática em Rede Nacional

Outubro/2018

Dedico este trabalho aos meus pais que me oportunizaram uma educação de qualidade, digna e significativa. E aos meus filhos, que sofreram os momentos de ausência. Amo muito vocês.

Agradecimentos

À Deus que com seu imenso amor fez-me forte e confiante na realização deste sonho.

À minha família, que sofreu todas as minhas angústias, estresses e ausências, sempre me motivando a dar o meu melhor.

Aos meus amigos, em especial a Sandra, por me confortarem nos momentos de dúvida e por me motivarem a acreditar na minha capacidade de superação.

Aos meus queridos amigos de turma do curso PROFMAT, pois sem as suas colaborações esse sonho não seria concretizado. Cada um teve uma grande importância, pois a nossa união nos possibilitou superar as grandes dificuldades ao longo do curso. Sou grata!

Aos meus orientadores, Prof. Dr. Luiz Antônio da Silva Medeiros e Prof. Dr. Gilberto Da Silva Matos, pela paciência, incentivo e contribuições para a realização desse trabalho.

Aos gestores, colegas de trabalho e queridos alunos das escolas que leciono por me incentivarem e compreenderem as ausências necessárias.

À SBM e a UFCG, por oportunizarem o curso PROFMAT, possibilitando ao docente de matemática qualificação profissional.

À Capes pelo auxílio financeiro durante o curso.

Aos professores do curso, pelo conhecimento compartilhado, os incentivos e as colaborações na minha formação profissional.

Faça o teu melhor, na condição que você tem,
enquanto você não tem condições melhores,
para fazer melhor ainda!

Mario Sergio Cortella

Resumo

Com o advento dos recursos computacionais, avanços tecnológicos e uso de novos instrumentos de medidas em exames para entrada em universidades (como o Exame Nacional do Ensino Médio - ENEM), surge a motivação em prepararmos o público - alunos do Ensino Médio - que sucederão estes exames. A Teoria de Resposta ao Item (TRI) vem sendo considerada o mecanismo mais viável de avaliar a proficiência dos alunos, devido a sua preocupação quanto ao item e o traço latente (habilidades) dos alunos, no entanto a maioria das instituições de ensino ainda usam a Teoria Clássica de Testes (TCT) como principal ferramenta de avaliação da aprendizagem, devido a sua praticidade. O uso da TCT é questionável, pois sua avaliação não é fidedigna, podendo o escore variar de teste para teste. Desse modo, propomos a construção de um instrumento de verificação de aprendizagem a partir do guia de Elaboração e Revisão de Itens do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), aplicando a TRI em duas turmas de 3º ano do Ensino Médio.

Palavras Chaves: Avaliação. Habilidade. Item.

Abstract

With the advent of computational devices, technological advances and the use of new assessment instruments in entrance exams at universities (such as the National High School Examination - ENEM), the motivation to prepare the public - high school students - will succeed these exams. The Item Response Theory (TRI) is considered to be the most feasible mechanism for assessing student proficiency due to their concern about the item and the latent ability (skills) of the students, however most schools still use the Theory Classical Testing (TCT) as the primary assessment tool learning due to its practicality. The use of TCT is questionable because its assessment is not reliable, and the score can vary from test to test. In this way, we propose the construction of a learning verification tool from the INEP Elaboration and Review of Items guide, applying the TRI two classes of the last year of High School.

Keywords: Evaluation. Skill. Rasch.

Lista de Siglas

ABNT	Associação Brasileira de Normas Técnicas
BNI	Banco Nacional de Itens
CCI	Curva Característica do Item
ENEM	Exame Nacional do Ensino Médio
FRI	Função de Resposta ao Item
INEP	Instituto Nacional de Estudo e Pesquisas Educacionais Anísio Teixeira
TCT	Teoria Clássica dos Testes
TRI	Teoria de Resposta ao Item

Sumário

1	Introdução	3
1.1	Objetivos	5
2	Psicometria clássica e TRI	7
2.1	Psicometria clássica	7
2.1.1	Modelo de medição de um respondente	8
2.2	Teoria de Resposta ao Item	11
2.2.1	Modelos para itens dicotômicos ou dicotomizados	12
2.3	Modelos envolvendo duas ou mais populações	21
2.3.1	Estimação dos parâmetros	21
3	O uso do R na TCT e na TRI	23
3.1	Aplicações do R na TCT	23
3.1.1	Coefficiente de correlação ponto-bisserial	23
3.1.2	Coefficiente de correlação bisserial	24
3.1.3	Coefficiente alfa de Cronbach	24
3.2	Aplicações do R na TRI	25
3.3	O R e a psicometria	25
4	Itens - Elaboração e revisão de itens	26
4.1	Elaboração do item de múltipla escolha	27
4.1.1	Quanto ao texto-base	28
4.1.2	Quanto ao enunciado	31
4.1.3	Quanto as alternativas	32
4.2	Análise de itens	33
4.2.1	(Item 161- ENEM 2017 - caderno 6 - prova cinza)	33
4.2.2	(Item 173- ENEM 2017 - caderno 6 - prova cinza)	34
4.2.3	(Item 164- ENEM 2017 - caderno 6 - prova cinza)	35
4.2.4	(Item 162- ENEM 2017 - caderno 6 - prova cinza)	36
4.2.5	((Item 180- ENEM 2017 - caderno 6 - prova cinza)	37

5	Itens elaborados para calibração	39
5.1	Itens Elaborados	39
5.1.1	Item 1	39
5.1.2	Item 2	40
5.1.3	Item 3	41
5.1.4	Item 4	42
5.1.5	Item 5	43
5.1.6	Item 6	43
5.1.7	Item 7	44
5.1.8	Item 8	45
5.1.9	Item 9	46
5.1.10	Item 10	46
5.1.11	Item 11	47
5.1.12	Item 12	48
5.1.13	Item 13	48
5.1.14	Item 14	49
5.1.15	Item 15	50
5.1.16	Item 16	51
5.1.17	Item 17	52
5.1.18	Item 18	53
5.1.19	Item 19	54
5.1.20	Item 20	54
5.1.21	Pesquisa exploratória	55
6	Análise dos itens e do teste usando o Ambiente Computacional e Estatístico R	57
6.1	Leitura e visualização breve da base de dados	59
6.2	Análise Clássica dos Itens e do Teste	60
6.2.1	Instalação do pacote "ltm" e análise pela TCT	60
6.3	Estimação Bayesiana dos Itens e das Habilidades - TRI	65
6.3.1	Instalação do pacote "MCMCpack" e estimação bayesiana via métodos MCMC	65
6.3.2	Comparando os resultados da TCT e da TRI	71
7	Conclusões	78
	Referências Bibliográficas	80
A	Primeiro Apêndice	81

Capítulo 1

Introdução

Corriqueiramente, vivencia-se situações do tipo: “Que caminho vou seguir para ir à farmácia?”; “Qual roupa visto para ir a entrevista de emprego?”; “Qual das propostas de venda daquele terreno é mais interessante?”. O ato de avaliar está presente no nosso dia-a-dia, naturalmente. Quando toma-se decisões, investiga-se, compara-se, classifica-se, definindo critérios e assim fazendo escolhas, mediante os parâmetros estabelecidos.

A avaliação é um tema frequentemente abordado nas rodas de conversa de professores, assim como nos cursos de graduação, devido à complexidade da avaliação educacional. O que avaliar? Como avaliar? Que instrumentos utilizar? As respostas para esses e outros questionamentos sobre avaliação, tem sido objeto de estudo para esse e muitos outros trabalhos de pesquisa, cujo propósito é melhorar a prática dos professores e minimizar os prejuízos de uma avaliação tradicional, classificatória e punitiva.

A maioria das instituições de ensino usam a TCT nos procedimentos avaliativos, por ser rápida e eficiente, facilitando o trabalho dos examinadores. O principal instrumento avaliativo consiste numa prova, em que a quantidade de acertos ou a média ponderada dos mesmos, determina o escore total. Na TCT, quanto mais questões acertadas, maior domínio sobre os conteúdos avaliados. Essa teoria possibilita a comparação entre alunos que fizeram a mesma prova, classificando-os, porém, não é possível compará-los se forem submetidos a provas diferentes.

Segundo Luckesi avaliar é:

Avaliar é o ato de diagnosticar uma experiência, tendo em vista reorientá-la para produzir o melhor resultado possível; por isso, não é classificatória nem seletiva, ao contrário, é diagnóstica e inclusiva. O ato de examinar, por outro lado, é classificatório e seletivo e, por isso mesmo, excludente, já que não se destina à construção do melhor resultado possível; tem a ver, sim, com a classificação estática do que é examinado. O ato de avaliar tem seu foco na construção dos melhores resultados possíveis, enquanto o ato de examinar está centrado no julgamento de aprovação ou reprovação. Por suas características e modos de ser, são atos praticamente opostos; no entanto, professores e professoras, em sua prática escolar cotidiana, não fazem essa distinção e, deste modo, praticam exames como se estivessem praticando avaliação. (LUCKESI, 2002, p. 90)

Portanto, o aluno é examinado em vez de ser avaliado. Uma avaliação quantitativa, que não permite ao docente o acompanhamento da construção da aprendizagem, já que habilidades e competências passam despercebidas, pois apenas se observa o que ele, o aprendiz, foi capaz de reproduzir.

Segundo Rabelo (1998), em uma proposta de avaliação, o importante não são apenas as respostas certas ou erradas que os alunos apresentam, mas a forma de como os alunos encontram as respostas também deve ser valorizada, pois esta é uma parte importante da aprendizagem. Quando o docente se propõe a observar a maneira que o aluno lida com as questões e como se posiciona diante delas, o foco da avaliação passa a ser o desenvolvimento dele no processo de construção do conhecimento, priorizando a qualidade do processo de ensino-aprendizagem.

Avaliar qualitativamente representa um grande desafio no processo educativo, neste sentido a TRI vem se destacando, como a ferramenta mais conveniente, por se distanciar da memorização e reprodução de conteúdo. A TRI avalia o aluno quanto as habilidades e competências adquiridas no decorrer da construção da aprendizagem, ou seja, quanto maior sua habilidade, maior será a probabilidade de acertar. Com ela é possível identificar os sucessos e deficiências significativas do caminho da aprendizagem.

Apesar de toda a importância da TRI, a mesma ainda não é muito conhecida e utilizada, devido à dificuldade de interpretação dos resultados, necessitando de avançados recursos tecnológicos para a realização da análise estatística. Esta teoria surgiu em meados do século 20, no entanto aqui no Brasil só ganhou atenção a pouco tempo, quando passou a estar presente nos principais instrumentos avaliativos do país como, por exemplo, o ENEM e o Sistema de Avaliação da Educação Básica (Saeb). A fim de proporcionar o uso da TRI no ensino médio se faz necessário instrumentos e alternativas que supram a complexidade em

utilizá-la, proporcionando aos avaliadores o uso mais frequente da TRI.

Com vista a uma aproximação do ensino básico com a TRI, propomos e elaboramos 20 itens que foram aplicados em turmas do Ensino Médio como simulado e avaliado segundo a TRI.

Esse trabalho está estruturado da seguinte forma: introdução (Capítulo 1), onde apresentam a motivação para a realização desse estudo, destacando as inquietações que o desencadearam, os objetivos e a metodologia utilizada; as considerações finais (Capítulo 7) e mais cinco capítulos relacionados a seguir:

No Capítulo 2, apresenta os conceitos e as características da TCT e da TRI, diferenciando-as. Também é feito um breve estudo sobre a TRI e seus modelos.

No Capítulo 3, destaca os principais procedimentos e métodos de calibração de itens e de verificação da qualidade de um instrumento avaliativo, na TCT e TRI. Também é apresentado informações sobre o R , sobre aspectos básicos utilizados no estudo.

No Capítulo 4, realiza uma análise criteriosa sobre a importância da elaboração de um item para a TRI, a partir do Guia de Elaboração e Revisão de Itens do INEP. Alguns itens do ENEM 2017 foram analisados segundo os critérios observados.

No Capítulo 5, apresenta vinte itens construídos a partir deste estudo, que constituíram um instrumento avaliativo, simulado, que representa a parte exploratória, sendo aplicado para alunos do 3º ano do Ensino Médio.

No Capítulo 6, aplica-se os modelos da TCT e da TRI - modelo dicotômico unidimensional de dois parâmetros - nos dados obtidos pelo simulado, realizando uma análise dos resultados pelas teorias apresentadas.

1.1 Objetivos

Geral:

- Aplicar a TRI como ferramenta facilitadora de avaliação da aprendizagem em Matemática no Ensino Médio.

Específicos:

- Destacar as diferenças entre a Teoria Clássica dos Testes (TCT) e a Teoria de Resposta ao Item (TRI);

- Aplicar a TRI, usando o modelo de um parâmetro - Rasch, para análise de itens e testes;
- Analisar e propor a construção de itens a partir do Guia de Elaboração e Revisão de Itens do INEP;
- Construir e aplicar um instrumento de medida (prova) a partir do guia de Elaboração e Revisão de Itens;
- Comparar os resultados na TCT e na TRI, usando o software livre *R*, com dados obtidos por respondentes do 3º ano do Ensino Médio.

Capítulo 2

Psicometria clássica e TRI

Facilmente é possível realizar observações sobre o mundo ao nosso redor: registrar a altura de um objeto, ou o seu peso, ou ainda a sua densidade; mas fazer observações sobre o comportamento humano não é tão simples. Frequentemente são aplicados questionários em várias populações, a fim de determinar características dos indivíduos em diferentes níveis, atribuindo uma quantificação a variável psicológica, ou seja, realizar uma mensuração da variável psicológica. Essa medida auxilia em diversas pesquisas tanto na área da psicologia, como em muitas outras, como no campo da saúde e da Educação.

2.1 Psicometria clássica

Quando avalia-se um aluno, realiza-se uma medição acerca do conhecimento adquirido por ele durante o processo de ensino aprendizagem. Querer medir inteligência ou outras variáveis psicológicas, não é tão simples e confiável. A representação numérica de variáveis psicológicas costuma ser determinada aos olhos da Psicometria, área da Psicologia que permite a mensuração de processos psíquicos, através da construção de instrumentos ou testes, aplicando conceitos matemáticos e estatísticos.

A atribuição de um valor a uma variável psicológica, associada a um instrumento avaliativo permite ao especialista realizar comparações entre resultados obtidos, apesar de ser muito discutida a confiabilidade de um resultado obtido. Um instrumento psicológico, para aferir valores em uma amostra ou população, deve apresentar características que vão além da teoria apresentada nos itens do teste (adequação teórica), o nível de alinhamento (coleta empírica), até suas propriedades estatísticas. Essas características estão relacionadas aos conceitos de validade e fidedignidade.

O teste mede realmente o que se espera medir, de forma consistente? O teste será válido se os seus itens e medidas expressarem as perspectivas dos constructos definidos. Se o instrumento apresentar estabilidade e consistência ele é fidedigno. O escore obtido no teste se aproxima do real escore do indivíduo em determinado traço latente, independente

dos fatores estranhos ao constructo, que possam causar uma variação nos escores. Portanto, a fidedignidade depende do conceito de variância do erro, do erro produzido pelo próprio instrumento, ou seja, quanto maior a variância verdadeira e menor a variância erro, mais fidedigno é o instrumento.

Para um instrumento ser considerado válido e confiável, precisa-se de uma medida psicométrica que envolva um conjunto de parâmetros, principalmente os que cercam a análise dos itens (discriminação e dificuldade) e também da validade e da fidedignidade do instrumento. É viável destacar que itens com muitos acertos ou itens com muitos erros, não representam um poder de discriminação.

No modelo psicométrico clássico, a matemática é muito simples. O escore total consiste na soma dos escores brutos ou padronizados dos acertos e erros e dependem em particular do conjunto de itens. O modelo Clássico não permite a comparação entre indivíduos que não foram submetidos aos mesmos instrumentos. O modelo matemático consiste em uma curva do modelo Normal, que se fundamenta no uso da média e do desvio padrão. A curva do modelo normal padronizada tem forma de sino, sendo unimodal - os eventos empíricos se concentram em torno de uma média (igual a zero) - e simétrica.

Na teoria clássica, o escore total não é consistente, podendo variar de acordo com o teste aplicado e do contexto histórico utilizado. Os itens do instrumento de verificação, tem parâmetros dependentes do nível de aptidão do sujeito, sendo assim a avaliação de desempenho é influenciada pela amostra analisada. Portanto, um mesmo item pode ser considerado fácil ou difícil, variando de teste para teste.

2.1.1 Modelo de medição de um respondente

No que segue vamos adotar as representações do modelo que explica o funcionamento da TRI com em [10], no qual, um teste quando destinado a medir um respondente com habilidade (θ), representa-se os níveis de dificuldades dos itens calibrados sobre uma reta. Na **Figura 2.1**, b_1, b_2, b_3 e b_4 representam as dificuldades de quatro itens calibrados. Os itens mais difíceis, que desafiam os respondentes com maior habilidade, definem a extremidade direita e os itens mais fáceis, que não necessitam de muita habilidade para responder corretamente, definem a extremidade esquerda.

Note que b_4 está mais à direita, indicando ser um item com alta dificuldade, portanto o mais difícil do teste. Já b_1 está mais à esquerda, portanto é o item mais fácil do teste. Ao analisarmos a reta empírica, notamos que o respondente possui habilidade (θ) acima dos itens b_1, b_2 e b_3 e abaixo de b_4 . Portanto pode-se estimar que seu resultado será responder corretamente os itens b_1, b_2 e b_3 e responder incorretamente o item b_4 , obtendo um escore igual a 3.

Ao calibrar os itens, com os quais um indivíduo foi testado, em uma variável na escala de fácil a difícil, esperamos um padrão consistente de resposta do respondente em relação

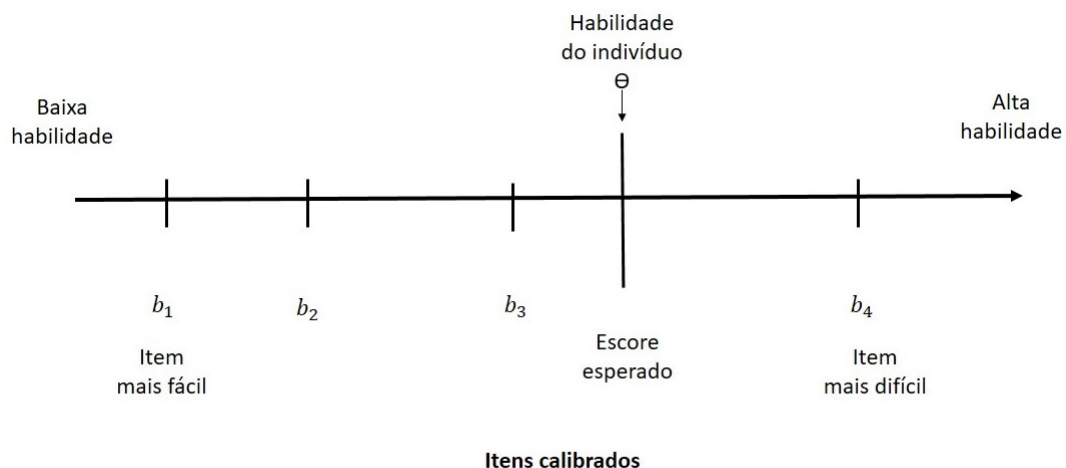


Figura 2.1: Medindo uma variável

a ordem de dificuldade desses itens ao longo da variável. Portanto, estima-se que ele tenha sucesso nos itens fáceis, que necessitam de uma habilidade inferior a dele, e insucesso nos itens difíceis, que necessitam de uma habilidade maior que a dele.

Observa-se, na **Figura 2.2**, dois padrões (A e B) de respostas para um mesmo teste de dez itens, ambos apresentando escore 6. Os itens estão localizados ao longo de uma variável em seus níveis de dificuldade.

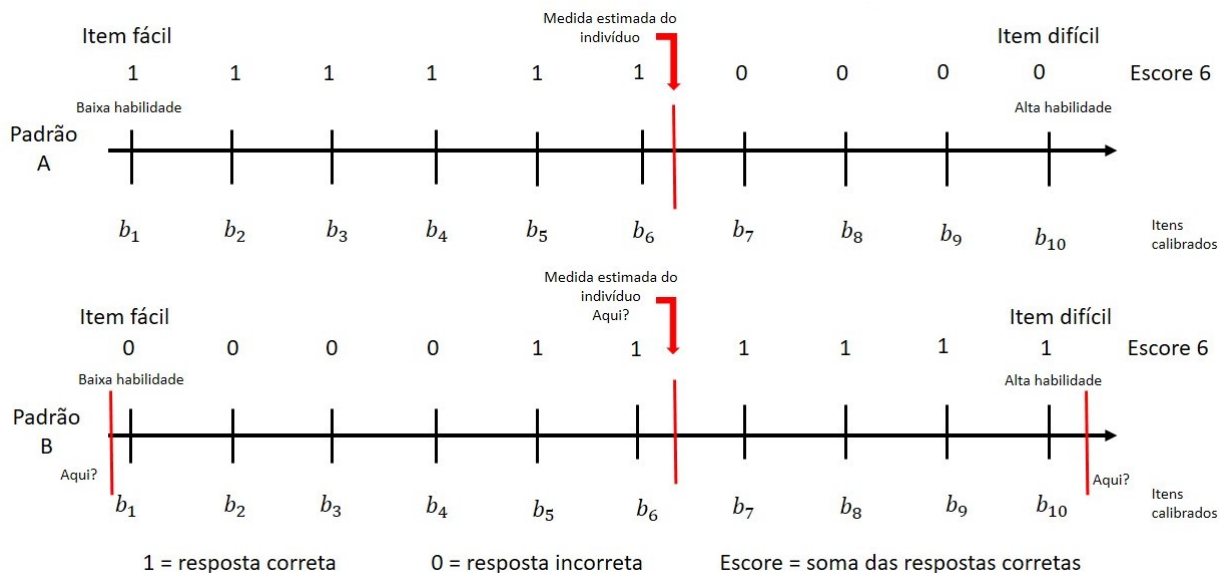


Figura 2.2: Validando o padrão de resposta

No padrão A, os seis itens que necessitam de uma habilidade inferior a proposta foram respondidos corretamente, e os outros quatro, mais difíceis foram respondidos incorretamente, o que sugere uma consistência no resultado. Portanto o indivíduo tem habilidade localizada entre b_6 e b_7 , assim seu escore aceitável é 6.

No padrão B, o escore 6 não é tão aceitável. O respondente respondeu corretamente os seis itens mais difíceis e respondeu incorretamente os quatro itens mais fáceis. Ao tentar localizar o respondente percebe-se alguns problemas:

- ao localizar sua habilidade acima de b_{10} , item mais difícil respondido corretamente, precisa-se justificar o erro nos quatro itens mais fáceis;
- ao localizá-la abaixo de b_1 , item mais fácil respondido incorretamente, como explicar o acerto dos itens mais difíceis?

O mesmo ocorre para qualquer nível ao longo da variável, como por exemplo, entre b_6 e b_7 , que seria o aceitável para o escore 6. Todas são insatisfatórias para a “medida” do respondente do padrão B, indicando que possivelmente exista algum problema, ou os itens usados no teste são não calibrados, ou essa pessoa não acertou de maneira pretendida (acerto casual), portanto nenhuma medida é aceitável para o padrão B.

Ao analisar esses padrões, destaca-se que ao construir itens que podem definir uma variável válida, precisa-se validar o padrão de resposta de cada respondente. Quando as calibrações dos itens são validadas por pessoas adequadas sapientemente, que apresentam as habilidades avaliadas, a maioria dos padrões de resposta se aproximam do padrão A. Mas, a ocorrência de padrões que se aproximam do padrão B nos obriga a examinar e validar rotineiramente o padrão de resposta de cada respondente antes que possam presumir uma medida a partir da sua pontuação no teste.

Quando um mesmo indivíduo responde testes diferentes é possível que seu escore seja diferente para cada teste analisado, pois o seu resultado não depende apenas de sua habilidade, mas das características dos itens dos testes. Na **Figura 2.3**, pode-se observar os escores esperados para o mesmo respondente submetido a cinco testes diferentes.

A habilidade do indivíduo está representada em cada uma das linhas da variável para cada teste, e dependendo das dificuldades dos seus itens, apresenta uma posição diferente sobre a variável, mas a sua localização é a mesma em cada linha. Na **Figura 2.3**, observa-se:

- O primeiro teste, muito fácil, tem itens tão fáceis que espera-se para essa pessoa um escore igual a oito.
- O segundo teste, limite do fácil, tem sete de seus itens abaixo da capacidade da pessoa e, portanto, espera-se um escore igual a sete.
- O terceiro teste, o mais amplo, tem cinco itens que devem ser fáceis para eles. Mesmo que este teste esteja centrado na mesma posição da variável como o segundo teste e assim tenha o mesmo nível de dificuldade médio, no entanto, devido a sua maior largura na dificuldade do item, espera-se apenas um escore cinco.
- O quarto teste, muito difícil, tem itens tão difíceis que espera-se um escore zero.

Como as pontuações dependem do nível e das dificuldades do item de teste

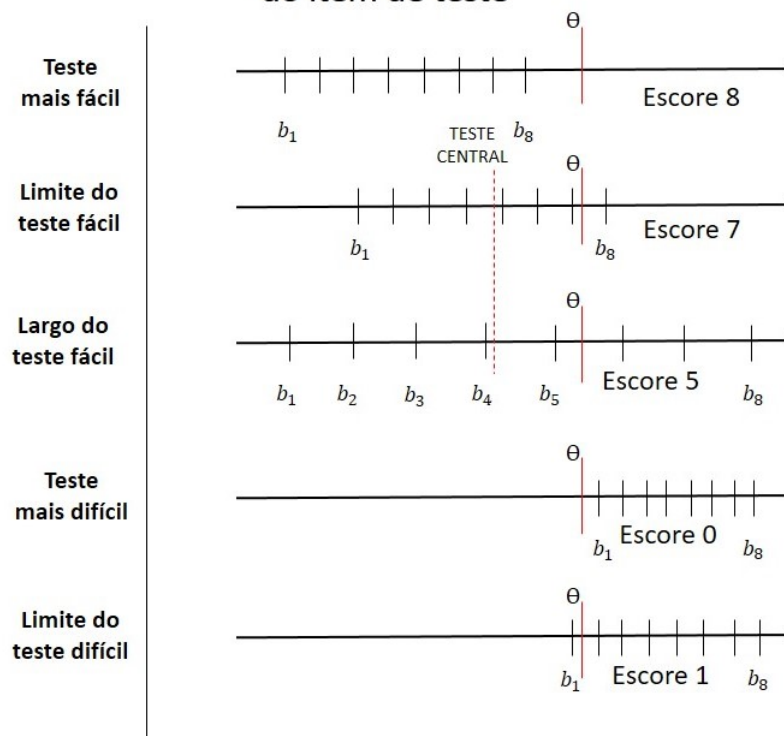


Figura 2.3: Escores do mesmo respondente para cinco testes diferentes

- O quinto teste, limite do difícil, tem sete itens acima da capacidade da pessoa e um abaixo. Nessa situação, o escore que se espera mais frequentemente seria um.

Para um mesmo respondente existem cinco possíveis resultados: 0, 1, 5, 7 e 8. Apesar da habilidade do respondente não mudar, os escores para cada teste são diferentes, pois dependem das características dos itens dos testes. Portanto, antes de determinar uma medida precisa-se “ajustar” a pontuação ao efeito dos itens, em particular, tornando as medidas de habilidade do respondente independente do teste. A dependência dos escores dos respondentes quanto a dificuldade dos itens é um problema comum na análise de testes.

2.2 Teoria de Resposta ao Item

A TRI propõe modelos para os traços latentes, que são características do indivíduo que não podemos observar diretamente, como proficiências e habilidades. Esses modelos relacionam a probabilidade de um indivíduo dar uma resposta correta com seus traços latentes. Dessa forma, o respondente com mais habilidade tem probabilidade maior de acertar o item.

Um ponto importante da TRI, é que o item tem papel fundamental na análise dos resultados. Devido a essa característica o uso desta teoria possibilita comparar populações

diferentes, desde que os testes apresentem alguns itens em comum, como também é possível comparar indivíduos de uma mesma população submetidos a provas diferentes.

Os primeiros modelos de resposta ao item surgiram em meados dos anos 50, eles consideravam uma única habilidade, um único grupo e as correções eram de maneira dicotômica. Mas os modelos vem evoluindo e apresentando adaptações que possibilitam observar outras características que podem influenciar na análise. Observamos bem isso na passagem do trecho de Andrade, Tavares & Valle (SINAPE 2000) abaixo.

Lord (1952) foi o primeiro a desenvolver o modelo unidimensional de 2 parâmetros, baseado na distribuição normal acumulada (ogiva normal). Após algumas aplicações desse modelo, o próprio Lord sentiu a necessidade da incorporação de um parâmetro que tratasse do problema do acerto casual. Assim, surgiu o modelo de 3 parâmetros. Anos mais tarde, Birnbaum (1968) substituiu, em ambos os modelos, a função ogiva normal pela função logística, matematicamente mais conveniente, pois é uma função explícita dos parâmetros do item e de habilidade e não envolve integração. Independentemente do trabalho de Lord, Rasch (1960) propôs o modelo unidimensional de 1 parâmetro, expresso também como modelo de ogiva normal e, também mais tarde descrito por um modelo logístico por Wright (1968).(SINAPE 2000)

Um dos dilemas da TRI é a necessidade de técnicas computacionais avançadas para realizar a estimação dos parâmetros envolvidos nos modelos, como por exemplo, para determinar a habilidade de um item. Para tanto, utilizava-se o método da máxima verossimilhança conjunta, porém, recentemente, os problemas de estimação dos parâmetros dos itens e de estimação da habilidade dos respondentes, tem nos métodos bayesianos uma possível alternativa aos problemas numéricos atribuído ao modelo que adota a máxima verossimilhança.

Os modelos encontrados na literatura dependem: da natureza do item (dicotômicos ou não dicotômicos), do número de populações envolvidas (apenas uma ou mais de uma) e da quantidade de traços latentes que está sendo medida (unidimensional ou multidimensional).

Nesta proposta de estudo, abordaremos o modelo unidimensional, visando avaliar um único traço latente.

2.2.1 Modelos para itens dicotômicos ou dicotomizados

Para realizar a análise de itens de múltipla escolha dicotomizados (corrigidos como certo ou errado), como também para análise de itens abertos, quando avaliados de forma dicotomizada, podemos utilizar alguns modelos de resposta ao item. Os mais utilizados são os

modelos logísticos para itens dicotômicos, que podem ser de três tipos e que se diferenciam pelo número de parâmetros que utilizam para descrever o item.

Eles são conhecidos como os modelos logísticos de 1, 2 e 3 parâmetros, que consideram, respectivamente:

- (i) somente a dificuldade do item;
- (ii) a dificuldade e a discriminação do item;
- (iii) a dificuldade, a discriminação e a probabilidade de resposta correta dada por indivíduos de baixa habilidade.

Segundo Andrade, Tavares & Valle (2000), dos modelos propostos pela TRI, o *modelo logístico unidimensional de 3 parâmetros (ML3)* é atualmente o mais utilizado e é dado por:

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

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

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

θ_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 θ_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, representa o ponto na escala de habilidade do respondente com 50% de probabilidade de responder corretamente ao item i .

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 assumindo frequentemente o valor 1. Utiliza-se o valor 1,7 quando deseja-se que a função logística forneça resultados semelhantes ao da

função ogiva normal.

A habilidade é medida arbitrariamente, pois o importante são as relações de ordem entre seus pontos. O parâmetro b é representado na mesma unidade da habilidade, visto que ele representa a habilidade necessária para uma probabilidade de acerto igual a $\frac{(1+c)}{2}$.

Como o parâmetro c representa a probabilidade de um indivíduo com baixa habilidade responder corretamente o item, ele assume valores entre 0 e 1. Quando no processo avaliativo não é permitido chutar, o parâmetro c assume o valor 0, e, portanto, b representará o ponto na escala de habilidade onde a probabilidade de acertar o item é 0,5.

A representação gráfica da Função $P(U_{ij} = 1|\theta_j)$ é representado por um "S", chamada de Curva Característica do Item (**Figura 2.4**). Segundo esse modelo, a relação entre indivíduos com maior habilidade possuem maior probabilidade de acertar o item não é linear.

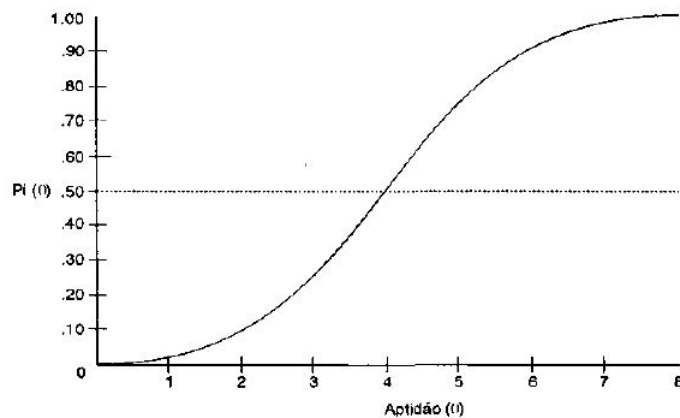


Figura 2.4: CCI (Pasquali, 1996)

Neste modelo é importante destacar que a medida que a habilidade do respondente aumenta, a probabilidade de responder corretamente o item também aumenta (veja a **Figura 2.4**). Portanto o parâmetro a não pode assumir valores negativos, uma vez que a é proporcional à derivada da curva no ponto. Outra observação importante, é que a discriminação entre bons alunos é feita a partir de itens considerados difíceis e não de itens considerados fáceis. A cada item é associado um intervalo na escala de habilidade no qual o item tem maior poder de discriminação. Este intervalo é definido em torno do valor do parâmetro b .

Na TRI, os parâmetros a e b podem, teoricamente, assumir qualquer valor real entre $-\infty$ e $+\infty$. Lembrando que a não pode assumir valores negativos. A **Figura 2.5** apresenta uma CCI destacando o significado geométrico dos parâmetros a e b .

Quando o parâmetro a assume valores muito baixos indicam que o item tem pouco poder de discriminação (alunos com habilidades bastante diferentes têm aproximadamente a mesma probabilidade de responder corretamente ao item). Por outro lado, quando o parâmetro a assume valores muito altos indicam itens com curvas características muito declivosas, que discriminam os alunos basicamente em dois grupos: os que possuem habilidades abaixo

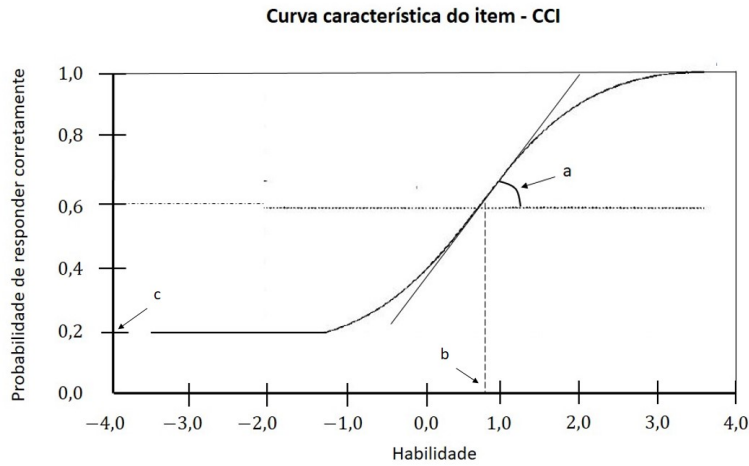


Figura 2.5: Parâmetros a, b e c na CCI(Adaptado de Andrade, Tavares & Valle (2000))

do valor do parâmetro b e os que possuem habilidades acima do valor do parâmetro b .

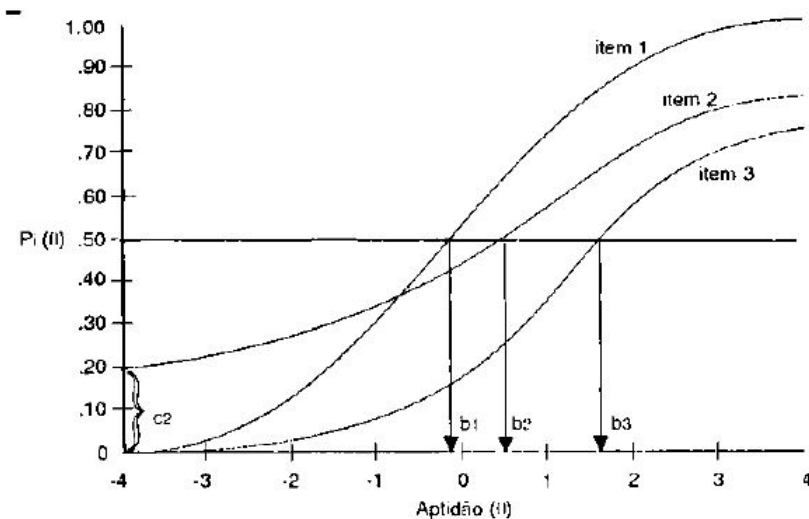


Figura 2.6: CCI do modelo de três parâmetros para três itens (Pasquali, 1996)

Na **Figura 2.6**, observa-se a representação de três Curvas características, referentes aos itens 1, 2 e 3. Quanto ao parâmetro a , que está relacionado a inclinação da CCI, portanto nota-se que o item 1 discrimina mais que os itens 2 e 3. O item 2 apresenta o menor poder de discriminação entre os itens analisados. Quanto ao parâmetro b , que representa a localização do indivíduo com relação a dificuldade, na escala de -4 a 4 , quando θ_i é igual a b_i , o indivíduo tem probabilidade de 0,50 de responder corretamente o item i . Nota-se que para responder corretamente o item 1, necessita-se de uma habilidade muito baixa, sendo assim considerado este item o mais fácil, já para o item 3, necessita-se que o respondente tenha habilidade superior a 1, requerendo do respondente uma maior habilidade, portanto podemos considerá-

lo um item difícil. Quanto ao parâmetro c , acerto casual, nota-se que para os itens 1 e 3, não há registro de acerto casual, enquanto que para o item 2, há uma probabilidade de 20% das respostas corretas terem sido dadas ao acaso.

Função de Informação do Item

A Função de Informação do item é uma medida bastante utilizada em conjunto com a CCI, que permite analisar quanto um item (ou teste) contém de informação para a medida de habilidade. Segundo Andrade & Dos Anjos (2012), a Função de Informação do Item, é dada por:

$$I_i(\theta) = \frac{[\frac{d}{d\theta}P_i(\theta)]^2}{P_i(\theta)Q_i(\theta)},$$

onde,

I_i é a "informação" fornecida pelo item i no nível de habilidade θ ;
 $P_i(\theta) = P(X_{ij} = 1|\theta)$ e $Q_i(\theta) = 1 - P_i(\theta)$.

Pode-se escrever esta equação para um modelo logístico de 3 parâmetros, como:

$$I_i(\theta) = D^2 a_i^2 \frac{Q_i(\theta)}{P_i(\theta)} \left[\frac{P_i(\theta) - c_i}{1 - c_i} \right]^2.$$

Os parâmetros a_i , b_i e c_i são fundamentais para a quantidade de informação do item. A informação será maior:

- (i) quando b_i se aproxima de θ , pois a probabilidade é, em média, de 0,50;
- (ii) quanto maior for o a_i ;
- (iii) e quanto mais c_i se aproximar de 0.

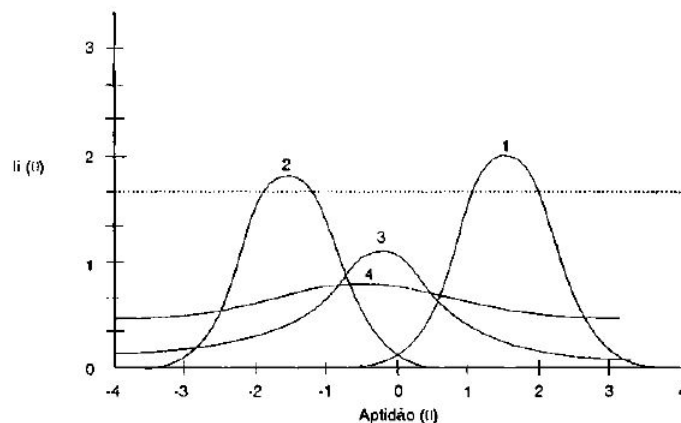


Figura 2.7: Funções de informação para quatro itens (Pasquali, 1996)

Na **Figura 2.7**, podemos observar que o item 1, é o que fornece mais informações para $\theta = 1,5$, já o item 2, oferece mais informações para $\theta = -1,5$. O item 3, oferece mais informações que os itens 1 e 2, quando $\theta = 0$. O item 4 é inútil ao teste, pois não produz

qualquer informação a mais do que a já produzida pelos itens 1, 2 e 3 em qualquer nível de θ .

Função de Informação do Teste

A soma das informações providas por cada item que compõe o teste, determina o montante de informação do mesmo:

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

A função de informação do teste também pode ser representada através do erro-padrão de estimação, dado por:

$$EP(\theta) = \frac{1}{\sqrt{I(\theta)}}.$$

As medidas de informações apresentadas dependem do valor de θ , assim com a amplitude do intervalo de confiança para θ , também depende do seu valor.

Quando se quer estimar a habilidade e os parâmetros de um item, precisa-se submeter o item a um grupo de indivíduos, e a partir de suas respostas estimá-los. Andrade, Tavares & Valle (SINAPE 2000) relata que uma vez estabelecida a escala de medida da habilidade, os valores dos parâmetros dos itens não mudam, isto é, seus valores são invariantes à diferentes grupos de respondentes, desde que os indivíduos destes grupos tenham suas habilidades medidas na mesma escala.

A Escala de Habilidade

Um teste com n questões do tipo certo/errado, define a medida escore com valores inteiros entre 0 e n , na TRI é diferente. Como já foi citado, a habilidade pode assumir qualquer valor real entre $-\infty$ e $+\infty$. O valor médio e o desvio-padrão das habilidades dos respondentes, são representados quando estabelecidos uma origem e uma unidade de medição da escala. Na TRI é muito utilizado a escala (0,1), que representa média igual a 0 e desvio-padrão igual a 1. Na escala (0,1), os valores do parâmetro b variam (tipicamente) entre -2 e +2, e para o parâmetro a , espera-se valores entre 0 e +2, sendo que os valores mais apropriados de a seriam aqueles maiores do que 1.

Mas outros valores da escala podem ser estabelecidos, pois o importante são as relações de ordem existente. Neste sentido, Andrade, Tavares & Valle (SINAPE 2000), aponta que a probabilidade de um indivíduo responder corretamente a um certo item é sempre a mesma, independentemente da escala utilizada para medir a sua habilidade, ou ainda, a habilidade de um indivíduo é invariante à escala de medida. Assim, não faz qualquer sentido querermos analisar itens a partir dos valores de seus parâmetros a e b sem conhecer a escala na qual eles foram determinados.

Suposições do Modelo: Unidimensionalidade e Independência Local

No modelo proposto, os itens de um teste devem nortear uma única habilidade, ou seja, o conjunto de itens devem supostamente medir um único traço latente. Assim propõe o postulado da unidimensionalidade, a existência de uma habilidade *dominante* (um fator dominante) responsável pelo conjunto de itens, logo supõe-se que num teste, está medindo esse fator.

A independência local ou independência condicional é uma suposição fundamental para o processo de estimação dos parâmetros do modelo. Ela determina que para uma dada habilidade, as respostas são independentes para os diferentes itens avaliados. É importante destacar que a unidimensionalidade implica independência local, portanto os itens devem satisfazer a suposição de unidimensionalidade, e assim também será contemplada a independência local (veja Andrade, Tavares & Valle (SINAPE 2000)).

Segundo Andrade, Tavares & Valle (SINAPE 2000),

As vantagens da utilização da TRI dependem fundamentalmente da adequação (ajuste) dos modelos e seus pressupostos. Por exemplo, somente a partir de modelos com bom ajuste é que pode-se garantir a obtenção de itens e habilidades invariantes. Uma excelente discussão das consequências da utilização de modelos inadequados aos dados e de métodos para verificação do ajuste e dos pressupostos do modelo utilizado, está apresentada no Capítulo 4 de Hambleton, Swaminathan & Rogers.(SINAPE 2000)

Outros modelos para itens dicotômicos

A partir do modelo logístico de 3 parâmetros podemos obter dois outros modelos:

- O *modelo logístico unidimensional de 2 parâmetros (ML2)*, dado por:

$$P(U_{ij} = 1|\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}, \quad (2.2)$$

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

Esse modelo é utilizado quando não existe possibilidade de acerto ao acaso, considerando $c = 0$ no modelo anterior.

O *modelo logístico unidimensional de 2 parâmetros (ML2)* (2.2) veio a substituir o *modelo da ogiva normal*, representado por:

$$P_i(\theta) = \int_{-\infty}^{a_i(\theta-b_i)} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt, \quad (2.3)$$

em que,

$P_i(t|\theta, a_i, b_i)$ é a probabilidade de um respondente com habilidade θ acertar um item i ; t é a realização da variável aleatória com distribuição normal padronizada;

- O *modelo logístico unidimensional de 1 parâmetro (ML1)*, também conhecido como modelo de **Rasch**, é dado por:

$$P(U_{ij} = 1|\theta_j) = \frac{1}{1 + e^{-D(\theta_j - b_i)}}, \quad (2.4)$$

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

Esse modelo é utilizado quando tivermos todos os itens com o *mesmo* poder de discriminação, com $a_i = 1$, além de não existir resposta ao acaso.

Sobre os modelos não dicotômicos, Andrade, Tavares & Valle (2000) informa:

Aqui são incluídos os modelos tanto para a análise de itens abertos (de resposta livre) quanto para a análise de itens de múltipla escolha que são avaliados de forma graduada, ou seja, itens que são elaborados ou corrigidos de modo a ter-se uma ou mais categorias intermediárias ordenadas entre as categorias certo e errado. Nesse tipo de item não se considera somente se o indivíduo respondeu à alternativa correta ou não, mas também leva-se em conta qual foi a resposta dada por ele. (SINAPE 2000)

Este trabalho tem como objeto de estudo os itens Dicotômicos, portanto listaremos alguns modelos não-dicotômicos, e sugerimos o Capítulo 2 de Andrade, Tavares & Valle (2000) para maior conhecimento desses modelos.

- Modelo de Resposta Nominal (Nominal Categories Model)
- Modelo de Resposta Gradual (Graded Response Model)
- Modelo de Escala Gradual (Rating Scale Model)
- Modelo de Crédito Parcial (Partial Credit Model)
- Modelo de Crédito Parcial Generalizado (Generalized Partial Credit Model)

2.3 Modelos envolvendo duas ou mais populações

Quando um teste é aplicado a dois ou mais grupos de respondentes, utiliza-se, na prática uma generalização dos modelos logísticos unidimensionais de 1, 2 e 3 parâmetros, visto que se encontra implementado computacionalmente. O modelo é dado por:

$$P(U_{ijk} = 1|\theta_{jk}) = c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_{jk} - b_i)}}, \quad (2.5)$$

com $i = 1, 2, \dots, I$, $j = 1, 2, \dots, n_k$ e $k = 1, 2, \dots, K$.

Onde:

U_{ijk} é uma variável dicotômica que assume os valores 1, quando o indivíduo j da população k responde corretamente ao item i , ou 0 quando o indivíduo não responde corretamente ao item;

θ_{jk} representa a habilidade do j -ésimo indivíduo da população k ;

$P(U_{ijk} = 1|\theta_{jk})$ é a probabilidade de um indivíduo j da população k , com habilidade θ_{jk} , responder corretamente ao item i .

Os demais parâmetros já foram descritos anteriormente.

Para comparar indivíduos de duas populações diferentes é necessário que os testes contenham alguns itens em comum, ou seja, os indivíduos não são submetidos todos aos mesmos itens, basta alguns itens em comum. Portanto o I representa o número de itens distintos apresentados. A comparação entre populações diferentes se tornou mais eficiente depois da implementação computacional desse modelo, representando um dos grandes avanços da TRI.

2.3.1 Estimação dos parâmetros

Na TRI, a estimação dos parâmetros dos itens e das habilidades dos respondentes representam um estágio importante. A estimação dos parâmetros dos itens é conhecida como *calibração*.

Na aplicação de um teste, normalmente sabemos apenas as respostas dos respondentes aos itens. Nos modelos de resposta ao item é preciso realizar a estimação dos parâmetros dos itens (1) e das habilidades dos indivíduos (2). Esta estimação, na base teórica, pode ser dividido em:

- É conhecido (1), precisa-se apenas estimar (2);
- É conhecido (2), precisa-se apenas estimar (1);

- Ambos desconhecidos, precisa-se estimar (1) e (2) simultaneamente.

Sobre as estimações citadas anteriormente, Andrade, Tavares & Valle (2000) ressalta

..., geralmente a estimação é feita pelo Método da Máxima Verossimilhança através da aplicação de algum processo iterativo, como o algoritmo Newton-Raphson (ver Issac & Keller (1966), por exemplo) ou "Scoring" de Fisher (ver Rao (1973), por exemplo). Alguns procedimentos bayesianos também são aplicados com bastante frequência (ver Mislevy (1986a), por exemplo).(SINAPE 2000)

Mais detalhes sobre a teoria, outros modelos, métodos de estimação podem ser vistos em Andrade, Tavares & Valle (2000), Andrade & Dos Anjos (1012), Pasquali (1996) etc.

Capítulo 3

O uso do *R* na TCT e na TRI

Tanto a TCT, como a TRI, contemplam a análise de itens e a análise do instrumento de medida, a partir da estimação dos seus parâmetros. Mas a aplicação dessas teorias, em particular a TRI, exige a utilização de recursos computacionais específicos que estão disponibilizados em vários programas, como o software *R*, o TESTFACT, o BILOG-MG, entre outros.

Neste trabalho, utilizaremos os recursos disponíveis no software *R*. O ambiente *R* é livre e utiliza uma linguagem de programação orientada ao objeto, permitindo ao usuário desenvolver funções específicas. Além da sua importância em análises de testes, por conter uma diversidade de métodos estatísticos. O *R* é gratuito e pode ser baixado pelo link: <http://www.r-project.org/>.

Sugere o software *Rstudio* como interface do software *R*. A versão compatível o computador utilizado pode ser baixado pelo link: <http://www.rstudio>.

As análises realizadas neste trabalho foram feitas na versão do *R*: *R* i386 3.2.4 Revised.

Mais detalhes sobre o software *R*, seus comandos e uso na psicometria podem ser observados em: Ora, Monteiro & Arica (2010); Ruch, Mair & Hatzinger (2013) e Zeireleis (2005), link: <https://CRAN.R-project.org/doc/Rnews/>.

3.1 Aplicações do *R* na TCT

O *R* é utilizado na TCT para obter algumas medidas que são utilizadas para avaliar a qualidade do instrumento de medida.

3.1.1 Coeficiente de correlação ponto-bisserial

O coeficiente de correlação ponto-bisserial (ρ_{pb}) é a correlação de Pearson entre uma variável dicotômica e o escore do teste e é definido por:

$$\rho_{pb} = \frac{\bar{X}_A - \bar{X}_T}{S_T} \sqrt{\frac{p}{1-p}} \quad (3.1)$$

em que,

\bar{X}_A é a média dos escores dos respondentes que acertaram o item;

\bar{X}_T é a média global dos escores do teste;

S_T é o desvio padrão do teste;

p é a proporção de respondentes que acertaram o item.

3.1.2 Coeficiente de correlação bisserial

O coeficiente de correlação biserial (ρ_b) é uma medida de associação entre uma variável dicotomizada e uma variável contínua, e é definido por:

$$\rho_b = \rho_{pb} \frac{\sqrt{p(1-p)}}{h(p)} \quad (3.2)$$

em que,

ρ_{pb} é a correlação ponto-biserial;

p é a proporção de respondentes que acertaram o item;

$h(p)$ é o valor da densidade da distribuição normal padrão no ponto em que a área da curva à esquerda deste ponto é igual a p .

O coeficiente de correlação biserial corresponde ao índice de discriminação do item.

3.1.3 Coeficiente alfa de Cronbach

O coeficiente alfa de Cronbach é utilizado para medir a consistência interna do instrumento de medida, ele avalia o quanto os itens que compõem o teste medem o mesmo constructo, ou seja, o quanto cada item correlaciona-se com o conjunto de itens do teste. É definido por:

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum s_i^2}{s_T^2} \right) \quad (3.3)$$

em que,

n é o número de itens;

$\sum s_i^2$ é a soma das variâncias dos n itens;

s_T^2 é a variância global dos escores dos testes.

O coeficiente alfa de Cronbach varia de 0 a 1. Quando o resultado está mais próximo de 0 menor a consistência e quando está mais próximo de 1 maior a consistência do teste.

3.2 Aplicações do *R* na TRI

O *R* é utilizado na TRI com o pacote "MCMCpack", e a estimação bayesiana revia métodos MCMC. Citamos anteriormente, no capítulo 2, os principais modelos logísticos unidimensionais da TRI. Entre eles os modelos de logísticos de 1 (2.3), 2 (2.2) e 3 (2.1) parâmetros. Assim como foram citadas a *Função de informação do item e função do teste*.

3.3 O *R* e a psicometria

A linguagem *R* é amplamente usada entre estatísticos e analistas de dados, pois disponibiliza de uma grande variedade de técnicas estatísticas e também de outros recursos, como gráficos. O *R* está disponível como software livre, nos termos da GNU GPL (General Public License) e pode ser encontrado no seguinte endereço: <http://cran.r-project.org/>.

Na psicometria, utiliza-se o pacote *Psychometrics*. Esse pacote pode ser acessado em: <http://cran-r.c3sl.ufpr.br/web/views/Psychometrics.html>.

As capacidades do *R* são estendidas, podendo baixar todos os pacotes listados em Psicometria de uma só vez. Primeiro instale o pacote *ctv*. Para instalar um pacote, utilize a função *install.packages('nomedopacote')* (com aspas). Em seguida, instale os pacotes da área de interesse. Para utilizar o pacote, utilize a função *library(nomedopacote)* (sem aspas).

```
> install.packages("ctv")
> library(ctv)
> install.views("Psychometrics")
```

No *R* existem conjuntos de dados (datasets) disponíveis em vários pacotes. Para saber quais os datasets instalados em seu computador utilize *data()*. Para utilizar dados de algum pacote, digite *data(nomedodataset)*.

Capítulo 4

Itens - Elaboração e revisão de itens

O Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) é responsável por importantes instrumentos de avaliação em larga escala do país, como o Exame Nacional do Ensino Médio (ENEM), principal mecanismo de seleção para o ingresso no ensino superior e de certificação do Ensino Médio, desde 2009. O ENEM ocorre anualmente submetendo vários jovens de todo o Brasil a este processo de avaliação.

As avaliações em larga escala são importantes por envolverem um grande público, e assim propor intervenções em um âmbito superior ao da instituição de ensino. Outro ponto importante deste tipo de avaliação é a aplicação de um instrumento de medida e a análise dos resultados, possibilitam medir o desempenho dos participantes amplamente e com ponderação.

O ENEM utiliza o teste como principal instrumento de avaliação. Neste dispositivo o item é de grande importância, pois deve estar em conformidade com o contexto educacional, além de ser fundamental na observação dos resultados, visto que se utiliza a TRI no processo avaliativo. Um teste deve ser constituído por itens fáceis, de média dificuldade e difíceis, porém nem todo item é interessante para a construção de instrumentos de avaliação deste tipo.

Os itens relevantes compõem um banco de itens segundo alguns critérios. A existência de um Banco Nacional de Itens (BNI) no Inep é necessária para que se tenha uma quantidade expressiva de itens com comprovada qualidade técnico-pedagógica e psicométrica para comporem os testes de uma avaliação em larga escala (Guia de Elaboração e Revisão de Itens, volume 1, pág. 12)

Na construção de um item, é preciso observar o que se espera que o indivíduo tenha adquirido em seu processo de ensino-aprendizagem, competências e habilidades. O Inep propõe uma Matriz de Referência, projetada a partir das competências e habilidades espera-

das na educação básica, a fim de promover uma clareza no processo avaliativo e possibilitar a todo participante preparar-se equiparavelmente para o exame.

No processo de elaboração dos itens de Matemática, leva-se em conta o que um estudante aprendeu e se ele é capaz de utilizar esse aprendizado para resolver um dado problema, pois a resolução de problemas possibilita o desenvolvimento de capacidades de raciocínio, tais como intuição, indução, dedução e estimativa. (Guia de elaboração de itens, MT, pág 14)

Os itens podem ser de resposta livre ou, de resposta orientada ou objetiva, mas o Inep usa estritamente os itens objetivos, que podem ser apresentadas de diferentes formas, mas sempre é priorizado os de múltipla escolha, onde o respondente escolhe uma das alternativas entre várias, sendo apenas uma correta.

Os itens objetivos apresentam mais facilidade e rapidez na correção, além da atribuição de escores não estarem sujeitos ao erro de julgamento. A agilidade e a objetividade na correção são características que sugerem os itens objetivos como mais recomendáveis a avaliação em grande escala.

Um item de qualidade deve ser contextualizado, propondo situações-problemas que incorporem experiências do dia-a-dia. Assim o respondente é levado a investigar e definir estratégias para resolução do item. É importante ficar atento a contextos que induzem ao erro, famosas pegadinhas, pois o participante é levado a errar por falta de atenção, e não por falta de domínio da habilidade testada.

4.1 Elaboração do item de múltipla escolha

Produzir um item de múltipla escolha é algo engenhoso e rústico, que demanda do elaborador domínio da área de conhecimento, compreensão de características educacionais e psicológicas dos participantes e criatividade ao sugerir situações-problemas. Várias revisões são necessárias para garantir a qualidade do item.

Segundo o Guia de Elaboração e revisão de Itens do INEP, volume 1:

O item deve ser estruturado de modo que se configure uma unidade de proposição e contemple uma única habilidade da Matriz de Referência. Para tanto, devem ser observadas a coerência e a coesão entre suas partes (texto-base, enunciado e alternativas), de modo que haja uma articulação entre elas e se explicita uma única situação problema e uma abordagem homogênea de conteúdo. (INEP, 2000)

Os testes do Inep utilizam itens de múltipla escolha constituídos por texto-base, enunciado e alternativas. Eles indicam a situação-problema relacionada a habilidade a ser avaliada, direcionando o comando para solucioná-la.

O texto-base instiga a situação-problema, que pode ser formulada a partir de um ou mais textos-bases (textos verbais ou não-verbais) e apresenta as informações necessárias para a resolução da situação-problema. O enunciado direciona a percepção da tarefa a ser realizada pelo respondente do item, a partir dele é tomada a decisão sobre a estratégia a ser utilizada. As alternativas são representadas por gabarito e distratores, e representam as possíveis respostas a ser escolhidas para solucionar a situação-problema.

Para entendermos melhor os critérios de construção de um item iremos analisar alguns itens utilizados no ENEM 2017. Elencando a seguir as principais características a serem observadas, baseados em alguns critérios do Guia de Elaboração e Revisão de Itens, INEP, volume 1.

4.1.1 Quanto ao texto-base

I. Habilidade da Matriz de Referência

O item deve contemplar uma única habilidade da Matriz de Referência do Inep (veja as **Figuras 4.1 e 4.2**).

II. Contextualização - as contextualizações são reais ou artificiais, sem ligação com a realidade

Um item contextualizado deve propor uma situação que o respondente pode relacionar a situações normalmente vivenciadas no dia a dia.

O contexto do item pode ser representado por um ou mais textos-base que podem ser: formulados pelo próprio elaborador ou referenciados por publicações de apropriação

MATEMÁTICA E SUAS TECNOLOGIAS – ENSINO MÉDIO

<p style="text-align: center;">EIXOS COGNITIVOS</p> <p style="text-align: center;">COMPETÊNCIAS DE MATEMÁTICA E SUAS TECNOLOGIAS</p>	<p>I - Dominar a norma culta da Língua Portuguesa e fazer uso das linguagens matemática, artística e científica.</p>	<p>II - Construir e aplicar conceitos das várias áreas do conhecimento para a compreensão de fenômenos naturais, de processos histórico-geográficos, da produção tecnológica e das manifestações artísticas.</p>	<p>III - Selecionar, organizar, relacionar, interpretar dados e informações representados de diferentes formas, para tomar decisões e enfrentar situações-problema.</p>	<p>IV - Relacionar informações, representadas em diferentes formas, e conhecimentos disponíveis em situações concretas, para construir argumentação consistente.</p>	<p>V - Recorrer aos conhecimentos desenvolvidos para a elaboração de propostas de intervenção solidária na realidade, respeitando os valores humanos e considerando a diversidade sociocultural.</p>
<p>M1 Construir significados e ampliar os já existentes para os números naturais, inteiros, racionais e reais.</p>	<p>H1 - Utilizar no contexto social diferentes significados e representações dos números — naturais, inteiros, racionais ou reais.</p>	<p>H2 - Utilizar algum procedimento de cálculo com números naturais, inteiros, racionais ou reais.</p>	<p>H3 - Resolver situação-problema com números naturais, inteiros racionais ou reais envolvendo significados da adição, subtração, multiplicação ou divisão, potenciação ou radiciação.</p>	<p>H4 - Avaliar a razoabilidade de um resultado numérico na constituição de argumentos sobre afirmações quantitativas.</p>	<p>H5 - Avaliar propostas de intervenção na realidade, utilizando conhecimentos numéricos.</p>
<p>M2 Utilizar o conhecimento geométrico para realizar a leitura e a representação da realidade e agir sobre ela.</p>	<p>H6 - Interpretar a localização e a movimentação de pessoas/objetos no espaço tridimensional e sua representação no espaço bidimensional.</p>	<p>H7 - Identificar características de polígonos ou sólidos (prismas, pirâmides, cilindros).</p>	<p>H8 - Resolver situação-problema que envolva noções geométricas (ângulo, paralelismo, perpendicularismo).</p>	<p>H9 - Utilizar o teorema de Pitágoras ou semelhança de triângulos na seleção de argumentos propostos como solução de problemas do cotidiano.</p>	
<p>M3 Construir e ampliar noções de grandezas e medidas para a compreensão da realidade e a solução de problemas do cotidiano.</p>	<p>H10 - Estabelecer relações entre diferentes unidades de medida (comprimento, massa, capacidade, área, volume).</p>	<p>H11 - Aplicar a noção de escalas na leitura de plantas ou mapas.</p>	<p>H12 - Resolver situação-problema que envolva medidas de arcos ou ângulos (grau e radiano), utilizando teorema de Pitágoras ou razão trigonométrica (seno de um ângulo agudo).</p>	<p>H13 - Avaliar a razoabilidade do resultado de uma medição, na construção de um argumento consistente.</p>	<p>H14 - Avaliar propostas de intervenção na realidade utilizando cálculos de perímetros, área de superfícies planas ou volume de blocos retangulares.</p>

Figura 4.1: Matriz de Referência do Ensino Médio - INEP

<p>EIXOS COGNITIVOS</p> <p>COMPETÊNCIAS DE MATEMÁTICA E SUAS TECNOLOGIAS</p>	<p>I - Dominar a norma culta da Língua Portuguesa e fazer uso das linguagens matemática, artística e científica.</p>	<p>II - Construir e aplicar conceitos das várias áreas do conhecimento para a compreensão de fenômenos naturais, de processos histórico-geográficos, da produção tecnológica e das manifestações artísticas.</p>	<p>III - Selecionar, organizar, relacionar, interpretar dados e informações representados de diferentes formas, para tomar decisões e enfrentar situações-problema.</p>	<p>IV - Relacionar informações, representadas em diferentes formas, e conhecimentos disponíveis em situações concretas, para construir argumentação consistente.</p>	<p>V - Recorrer aos conhecimentos desenvolvidos para elaboração de propostas de intervenção solidária na realidade, respeitando os valores humanos e considerando a diversidade sociocultural.</p>
<p>M4 Construir e ampliar noções de variação de grandeza para a compreensão da realidade e a solução de problemas do cotidiano.</p>	<p>H15 Identificar leis matemáticas que expressem a relação de dependência entre duas grandezas.</p>	<p>_____</p>	<p>H16 Resolver situação-problema envolvendo a variação de grandezas direta ou inversamente proporcionais.</p>	<p>H17 Utilizar informações expressas em forma de juros (simples ou composto) como recurso para a construção de argumentação (aumentos e descontos sucessivos).</p>	<p>H18 Avaliar propostas de intervenção na realidade, utilizando cálculos de porcentagem e/ou juros.</p>
<p>M5 Aplicar expressões algébricas para modelar e resolver problemas, envolvendo variáveis socioeconômicas ou técnico-científicas.</p>	<p>H19 - Identificar representações algébricas que expressem a relação de interdependência entre duas grandezas.</p>	<p>H20 - Identificar gráfico cartesiano que represente a relação de interdependência entre duas grandezas (variação linear).</p>	<p>H21 - Resolver situação-problema cujos dados estejam expressos em gráfico cartesiano que mostre a variação de duas grandezas.</p>	<p>_____</p>	<p>_____</p>
<p>M6 Interpretar informações de natureza científica e social obtidas da leitura de gráficos e tabelas, realizando previsão de tendência, extrapolação, interpolação e interpretação.</p>	<p>H22 - Identificar informações apresentadas em tabelas ou gráficos (de coluna, de setores e de linha).</p>	<p>H23 - Utilizar informações expressas em gráficos ou tabelas para fazer inferências.</p>	<p>H24 - Resolver situação-problema com dados apresentados em forma de tabela de dupla entrada ou gráfico.</p>	<p>H25 - Utilizar informações expressas em gráficos ou tabelas como recurso para a construção de argumentos.</p>	<p>H26 - Avaliar propostas de intervenção na realidade, utilizando informações expressas em gráficos ou tabelas.</p>
<p>M7 Compreender o caráter aleatório e não-determinístico dos fenômenos naturais e sociais e utilizar instrumentos adequados para medidas e cálculos de probabilidade, para interpretar informações de variáveis apresentadas em uma distribuição estatística.</p>	<p>_____</p>	<p>H27 - Calcular a média aritmética de um conjunto de dados expressos em uma tabela de frequências de dados agrupados (não em classes) ou gráficos de colunas.</p>	<p>H28 - Resolver situação-problema que envolva processos de contagem ou noções de probabilidade.</p>	<p>H29 - Utilizar médias aritméticas, noção de probabilidade ou conhecimentos estatísticos como recurso para a construção de argumentação.</p>	<p>H30 - Avaliar propostas de intervenção na realidade utilizando probabilidade e/ou conhecimentos estatísticos (porcentagem, gráficos, médias).</p>

Figura 4.2: Matriz de Referência do Ensino Médio - INEP

pública. Neste último caso o texto-base deverá estar conforme a redação original e não poderá ter tradução livre, além de implicar em citação da respectiva fonte, conforme as normas da Associação Brasileira de Normas Técnicas (ABNT). Destacando também que a referência bibliográfica deve ser fidedigna e não podendo ser utilizado livro didático como fonte para o texto-base.

III. Extensão do texto

Deve-se evitar textos muito extensos que demandem dispendioso tempo de leitura, levando em consideração o tempo de leitura do item durante a realização do exame.

IV. Afirmações ou palavras que atrapalham o entendimento do texto

O texto-base deve ser claro e objetivo, não abordando temas que suscitem polêmicas ou que induzam o participante ao erro ("pegadinhas"). Evitar usar termos que confirmam ambiguidade à interpretação da tarefa a ser realizada.

V. Qualidade dos textos não verbais

O texto-base pode se apresentar por textos verbais e não verbais, como imagens, figuras, tabelas, gráficos ou infográficos, esquemas, quadros, experimentos, entre outros. Estes elementos não verbais podem ser fundamentais na compreensão da situação-problema e no desenvolvimento de estratégias de resolução do item, portanto devem estar adequados ao entendimento dos respondentes.

4.1.2 Quanto ao enunciado

I. Considera a totalidade das informações previamente oferecidas

O enunciado consiste numa instrução que poderá ser expressa como pergunta ou frase a ser completada pela alternativa correta. Deverá considerar exatamente a totalidade das informações previamente oferecidas, não podendo apresentar informações adicionais ou complementares ao texto base.

II. Clareza e Objetividade

A tarefa a ser realizada deve estar expressa de forma clara e objetiva, devendo utilizar termos impessoais. Alguns termos podem induzir ao erro por falta de atenção, portanto deve-se evitar termos como: "falso", "exceto", "incorreto", "não", "errado"; como também termos absolutos. Sentenças tais: "pode-se afirmar que", "É correto afirmar que" etc.; não devem ser utilizadas.

4.1.3 Quanto as alternativas

I. Coerência com o enunciado

As alternativas compreendem gabarito e distratores. Onde o gabarito, única alternativa correta, deve estar exposto de forma clara e não deve estar mais atrativo que os distratores. Deve-se evitar repetições de palavras que aparecem no enunciado, além de evitar alternativas demasiadamente longas. A pontuação das alternativas e as regras para a sua redação dependem do formato do enunciado. Se a instrução for uma frase incompleta, as alternativas devem começar com letras minúsculas e terminar com ponto apropriado para a frase. Caso o enunciado seja uma pergunta, as alternativas devem começar com letras maiúsculas.

II. Distratores plausíveis

Os distratores não devem ser absurdos em relação à situação-problema e os mesmos devem parecer corretas para os respondentes do teste que não desenvolveram a habilidade contemplada no item. Cada distrator representa uma dificuldade apresentada pelo participante em relação a habilidade em questão.

III. Disposição lógica

As alternativas devem estar dispostas de maneira lógica (sequência narrativa, alfabética, ordem crescente/decrecente etc.).

4.2 Análise de itens

Nesta seção apresenta-se a análise de alguns itens do ENEM 2017 segundo o Guia de Elaboração e Revisão de Itens do INEP. A fim de apresentar os itens à luz da matriz de referência e ganhar experiência na elaboração de itens.

4.2.1 (Item 161- ENEM 2017 - caderno 6 - prova cinza)

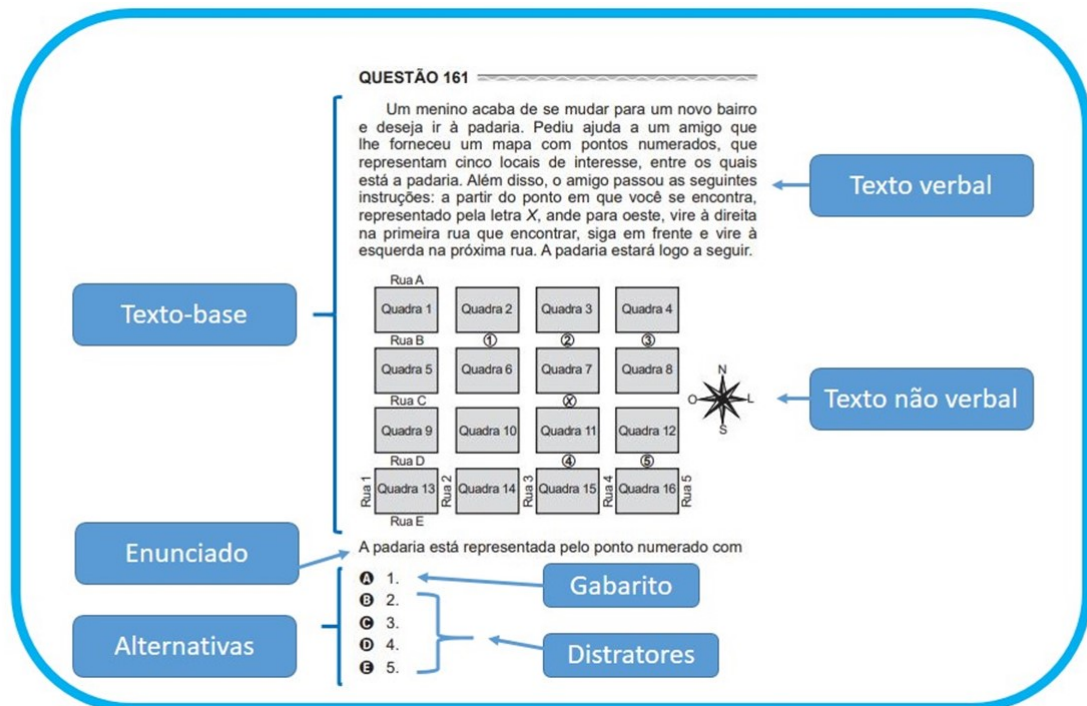


Figura 4.3: Item 161 - ENEM 2017

O item contempla a habilidade H6 - Interpretar a localização e a movimentação de pessoas/objetos no espaço tridimensional e sua representação no espaço bidimensional, que correlaciona a Competência M2 - Utilizar o conhecimento geométrico para realizar a leitura e a representação da realidade e agir sobre ela, com o Eixo Cognitivo I - Dominar a norma culta da Língua Portuguesa e fazer uso das linguagens matemática, artística e científica, segundo a Matriz de Referência de Matemática e suas Tecnologias - Ensino Médio, INEP.

O texto-base está representado por texto e imagem, ambos importantes para a compreensão da situação-problema. O texto é conciso, não aborda temas polêmicos e não induz o respondente ao erro, pois não utiliza termos que favorecem a dúvida ou cria sentido ambíguo. A imagem é fundamental para a resolução da questão e se apresenta com boa qualidade, com recursos que facilitam as estratégias de resolução.

A contextualização sugere uma situação-problema que atenta para uma situação hipotética, apresentando objetividade e clareza nas definições. O enunciado considera a totalidade das informações sugeridas no texto-base, com um comando pragmático em busca da alternativa correta.

As alternativas são coerentes com o enunciado e independentes umas das outras. Elas completam a sentença do enunciado, portanto finalizadas por um ponto, e estão expressas por números dispostos em ordem crescente. O gabarito está claro, e os distratores não estão representados de forma absurda em relação a situação-problema.

4.2.2 (Item 173- ENEM 2017 - caderno 6 - prova cinza)

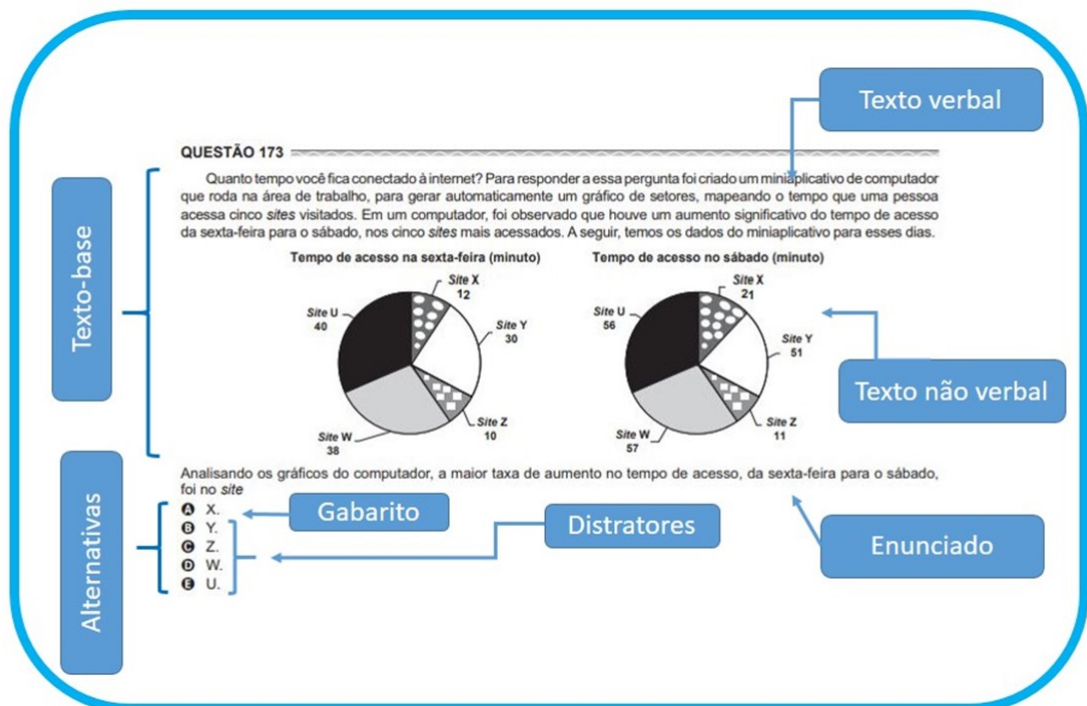


Figura 4.4: Item 173 - ENEM 2017

O item contempla a habilidade H22 - Identificar informações apresentadas em tabelas ou gráficos (de coluna, de setores e de linha) que correlaciona a Competência M6 - Interpretar informações de natureza científica e social obtidas da leitura de gráficos e tabelas, realizando previsão de tendência, extrapolação, interpolação e interpretação, com o Eixo Cognitivo I - Dominar a norma culta da Língua Portuguesa e fazer uso das linguagens matemática, artística e científica, segundo a Matriz de Referência de Matemática e suas Tecnologias - Ensino Médio, INEP.

O texto-base aborda um tema atual e presente no cotidiano da maioria do público-alvo, despertando a curiosidade e a materialização da situação-problema. Está composto de um texto explicativo associado a dados expressos em gráficos que possibilitam o entendimento da situação-problema e o desenvolvimento de estratégias de resolução.

Os gráficos estão claros e com boa qualidade, como não há citações de fonte referenciado por publicações de apropriação pública, entende-se que o texto-base é de autoria do próprio elaborador. O enunciado está representado de forma clara e objetiva, não apresentando termos que confirmam ambiguidade à interpretação da tarefa a ser realizada.

As alternativas estão coerentes e completam a sentença do enunciado. São independentes umas das outras e estão dispostas de maneira lógica, seguindo o sentido horário das informações citadas no gráfico de setores. O gabarito está exposto de forma clara, apresentando paralelismo com os distratores.

4.2.3 (Item 164- ENEM 2017 - caderno 6 - prova cinza)

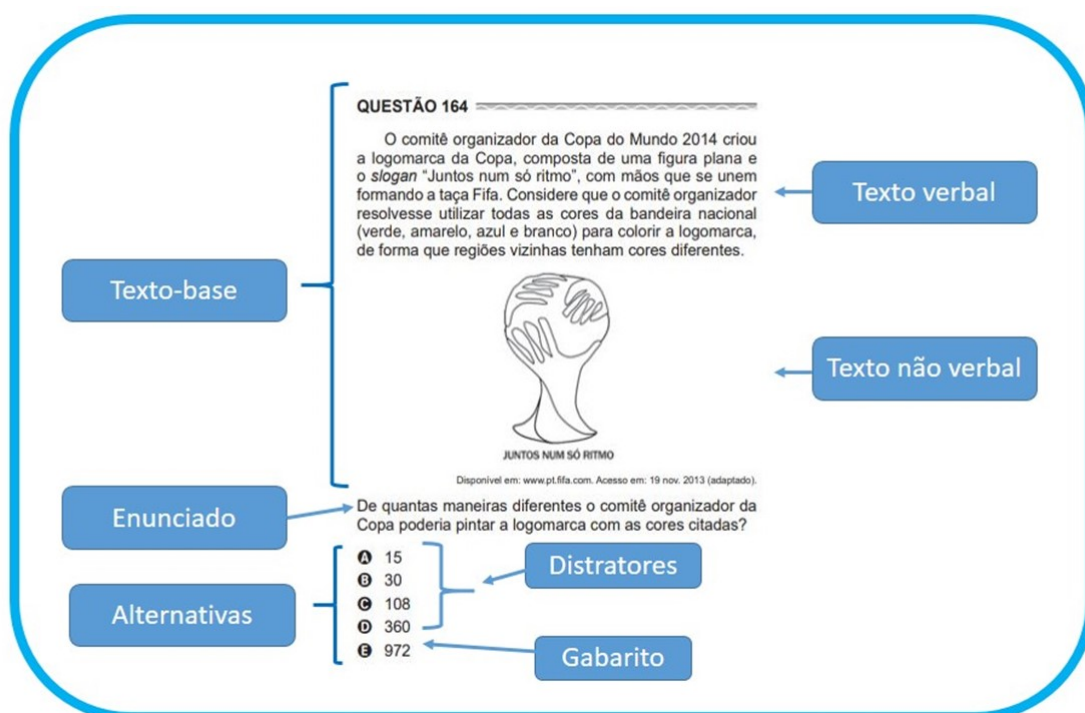


Figura 4.5: Item 164 - ENEM 2017

O item contempla a habilidade H28 - Resolver situação-problema que envolva processos de contagem ou noções de probabilidade, que correlaciona a Competência M7 - Compreender o caráter aleatório e não-determinístico dos fenômenos naturais e sociais e utilizar

instrumentos adequados para medidas e cálculos de probabilidade, para interpretar informações de variáveis apresentadas em uma distribuição estatística, com o Eixo Cognitivo III - Selecionar, organizar, relacionar, interpretar dados e informações representados de diferentes formas, para tomar decisões e enfrentar situações-problema, segundo a Matriz de Referência de Matemática e suas Tecnologias - Ensino Médio, INEP.

O texto-base utiliza uma publicação de apropriação pública adaptada, onde a fonte está citada conforme as normas da ABNT. O texto-base é formado por um texto e uma ilustração que apresentam informações importantes para a resolução da situação problema. O texto é claro e não possui termos que conferem ambiguidade à interpretação, como também não necessita de muito tempo para a compreensão.

O enunciado está representado por uma pergunta que considera as informações previamente oferecidas, não apresenta termos que confundam o entendimento da situação-problema. A tarefa a ser realizada está limpidamente expressa no enunciado.

As alternativas representam resposta do questionamento e assim não devem ser pontuadas, essa regra é característica da área de conhecimento Matemática e suas Tecnologias. Estão representadas por números inteiros dispostos em ordem crescente e são independentes umas das outras. Os distratores são plausíveis e indicam possíveis erros que ocorreram durante o processo de ensino-aprendizagem e indicam que o respondente não desenvolveu a habilidade proposta.

4.2.4 *(Item 162- ENEM 2017 - caderno 6 - prova cinza)*

O item contempla a habilidade H27 - Calcular a média aritmética de um conjunto de dados expressos em uma tabela de frequências de dados agrupados (não em classes) ou gráficos de colunas, que correlaciona a Competência M7 - Compreender o caráter aleatório e não-determinístico dos fenômenos naturais e sociais e utilizar instrumentos adequados para medidas e cálculos de probabilidade, para interpretar informações de variáveis apresentadas em uma distribuição estatística, com o Eixo Cognitivo II - Construir e aplicar conceitos das várias áreas do conhecimento para a compreensão de fenômenos naturais, de processos histórico-geográficos, da produção tecnológica e das manifestações artísticas, segundo a Matriz de Referência de Matemática e suas Tecnologias - Ensino Médio, INEP.

O texto-base é constituído por um texto e uma tabela formulados pelo próprio elaborador e de fácil interpretação, pois aborda um tema presente no cotidiano no ambiente escolar. Não apresenta informações que despertem dúvida ou induzam ao erro, que demandem mais tempo pra leitura. A tabela apresenta dados que são fundamentais para a resolução

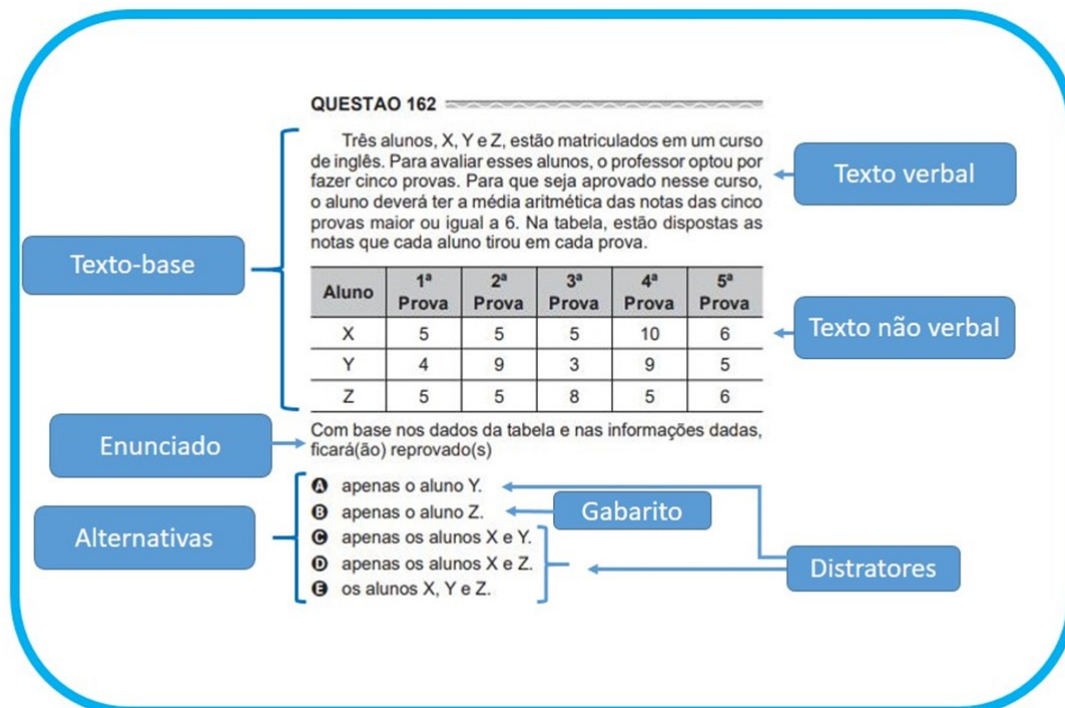


Figura 4.6: Item 162 - ENEM 2017

da situação-problema e estão representados claramente.

O enunciado considera a totalidade das informações previamente oferecidas e instrui de forma objetiva a ação a ser realizada. As alternativas completam e são coerentes com o enunciado, e o gabarito está exposto de forma clara, sem destacar-se em relação aos distratores. Os termos utilizados não causam confusão ou dúvida na de identificar a resposta certa. Os distratores são plausíveis e indicam que o respondente não construiu a habilidade esperada.

4.2.5 ((Item 180- ENEM 2017 - caderno 6 - prova cinza)

O item contempla a habilidade H16 - Resolver situação-problema envolvendo a variação de grandezas direta ou inversamente proporcionais, que correlaciona a Competência M4 - Construir e ampliar noções de variação de grandeza para a compreensão da realidade e a solução de problemas do cotidiano, com o Eixo Cognitivo III - Selecionar, organizar, relacionar, interpretar dados e informações representados de diferentes formas, para tomar decisões e enfrentar situações-problema, segundo a Matriz de Referência de Matemática e suas Tecnologias - Ensino Médio, INEP.

O texto-base é formado por um texto um tanto extenso, mas que não demanda dispendioso tempo para a compreensão. Utiliza uma publicação de apropriação pública adaptada,

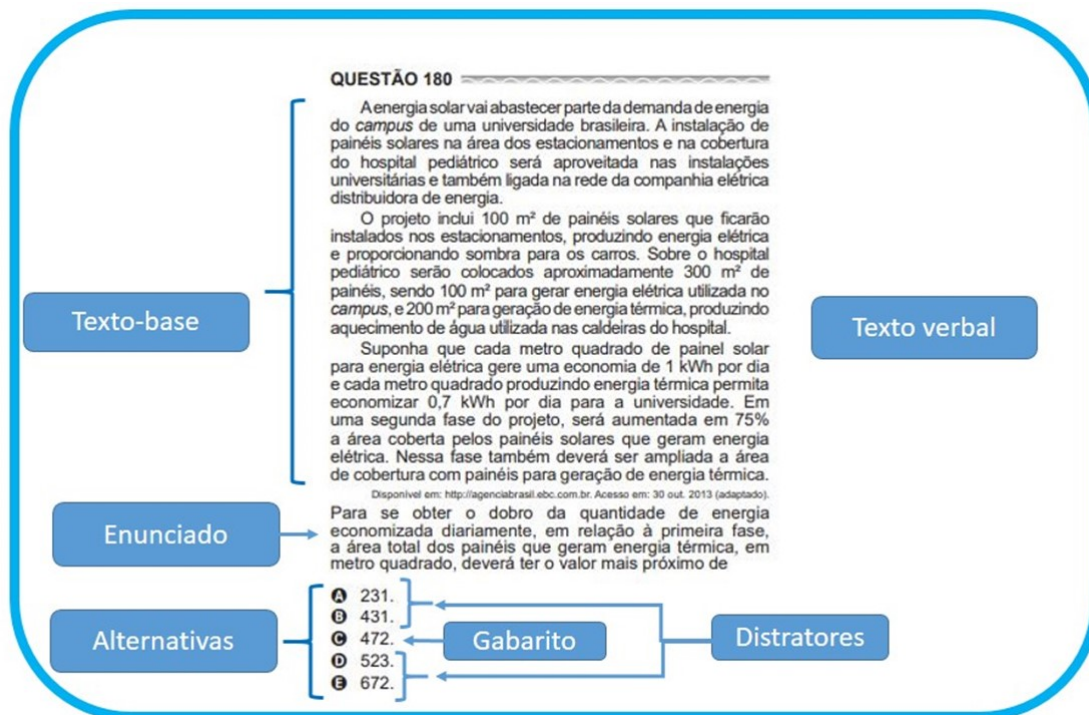


Figura 4.7: Item 180 - ENEM 2017

onde a fonte está citada conforme as normas da ABNT. A situação-problema aborda um tema atual que não suscita polêmica, como propõe situações capazes de induzir o respondente ao erro.

O enunciado é extenso, mas não apresenta informações adicionais ao texto-base, compreende a totalidade das informações ofertadas anteriormente e instrui claramente a tarefa a ser realizada. Está representada por uma frase a ser completada por uma das alternativas sugeridas.

As alternativas estão representadas por números inteiros dispostos em ordem crescente, com o gabarito inquestionável e distratores plausíveis, retratando hipóteses de raciocínio utilizadas na busca da solução da situação problema apresentada. O gabarito não se destaca em relação as demais alternativas, assim como os distratores não são absurdos em relação à situação-problema.

Capítulo 5

Itens elaborados para calibração

Os itens a seguir foram formulados segundo os critérios estabelecidos pelo Guia de Elaboração e Revisão de Itens e Matriz de Referência do INEP, a fim de constituir um banco de dados para compor um simulado, cuja avaliação é medida à luz da TRI.

5.1 Itens Elaborados

5.1.1 Item 1

Contempla a Habilidade H22 - Identificar informações apresentadas em tabelas ou gráficos (de coluna, de setores e de linha).

A microcefalia é uma condição em que a cabeça de um bebê é significativamente menor que o esperado, muitas vezes devido ao desenvolvimento anormal do cérebro. A microcefalia pode ter como causa diversos agentes infecciosos além do zika como sífilis, toxoplasmose e outros agentes infecciosos. Entre os anos de 2015 e 2016, houve um aumento assustador no número de casos diagnosticados. Até março de 2016, o Ministério da Saúde havia notificado 4293 casos sobre investigação suspeitos de microcefalia e outras alterações do sistema nervoso, sugestivos de infecção congênita. A região Nordeste apresentou o maior número de casos em investigação.

O estado com maior razão entre o número de casos confirmados e casos descartados é

- a) Bahia.
- b) Maranhão.
- c) Piauí.
- d) Rio Grande do Norte.
- e) Sergipe.

Distribuição dos casos notificados de microcefalia por UF, até 19 de março de 2016				
Regiões e Unidades Federadas	Casos de Microcefalia e/ou malformações, sugestivos de infecção congênita			Total acumulado de casos notificados de 2015 a 2016
	Em investigação	Confirmados	Descartados	
Brasil	4 293	907	1 471	6 671
Alagoas	103	41	106	250
Bahia	670	170	120	960
Ceará	249	68	100	417
Maranhão	146	53	31	230
Paraíba	417	91	334	842
Pernambuco	1 210	268	341	1 819
Piauí	52	62	31	145
Rio Grande do Norte	290	81	35	406
Sergipe	161	26	14	201
Região Nordeste	3 298	860	1 112	5 270

Fonte: www.tribunafeirense.com.br Acesso em 25/07/2018 (Adaptado).

Figura 5.1: Casos notificados de microcefalia no Brasil

5.1.2 Item 2

Contempla a Habilidade H10 - Estabelecer relações entre diferentes unidades de medida (comprimento, massa, capacidade, área, volume).

Para melhorar a infraestrutura e se adequar as normas de vigilância sanitária, o proprietário de um parque aquático deseja construir um lava-pés no acesso a principal piscina do seu empreendimento. O ambiente terá as seguintes dimensões 2 m de comprimento, 1,75 m de largura e mantendo o nível da água em 20 cm de altura. Pelas normas sanitárias, o lava-pés deve conter uma torneira e um ralo, ser esvaziado e lavado diariamente além da realização da medição de cloro residual, a cada duas horas, conservando 25 mg/L.

Ao encher o lava-pés logo após ter sido lavado, quanto de cloro residual, em gramas, deve ser acrescentado a água para garantir a solicitação da Norma Sanitária?

- a) 2,8 g
- b) 17,5 g
- c) 98,75 g
- d) 437,5 g
- e) 1 750 g



Fonte: http://www.prefeitura.sp.gov.br/cidade/secretarias/upload/saude/vigilancia_em_saude/arquivos/Praticas_Sanitarias_nas_Piscinas_dos_CEU.pdf. Acesso em 25/07/2018 (Adaptado).

Figura 5.2: Exemplo de lava-pés

5.1.3 Item 3

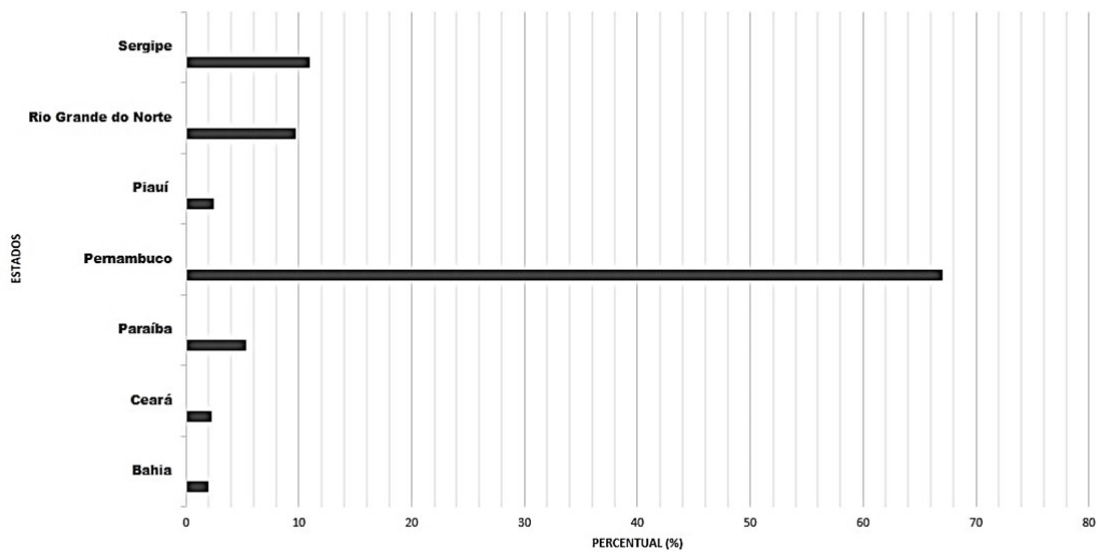
Contempla a Habilidade H24 - Resolver situação-problema com dados apresentados em forma de tabela de dupla entrada ou gráfico.

Em novembro de 2015, o Ministério da Saúde informou que os casos de contaminação pelo vírus da zika registrados no primeiro semestre deste mesmo ano, são "principal hipótese" para explicar o aumento de ocorrência de microcefalia na região Nordeste. Maierovitch disse que a relação entre o vírus da zika e a microcefalia "é inédita no mundo" e não consta na literatura científica até o momento. A relação entre o vírus da zika e a má-formação genética ganhou força porque o micro-organismo foi identificado em duas gestantes da Paraíba. O boletim epidemiológico com dados reunidos até o dia 16/11/2015, aponta a ocorrência de 399 casos em 2015, em sete estados da região Nordeste.

O número de casos de microcefalia registrados no estado da Paraíba foi de, aproximadamente

- a) 9.
- b) 16.
- c) 21.
- d) 39.

Casos de microcefalia registrados em sete estados da região Nordeste



Fonte: <http://g1.globo.com/bemestar/noticia/2015/11/virus-zika-e-principal-hipotese-para-o-aumento-de-microcefalia--diz-saude.html>. Acesso em 26/08/2018 (adaptado)

Figura 5.3: Casos de microcefalia registrados em sete estados da região Nordeste

e) 44.

5.1.4 Item 4

Contempla a Habilidade H3 - Resolver situação-problema com números naturais, inteiros racionais ou reais envolvendo significados da adição, subtração, multiplicação ou divisão, potenciação ou radiciação.

Luiz tem o hobby de colecionar figurinhas do álbum oficial da copa do mundo. Nesta copa, ele comprou o pacote com cinco figurinhas por R\$ 2,00, na banca próxima a sua casa, em Maceió. Para completar seu álbum ele precisa de uma figura rara, que encontrou a venda na internet por um colecionador suíço. O colecionador está vendendo a figurinha por 750% do valor comercializado na Suíça, onde cada pacote de figurinhas é vendido por R\$ 6,00. Desconsiderando os custos com transporte e taxa de importação, quanto custará a figura?

- a) R\$ 1,25
- b) R\$ 3,00
- c) R\$ 4,50
- d) R\$ 9,00
- e) R\$ 18,00

5.1.5 Item 5

Contempla a Habilidade H27 - Calcular a média aritmética de um conjunto de dados expressos em uma tabela de frequências de dados agrupados (não em classes) ou gráficos de colunas.

Todo sábado ocorre uma feira na comunidade onde Emanuel mora. A principal característica da feira é a diversidade dos produtos comercializados que rende fama e grande público. Emanuel pretende comercializar seus produtos na feira e para verificar a viabilidade do seu produto ele resolveu entrevistar alguns dos frequentadores para saber sobre que produto veio adquirir na feira ou prioridade de compra. No mês de maio de 2018, Emanuel realizou a pesquisa e organizou o número de entrevistados por segmento numa tabela.

<i>Data/segmento</i>	<i>Alimentação</i>	<i>Utensílios</i>	<i>Vestuário</i>
<i>05/05/2018</i>	198	75	104
<i>12/05/2018</i>	207	51	79
<i>19/05/2018</i>	153	63	132
<i>26/05/2018</i>	136	46	76

Segundo os dados coletados, qual a média de pessoas entrevistadas em cada sábado?

- a) 330
- b) 348
- c) 377
- d) 391
- e) 440

5.1.6 Item 6

Contempla a Habilidade H28 - Resolver situação-problema que envolva processos de contagem ou noções de probabilidade.

Um jogo de dominó contém 28 peças, conhecidas por "pedras", cada uma dividida ao meio e apresentando ou não de um a seis pontinhos, em cada lado da peça. As pedras que apresentam a mesma quantidade de pontinhos nas duas metades são chamadas de carroças.

Qual a probabilidade de ao escolher ao acaso uma pedra de dominó, sair uma carroça?

- a) $\frac{1}{28}$

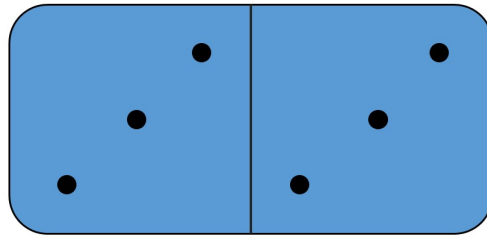


Figura 5.4: Carroça de terno

- b) $\frac{1}{7}$
- c) $\frac{1}{6}$
- d) $\frac{3}{14}$
- e) $\frac{1}{4}$

5.1.7 Item 7

Contempla a Habilidade H10 - Estabelecer relações entre diferentes unidades de medida (comprimento, massa, capacidade, área, volume).

Na aula de educação física o professor propôs uma caça ao tesouro. Ele dividiu a turma em equipes e entregou, a cada uma, um mapa com orientações de como chegar ao tesouro. A equipe A recebeu as seguintes instruções:

Partindo da árvore ao lado do refeitório, dê 25 passos no sentido da quadra de esportes, depois mais 18 passos no sentido da biblioteca, mais 33 passos no sentido dos banheiros e para chegarem ao destino dê mais 9 passos no sentido do auditório.

Pelo Sistema Internacional de Unidades, um passo equivale a 82 cm. Portanto a distância percorrida pela equipe A, em metros, é de:

- a) 62,3.
- b) 69,7.

- c) 85,0.
- d) 103,6.
- e) 136,5.

5.1.8 Item 8

Contempla a Habilidade H29 - Utilizar médias aritméticas, noção de probabilidade ou conhecimentos estatísticos como recurso para a construção de argumentação.

A premiação de uma gincana estudantil é ofertada a partir do empenho e participação das equipes nas provas sugeridas. A premiação consiste na atribuição de 1 ou 2 pontos, segundo a norma: As equipes com pontuação igual ou superior a mediana de todas as pontuações das equipes receberam 2 pontos extras nas disciplinas cursadas e os que obtiveram pontuação inferior a mediana receberam 1 ponto extra.

Resultado da gincana estudantil	
Equipes	Pontuação
1A	445
1B	316
1C	280
2A	302
2B	420
2C	360
3A	405
3B	320

A partir dos dados observados, podemos afirmar que:

- a) A equipe 1 A recebeu 2 pontos e a equipe 2 C, 1 ponto.
- b) A equipe 3 B recebeu 1 ponto e a equipe 1 B, 2 pontos.
- c) As equipes 2 B e 2 C receberam ambas 2 pontos.
- d) As equipes 1 C e 2 B receberam ambas 2 pontos.
- e) As equipes 2 A e 3 A receberam ambas 1 ponto.

5.1.9 Item 9

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

Um objeto radioativo foi abandonado num terreno baldio, o qual foi representado num plano cartesiano pelo ponto A. Devido ao alto risco de contaminação pela radiação, uma medida de segurança foi tomada, isolando uma região circular em torno do objeto, representada pela equação da circunferência dada por $x^2 + y^2 - 8x - 4y + 11 = 0$. Todas as pessoas que estiverem nesta região serão submetidas a uma quarentena e passarão por alguns exames para investigação de possível contaminação.

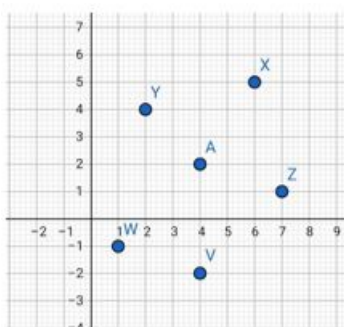


Figura 5.5: Localização dos pontos no plano cartesiano - autor

No gráfico também estão representadas cinco pessoas nas suas respectivas posições no momento da averiguação, será submetido a quarentena aquele que está em:

- a) V.
- b) W.
- c) X.
- d) Y.
- e) Z.

5.1.10 Item 10

Contempla a Habilidade H3 - Resolver situação-problema com números naturais, inteiros racionais ou reais envolvendo significados da adição, subtração, multiplicação ou divisão, potenciação ou radiciação.

Sofia deseja comprar uma moto, para tanto ela tem em mãos R\$ 1 000,00 que foi um presente do seu avô e pretende completar o valor poupando dinheiro. No dia primeiro de

junho ela colocou em seu cofrinho R\$ 1,00; no segundo dia ela colocou R\$ 2,00; no terceiro R\$ 3,00 e continuou até o n-ésimo dia, com R\$ n,00. A fim de poupar o valor desejado. (Considere todos os meses com trinta dias) Sendo o n-ésimo valor colocado no dia 25 de agosto, quanto Sofia pagou pela moto?

- a) R\$ 1 850,00
- b) R\$ 2 730,00
- c) R\$ 3 570,00
- d) R\$ 4 655,00
- e) R\$ 9 400,00

5.1.11 Item 11

Contempla a Habilidade H14 - Avaliar propostas de intervenção na realidade utilizando cálculos de perímetros, área de superfícies planas ou volume de blocos retangulares.

Carlos foi vítima de um acidente de moto que lhe causou uma lesão nos membros inferiores, impossibilitando-o de andar, sendo necessário o uso da cadeira de rodas. Priorizando a acessibilidade, a família de Carlos está reformando a casa. Uma das melhorias é eliminação de uma série de degraus construindo no lugar uma rampa. São ao todo 20 degraus, revestidos com um piso cerâmico, que serão substituídos por uma rampa revestida por um piso antiderrapante.

Quantos metros quadrados de revestimento antiderrapante serão necessários para revestir toda a rampa? (Considere $\sqrt{5} = 2,2$)

- a) 7,92
- b) 13,20
- c) 18,00
- d) 23,76
- e) 27,00

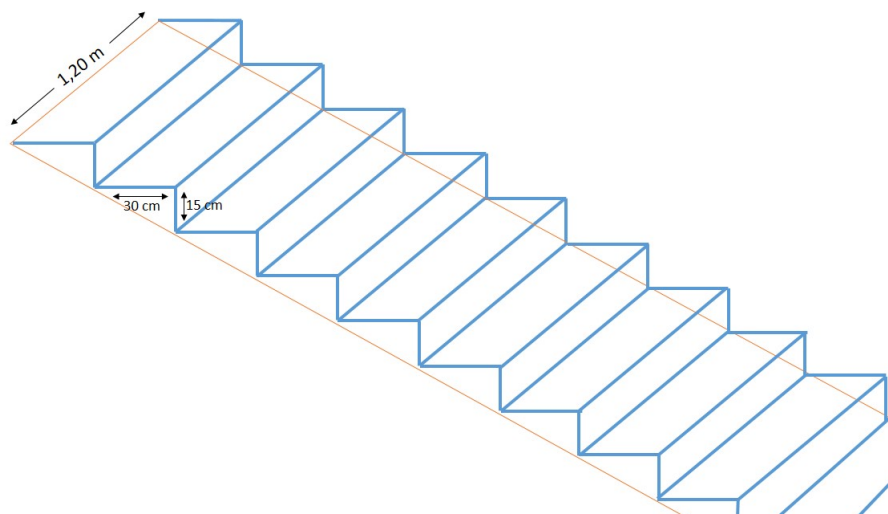


Figura 5.6: autor

5.1.12 Item 12

Contempla a Habilidade H11 - Aplicar a noção de escalas na leitura de plantas ou mapas.

A família Gonçalves precisa se deslocar do interior para a capital de Pernambuco, Recife, e usou um aplicativo para definir uma rota. No aplicativo a rota está representada na escala 1 : 3 300 000, e tem uma extensão de 6,5 cm. A distância a ser percorrida, em quilômetros, segundo o aplicativo é de:

- a) 196,9.
- b) 214,5.
- c) 5 076,9.
- d) 6 500.
- e) 33 000.

5.1.13 Item 13

Contempla a Habilidade H2 - Utilizar algum procedimento de cálculo com números naturais, inteiros, racionais ou reais.

A adoção do fracionamento das vacinas é uma medida preventiva e recomendada pela Organização Mundial de Saúde (OMS) quando há um aumento de epizootias e casos de

febre amarela silvestre de forma intensa, com risco de expansão da doença em cidades com elevado índice populacional. A dose padrão da vacina contra febre amarela protege uma pessoa por toda a vida, enquanto a fracionada dura pelo menos oito anos. O fracionamento nada mais é do que a aplicação de uma dose menor do que a prescrita tradicionalmente. Dessa forma, um frasco com cinco doses da vacina padrão pode imunizar até 25 brasileiros com esse método. Ao todo, 19,7 milhões de pessoas devem ser imunizadas nos estados de São Paulo, Rio de Janeiro e Bahia contra a febre amarela. Sendo 15 milhões com doses fracionadas e 4,7 milhões com a concentração padrão. Quantos frascos, no mínimo, devem ser utilizados durante a campanha de imunização contra a febre amarela nos três estados?

- a) 788 000
- b) 980 000
- c) 1 540 000
- d) 3 940 000
- e) 5 300 000

5.1.14 Item 14

Contempla a Habilidade H8 - Resolver situação-problema que envolva noções geométricas (ângulo, paralelismo, perpendicularismo).

A pirâmide de base quadrangular representada na figura 5.7 foi seccionada paralelamente ao plano que contém a sua base, determinando uma região poligonal semelhante a essa base. A secção foi realizada a partir do vértice representando $\frac{3}{8}$ da altura da pirâmide.

A aresta da base da pirâmide é dada por 36 cm, portanto a medida da aresta da região poligonal determinada pela secção é de:

- a) 12 cm.
- b) 13,5 cm.
- c) 22,5 cm.
- d) 49,5 cm.
- e) 96 cm.

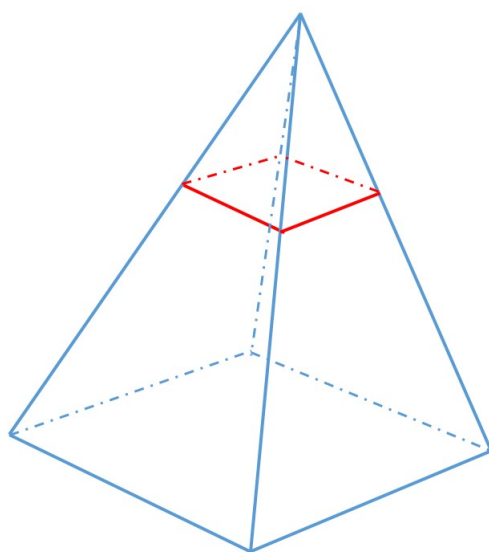


Figura 5.7: Pirâmide de base quadrangular - autor

5.1.15 Item 15

Contempla a Habilidade H3 - Resolver situação-problema com números naturais, inteiros racionais ou reais envolvendo significados da adição, subtração, multiplicação ou divisão, potenciação ou radiciação.

Sandra tem um aquário e realiza periodicamente a sua manutenção que consiste em retirar parte da água presente no aquário e repor com água da torneira. O uso da água da torneira pode causar danos à saúde dos peixes, então ela utiliza um medicamento que condiciona água da torneira para o uso em aquários.

Instruções de uso:
Adicionar 3 mL para cada 24 Litros de água ou 2 gotas para cada litro de água.

Segundo as instruções de uso do medicamento, podemos estabelecer que cada mL equivale a quantas gotas?

- a) 16
- b) 24
- c) 36
- d) 48
- e) 72

5.1.16 Item 16

Contempla a Habilidade H22 - Identificar informações apresentadas em tabelas ou gráficos (de coluna, de setores e de linha).

Doze mulheres são assassinadas todos os dias, em média, no Brasil. É o que mostra um levantamento feito pelo G1 considerando os dados oficiais dos estados em 2017. São 4.473 homicídios dolosos, sendo 946 feminicídios, ou seja, casos de mulheres mortas em crimes de ódio motivados pela condição de gênero. Os dados da pesquisa estão representados a seguir:

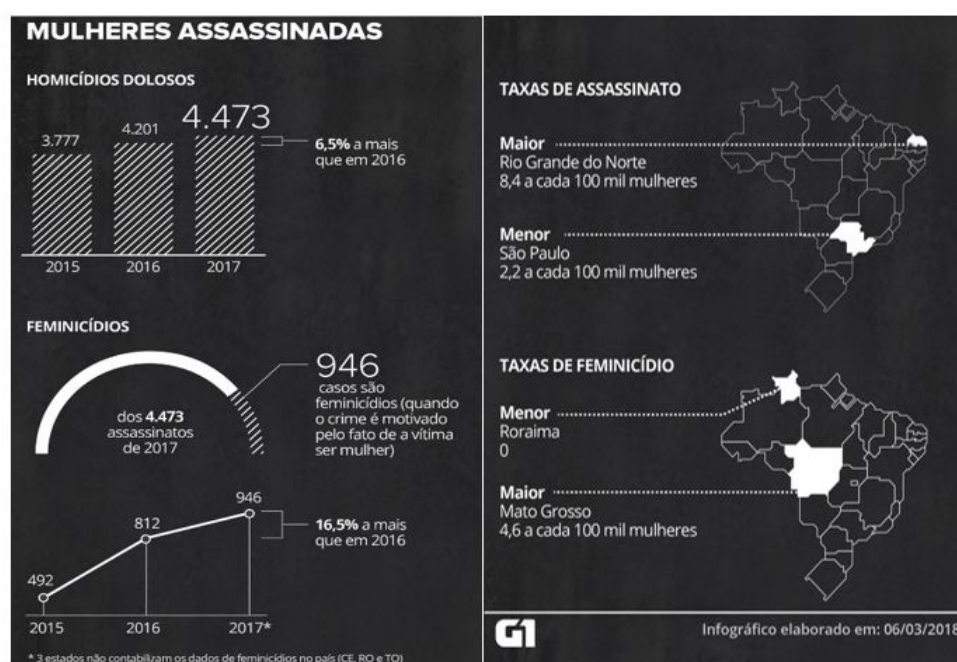


Figura 5.8: Dados sobre os Homicídios contra mulheres

Fonte: <https://g1.globo.com/monitor-da-violencia/noticia/cresce-n-de-mulheres-vitimas-de-homicidio-no-brasil-dados-de-feminicidio-sao-subnotificados.ghtml> Acesso em 05/08/18.

Com base nos dados apresentados anteriormente, podemos observar que:

- O estado de São Paulo apresenta menor taxa de feminicídio do país, com 2,2 casos a cada 100 mil mulheres.
- O número de feminicídios no Brasil, em 2017, aumentou em 454 casos em relação a 2016.
- O número de homicídios dolosos em 2017 foram de 6,5% a mais que em 2015.
- Os casos de feminicídio representam 21% dos assassinatos de mulheres em 2017.
- O estado me Mato Grosso apresenta a maior taxa de feminicídio, com 8,4 casos a cada 100 mil mulheres.

5.1.17 Item 17

Contempla a Habilidade H24 - Resolver situação-problema com dados apresentados em forma de tabela de dupla entrada ou gráfico.

O ar atmosférico tem em sua composição 21% de oxigênio. A pressão atmosférica, ao nível do mar, é de 760 mmHg (milímetros de mercúrio), o que corresponde a uma pressão parcial de oxigênio de 159,6 mmHg (21% de 760). A disponibilidade do oxigênio depende de sua pressão parcial na atmosfera. Podemos observar nos dados relacionados no quadro a seguir que a pressão atmosférica diminui à medida que a altitude se eleva, a pressão parcial (e a disponibilidade) do oxigênio também se reduz parcialmente.

Pressão atmosférica em função da altitude	
Altitude (m)	Pressão (mmHg)
900	681
1000	668
1200	656
1350	644
1500	632

Rafael pretende visitar o ponto mais alto do Estado de Pernambuco, a Serra da Boa Vista (Serra do Ponto) no município de Brejo da Madre de Deus, com altitude de 1195 metros conforme consta no Anuário Estatístico do Brasil (IBGE 2010). Para tanto, ele verificou a pressão parcial de oxigênio presente no ar atmosférico na Serra da Boa Vista.

Fonte: <http://www.estacaonoticias.com.br/2011/06/o-ponto-mais-alto-de-pernambuco-fica-em.html>
<http://www.cives.ufrj.br/informacao/altitude/altitude-iv.html> Acesso em 05/08/2018.

A variação da pressão parcial de oxigênio (em mmHg) na Serra da Boa Vista em relação ao nível do mar é de

- a) 16,54.
- b) 19,32.
- c) 21,84.
- d) 24,36.
- e) 26,88.

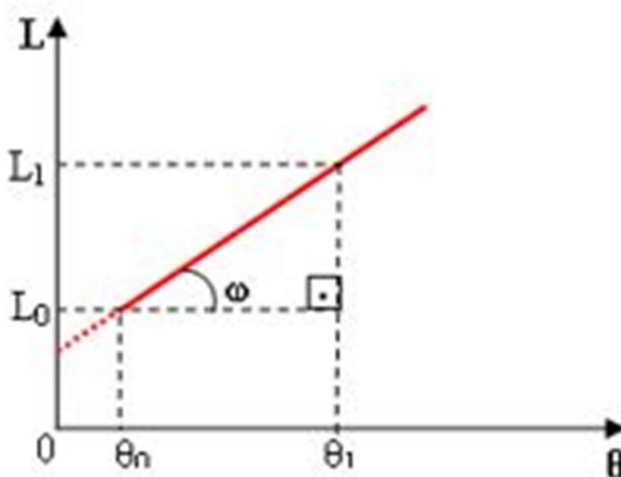
5.1.18 Item 18

Contempla a Habilidade H21 - Resolver situação-problema cujos dados estejam expressos em gráfico cartesiano que mostre a variação de duas grandezas.

A dilatação térmica é um dos efeitos da variação da temperatura, ocasionando a variação de dimensões em corpos sólidos e líquidos. Ao considerarmos uma barra homogênea, por exemplo, de comprimento L_0 a uma temperatura inicial θ_0 . Quando esta temperatura é aumentada até uma $\theta(> \theta_0)$, observa-se que esta barra passa a ter um comprimento $L(> L_0)$. Representamos a variação do comprimento por:

$$\Delta L = L_0 \cdot \alpha \cdot \Delta \theta$$

. A dilatação linear ocorre de maneira proporcional à variação de temperatura e ao comprimento inicial. Podemos expressar a dilatação linear de um corpo através de um gráfico de seu comprimento (L) em função da temperatura (θ), desta forma:



Fonte: <https://www.sofisica.com.br/conteudos/Termologia/Dilatacao/linear.php>. Acesso em 05/08/2018

Figura 5.9: Variação de comprimento em função da variação de temperatura

Considerando um ângulo φ como a inclinação da reta em relação ao eixo horizontal. Podemos estabelecer a relação

- a) $\tan \varphi = L_0 \cdot \alpha$
- b) $\tan \varphi = \Delta L \cdot \alpha$
- c) $\tan \varphi = \Delta \cdot \alpha$
- d) $\tan \varphi = \frac{\Delta L}{\alpha}$
- e) $\tan \varphi = \frac{L_0}{\Delta L}$

5.1.19 Item 19

Contempla a Habilidade H19 - Identificar representações algébricas que expressem a relação de interdependência entre duas grandezas.

Bethânia comercializa doces e salgados na sua vizinhança. O valor do produto comercializado depende do valor total gasto na produção e do lucro esperado. Para fazer uma certa quantidade de sonhos ela gasta R\$ 50,00 e o comercializa no valor de R\$ 3,00 cada um. Qual expressão abaixo representa a relação entre o Lucro L , em reais, em função dos sonhos vendidos, representado por x ?

- a) $L = 50 - 3x$
- b) $L = 3 - 50x$
- c) $L = 50x - 3$
- d) $L = 3x$
- e) $L = 3x - 50$

5.1.20 Item 20

Contempla a Habilidade H7 - Identificar características de polígonos ou sólidos (prismas, pirâmides, cilindros).

Na figura 5.10, temos um cubo onde a partir dos vértices A, B, C e D foram traçados segmentos ligando-os ao ponto N, e a partir dos vértices E, F, G, H, segmentos ligando-os ao ponto M. Sendo M e N, pontos de intercessão das diagonais das faces ABCD e EFGH, respectivamente.

Um poliedro é formado a partir dos segmentos que partem dos pontos M e N até os pontos onde os segmentos traçados anteriormente se interceptam. O poliedro formado é um

- a) tetraedro.
- b) hexaedro.
- c) octaedro.
- d) decaedro.
- e) dodecaedro.

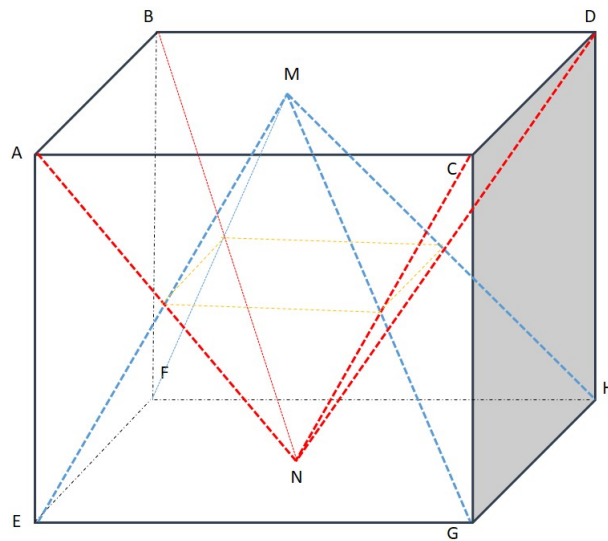


Figura 5.10: Cubo - autor

5.1.21 Pesquisa exploratória

O teste constituído pelos 20 itens formulados e apresentados neste trabalho foram submetidos a oitenta respondentes, alunos de duas turmas de 3^o ano (C e D) do Ensino Médio, da Escola de Referência em Ensino Médio André Cordeiro, na cidade de Brejo da Madre de Deus, Pernambuco. A fim de observar as características dos itens, como o poder de discriminação e a dificuldade, por exemplo. Para tanto, os resultados serão analisados pela TCT e TRI, utilizando o Ambiente Computacional e Estatístico R no processo de calibração dos itens.

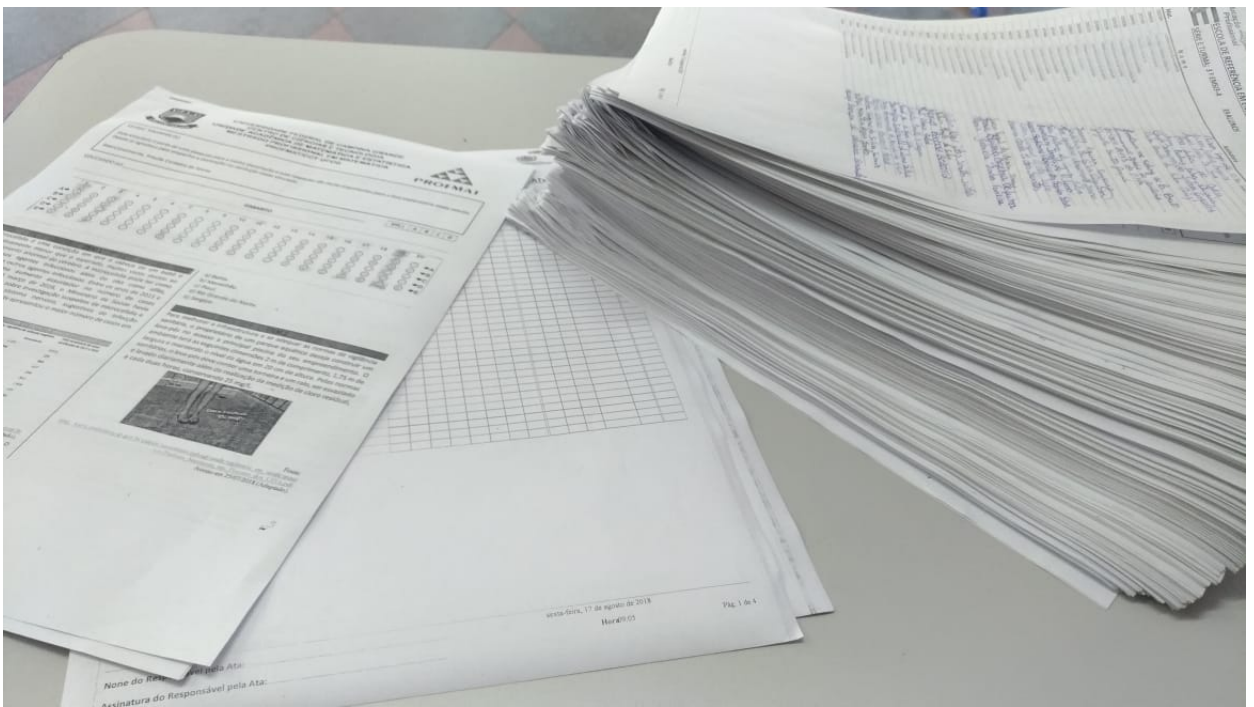


Figura 5.11: Simulados - Arquivo pessoal

Capítulo 6

Análise dos itens e do teste usando o Ambiente Computacional e Estatístico *R*

Nesta seção está apresentado os resultados obtidos na aplicação do simulado elaborado com os 20 itens apresentados no capítulo anterior. Apresenta-se a análise dos dados na TCT e na TRI, comparando-os, quando possível.

Após aplicado o simulado, onde os itens analisados são dicotômicos, uma correção foi realizada de modo que a resposta de cada pessoa para cada item é indicada por "1" para uma resposta correta e "0" por uma resposta incorreta.

As **Figuras 6.1 e 6.2** mostram uma matriz de dados simples contendo as respostas dos oitenta respondentes para o teste de 20 itens. Os 20 itens são nomeados no topo da matriz. As oitenta pessoas estão numeradas à esquerda. Observe que as respostas foram somadas entre os itens e inseridas no lado direito da matriz como o escore total de cada respondente e inseridas na parte inferior da matriz como escore total de cada item.

Respondente	Matriz de datos observados dos respondentes																				Escore
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	0	1	0	0	1	0	0	1	0	0	1	0	0	0	1	1	0	0	0	1	8
2	0	0	0	1	1	1	1	1	0	0	1	0	1	1	1	0	0	0	1	0	10
3	0	1	0	0	1	1	1	1	1	1	0	1	1	1	0	1	0	0	1	1	11
4	0	1	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	7
5	0	0	1	1	1	1	0	1	1	0	0	1	0	0	1	0	1	0	0	1	10
6	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	1	5
7	0	1	0	0	0	0	1	0	0	1	1	1	0	0	1	0	0	0	1	1	7
8	1	1	1	1	1	1	0	1	0	0	1	0	0	0	1	1	0	0	0	0	9
9	0	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	4
10	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	1	0	0	6
11	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	0	1	10
12	0	0	0	1	1	0	1	1	0	0	1	0	0	0	0	0	0	0	1	1	7
13	0	1	0	1	1	1	0	1	1	0	0	1	1	0	0	0	0	0	0	0	7
14	1	1	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	6
15	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	1	1	0	0	1	5
16	0	1	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	5
17	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	3
18	1	1	0	1	0	1	1	1	1	0	0	0	0	0	1	1	0	1	0	0	11
19	0	1	0	1	1	0	1	1	1	0	1	0	0	0	1	0	0	1	0	0	8
20	0	1	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	1	0	0	4
21	0	1	0	1	1	1	0	1	1	0	1	0	0	0	1	0	0	1	0	1	10
22	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	1	0	0	6
23	1	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	4
24	0	1	1	1	1	1	0	1	0	1	0	1	0	0	0	0	1	1	0	1	10
25	0	0	0	1	1	1	0	1	1	0	0	0	0	0	1	0	1	1	0	1	9
26	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	1	5
27	0	1	0	0	0	0	0	1	0	1	1	1	0	0	1	1	0	0	1	1	9
28	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0	4
29	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	5
30	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	4
31	1	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	1	1	0	0	6
32	0	1	0	0	1	1	1	1	0	0	0	1	1	0	1	0	1	1	0	0	11
33	0	1	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	1	1	6
34	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	4
35	0	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	1	5
36	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	3
37	0	0	0	1	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	5
38	0	1	0	1	0	0	1	1	0	1	0	0	1	0	0	0	0	0	1	0	7
39	0	1	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	4

Figura 6.1: Matriz de Resultados

40	1	1	1	0	1	0	1	0	0	0	1	1	0	1	0	0	0	1	0	
41	0	0	0	0	1	0	1	1	0	0	1	0	1	0	1	0	1	1	0	
42	0	1	0	0	0	0	1	0	1	1	0	0	0	1	1	0	0	0	0	
43	0	1	0	0	1	0	1	0	0	1	0	1	0	1	1	0	1	0	0	
44	0	0	1	0	1	0	1	0	1	0	0	0	0	1	0	0	1	0	0	
45	0	0	0	0	1	0	1	0	0	1	0	1	1	0	0	0	0	0	1	
46	1	1	0	1	1	0	1	1	0	1	0	0	1	0	1	1	1	0	0	
47	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	1	1	0	1	
48	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	0	
49	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0	
50	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	
51	0	0	1	0	1	0	1	0	1	1	0	1	1	1	1	1	1	0	0	
52	0	0	1	1	1	0	1	0	0	0	0	1	0	0	0	0	1	0	1	
53	0	1	1	1	1	0	0	1	0	1	0	0	1	0	0	0	0	1	1	
54	0	1	0	1	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	
55	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	1	
56	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	
57	0	1	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	
58	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	
59	1	0	0	0	1	0	1	1	0	1	0	0	0	0	1	0	0	0	0	
60	0	0	0	0	0	0	0	1	0	0	0	1	1	1	1	0	1	0	1	
61	0	1	1	1	1	0	0	1	0	1	0	0	1	1	0	0	0	0	0	
62	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0	
63	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	
64	0	0	0	0	0	0	1	1	0	0	1	0	0	1	1	1	0	0	1	
65	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	
66	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
67	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1	0	
68	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	
69	0	1	0	0	0	0	1	1	1	0	0	1	0	0	0	0	0	0	0	
70	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	1	
71	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	1	1	
72	0	0	0	0	1	0	1	0	0	0	0	1	0	1	0	0	0	0	0	
73	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	
74	0	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0	0	
75	0	0	0	1	0	0	1	1	1	0	1	1	0	1	1	0	0	1	0	
76	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	1	
77	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	
78	0	1	0	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	
79	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	1	0	0	
80	0	0	0	1	1	0	1	0	1	1	0	0	0	0	1	0	1	0	0	
	11	38	12	29	38	10	56	37	22	22	13	24	22	28	29	16	22	24	29	37

Figura 6.2: Matriz de Resultados

6.1 Leitura e visualização breve da base de dados

A partir dos dados apresentados na matriz de resultados (**Figura 6.1 e 6.2**) utiliza-se o ambiente *R* para realização das estimações.

Iniciando o Ambiente Computacional e Estatístico R

Iniciar as análises inserindo o arquivo de dados apresentado no formato .txt no *R*, através dos comandos:

```
> #Leitura de uma base de dados armazenado num arquivo .txt
> data = read.table("C:/Users/Gilberto/Dropbox
(Pessoal)/04_Orientações/17_PROFMAT
- Priscila - Rasch model/Análise dos Itens/Análise Clássica/dados.txt",
+ header = T)
> # Visualiza as primeiras linhas da base de dados
> head(data)
```

O comando *head(data)* exibe a matriz inicial de dados com as respostas obtidas seguir

Num	Item1	Item2	Item3	Item4	Item5	Item6	Item7	Item8	Item9	Item10	Item11	
1	1	0	1	0	1	0	0	1	0	0	1	0
2	2	0	0	0	1	1	1	1	0	0	1	0
3	3	0	1	0	0	1	1	1	1	1	0	0
4	4	0	1	0	0	1	1	1	1	1	0	0
5	5	0	0	1	1	1	0	1	1	0	0	1
6	6	0	1	0	0	0	0	0	1	0	0	0

	Item12	Item13	Item14	Item15	Item16	Item17	Item18	Item19	Item20
1	0	0	1	1	1	0	0	0	1
2	1	1	1	1	0	0	0	1	0
3	1	1	0	0	1	0	0	1	1
4	0	0	0	0	0	0	1	0	0
5	0	0	1	0	1	0	0	1	1
6	0	0	0	1	0	0	0	1	1

Figura 6.3: Início da matriz de dados

Para iniciar as análises dos resultados, excluí a coluna que identifica os alunos, utilizando o comando a seguir, ficando apenas as respostas aos itens.

```
> respostas = data[, -1]
```

6.2 Análise Clássica dos Itens e do Teste

Para realizar a análise Clássica descritiva, foi utilizado o pacote "ltm" do ambiente R.

6.2.1 Instalação do pacote "ltm" e análise pela TCT

Instala-se o pacote *ltm* pelo comando:

```
install.packages("ltm"),
```

Uma vez instalado o pacote *ltm*, a função *descript* realiza uma análise clássica dos dados. como executado a seguir

Este comando gera uma matriz com os percentuais de acertos dos itens.

```
> analise.classica$perc
```

Observando-se a saída do R acima (veja Tabela 6.2.1), verifica-se que os itens com os menores percentuais de acerto, ou seja, os considerados mais difíceis foram os itens 6 e 1, com percentuais 12,5% e 13,75%, respectivamente. Os Itens 3 e 11 também podem ser considerados difíceis, pois apresentam nível de dificuldade 15% e 16,25%, respectivamente. O item 7 apresenta o maior percentual de acerto, 70%, o que indica ser o item mais fácil do teste.

Análise Clássica - Representação decimal

	0	1	logit
Item 1	0.8625	0.1375	-1.8362112
Item 2	0.5250	0.4750	-0.1000835
Item 3	0.8500	0.1500	-1.7346011
Item 4	0.6375	0.3625	-0.5645298
Item 5	0.5250	0.4750	-0.1000835
Item 6	0.8750	0.1250	-1.9459101
Item 7	0.3000	0.7000	0.8472979
Item 8	0.5375	0.4625	-0.1502822
Item 9	0.7250	0.2750	-0.9694006
Item 10	0.7250	0.2750	-0.9694006
Item 11	0.8375	0.1625	-1.6397433
Item 12	0.7000	0.3000	-0.8472979
Item 13	0.7250	0.2750	-0.9694006
Item 14	0.6500	0.3500	-0.6190392
Item 15	0.6375	0.3625	-0.5645298
Item 16	0.8000	0.2000	-1.3862944
Item 17	0.7250	0.2750	-0.9694006
Item 18	0.7000	0.3000	-0.8472979
Item 19	0.6375	0.3625	-0.5645298
Item 20	0.5625	0.4375	-0.2513144

Tabela 6.1: Acertos por item - Representação decimal

Discriminação dos itens

Item 1	0.053023095
Item 2	0.359111277
Item 3	0.450272846
Item 4	0.403747957
Item 5	0.424531679
Item 6	0.358087072
Item 7	0.398206857
Item 8	0.354643881
Item 9	0.019075202
Item 10	0.447613983
Item 11	0.202256781
Item 12	0.396169993
Item 13	0.259475006
Item 14	0.440804195
Item 15	0.413456354
Item 16	0.368696245
Item 17	0.175857683
Item 18	0.009165887
Item 19	0.219288415
Item 20	0.224024917

Tabela 6.2: Discriminação dos itens - Coeficiente de correlação Bisserial


```

> # Carrega o pacote "ltm"
> require(ltm)

> # O comando "descript" do pacote ltm executa e cria o objeto com a
análise clássica
> analise.classica = descript(respostas)

> # Exibi-se os nomes dos elementos do objeto/resultado da análise
clássica
> names(analise.classica)

[1] "sample"    "perc"      "items"     "pw.ass"    "n.print"
"name"
[7] "missin"    "data"      "bisCorr"   "ExBisCorr" "alpha"

> # Percentual de acertos aos itens
> analise.classica$perc

```

A discriminação é calculada a partir do coeficiente de correlação bisserial (3.2). Observa-se na Tabela 6.2.1 que os Itens 18, 9 e 1 possuem fortes problemas pois os mesmos apresentaram valores muito próximos de zero.

Destaca-se que o Item 1 como problemático pois o mesmo além de não discriminar, com valor 0,05 (veja Tabela 6.2.1), também é considerado muito difícil, 13,75% de acertos (veja Tabela 6.2.1) O comando a seguir determina o valor do coeficiente alfa de Cronbach (3.3), determinando a confiabilidade para cada item analisado.

```
> analise.classica$alpha
```

O coeficiente alfa de Cronbach (3.3) é uma medida estatística que revela o quanto a prova mede de forma precisa o que se propõe a medir, o qual pode assumir um valor dentro do intervalo $[0, 1]$. Quanto mais próximo de 1, essa medida indica alta confiabilidade e o teste é considerado aceitável a partir de 0,70. Para a presente prova, observa-se na Tabela 6.2.1 o valor igual a 0,466, o que nos indica uma confiabilidade não muito boa, porém moderada.

Ao calcular outras confiabilidades, excluindo cada item da prova, espera-se que ao se excluir um item, o coeficiente alfa de Cronbach não aumente muito com referência ao valor global, pois caso isto aconteça, temos a indicação que tal item está diminuindo a confiabilidade da prova e, portanto, tal item apresenta problema.

Ao realizar essa análise no simulado proposto, observa-se que os itens 9 e 18, são itens que aumentam um pouco a confiabilidade quando são excluídos da prova, entretanto, tais aumentos não são tão significativos. É bom reforçar, que estes dois itens também apresentaram problemas com relação à discriminação.

```
> analise.classica$items
```

Coeficiente alfa de Cronbach	
Todos os itens	0.4666695
Excluindo Item 1	0.4870631
Excluindo Item 2	0.4440280
Excluindo Item 3	0.4219149
Excluindo Item 4	0.4316439
Excluindo Item 5	0.4270070
Excluindo Item 6	0.4389321
Excluindo Item 7	0.4321704
Excluindo Item 8	0.4450960
Excluindo Item 9	0.5090166
Excluindo Item 10	0.4201286
Excluindo Item 11	0.4663405
Excluindo Item 12	0.4326566
Excluindo Item 13	0.4624757
Excluindo Item 14	0.4220750
Excluindo Item 15	0.4291836
Excluindo Item 16	0.4372040
Excluindo Item 17	0.4795370
Excluindo Item 18	0.5130460
Excluindo Item 19	0.4749711
Excluindo Item 20	0.4759619

Tabela 6.3: Coeficiente alfa de Cronbach

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Freq	0	0	1	8	12	16	8	11	4	7	6	4	2	0	0	1	0	0	0	0	0

Figura 6.4: Distribuição dos escores totais dos respondentes

Nos resultados obtidos pelos respondentes neste teste, observa-se notas variando entre 2 e 15 acertos, de um total de vinte itens. Destaca-se que não ocorreu escore nulo (nenhum acerto), como também não existiram respondentes que obtiveram escore total. A maioria acertou entre 3 e 11 questões, como pode ser observado na **Figura 6.4**. Quanto aos escores dos itens, destaca-se que nenhum item teve escore nulo ou escore total.

6.3 Estimação Bayesiana dos Itens e das Habilidades - TRI

Para realização da análise TRI, foi utilizado o pacote *MCMCpack* do Ambiente Computacional e Estatístico *R*.

6.3.1 Instalação do pacote "MCMCpack" e estimação bayesiana via métodos MCMC

Instala-se o pacote *MCMCpack* pelo comando:

```
install.packages("MCMCpack"),
```

Inicia-se com a estimação dos parâmetros dos itens via inferência estatística bayesiana por geração de valores amostrais da distribuição a posteriori pelo método MCMC (Monte Carlo via Cadeias de Markov). Para tanto, utiliza-se a função *MCMCirtKd* do pacote *MCMCpack* do *R* que estima parâmetros de um modelo multidimensional da TRI e que permite ao usuário especificar restrições nos parâmetros de discriminação dos itens. Tal função foi utilizada fixando a dimensão das habilidades em $K = 1$ por estar avaliando preponderantemente o conhecimento em Matemática e, além disso, especificar na função que as discriminações somente assumem valores não negativos, uma vez que valores negativos foram obtidos em outros métodos de estimação, sejam eles por máxima verossimilhança ou bayesianos sem restrição, com indicações de instabilidade na convergência dos algoritmos de estimação (Matos, 2008; ver pdf em <http://www.teses.usp.br/teses/disponiveis/45/45133/tde-21102009-104941/pt-br.php>).

list()		
Item1=list(2,"+")	Item2=list(2,"+")	Item3=list(2,"+")
Item4=list(2,"+")	Item5=list(2,"+")	Item6=list(2,"+")
Item7=list(2,"+")	Item8=list(2,"+")	Item9=list(2,"+")
Item10=list(2,"+")	Item11=list(2,"+")	Item12=list(2,"+")
Item13=list(2,"+")	Item14=list(2,"+")	Item15=list(2,"+")
Item16=list(2,"+")	Item17=list(2,"+")	Item18=list(2,"+")
Item19=list(2,"+")	Item20=list(2,"+")	

Tabela 6.4: Restrição para garantir a não-negatividade das discriminações

É importante destacar que os parâmetros relativos ao acerto casual não foram estimados, pois os algoritmos, tanto de estimação por máxima verossimilhança quanto bayesianos, demonstraram-se extremamente instáveis, não acontecendo convergência e, conseqüentemente, não permitindo a obtenção das estimativas para tais parâmetros.

Uma única vez instalado o pacote *MCMCpack*, podemos executar os comandos:

```
> # Carrega o pacote "MCMCpack"
> require(MCMCpack)
```

Como a discriminação do item não pode admitir valores negativos, então usamos (parâmetros=2,"+"), impondo restrições para garantir não-negatividade das discriminações (veja a **Figura 6.4**).

```
> restricoes <- list()
```

Estimando parâmetros associados às dificuldades e discriminações pelo modelo K-dimensional. Para tanto, fazemos k=1 (uma dimensão preponderante) com restrições de não-negatividade nas discriminações.

```
> bayes1 <- MCMCirtKd(respostas, dimensions=1,
  item.constraints = restricoes,
+                   burnin = 5000, mcmc = 15000, thin = 10,
+                   verbose = 0, seed = NA,
+                   alphabeta.start = NA,
+                   b0 = 0, B0 = 0,
+                   store.item = TRUE,
+                   store.ability = TRUE,
+                   drop.constant.items = TRUE)
```

Dificuldades dos itens (b)

	Dificuldade	Desvio padrão
b_1	1,10585842	0,1771439
b_2	0,07021133	0,1526550
b_3	1,81180370	0,9843905
b_4	0,40312932	0,1767047
b_5	0,07660544	0,1848819
b_6	1,25390941	0,2204695
b_7	-0,60636285	0,1942287
b_8	0,10021490	0,1493875
b_9	0,60205395	0,1503701
b_{10}	0,73875783	0,2287389
b_{11}	1,02503324	0,1749867
b_{12}	0,60389663	0,1880973
b_{13}	0,62497423	0,1590358
b_{14}	0,48439290	0,2098338
b_{15}	0,39132481	0,1688288
b_{16}	0,93466632	0,1963434
b_{17}	0,61214685	0,1526477
b_{18}	0,53139055	0,1518539
b_{19}	0,35782945	0,1502378
b_{20}	0,17055524	0,1504114

A partir do comando a seguir, extraí do objeto do R "bayes1" os parâmetros dos itens estimados e desvios padrões.

```
> resumo = summary(bayes1)
> estatisticas = resumo$statistics
> itens = estatisticas[1:40,1:2]
```

O comando a seguir extraí do objeto a ser estimado no R, as dificuldades dos itens.

```
> dificuldades = itens[seq(1, 40, by=2),]
```

Discriminações dos itens (*a*)

	Discriminação	Desvio padrão
a_1	0.1733144	0.14539531
a_2	0.3361709	0.20608580
a_3	1.5720184	1.30831676
a_4	0.6131702	0.30745206
a_5	0.8445604	0.40340147
a_6	0.4359630	0.28194446
a_7	0.6160026	0.36271429
a_8	0.3264761	0.19714413
a_9	0.1259387	0.10475325
a_{10}	0.7912984	0.41081546
a_{11}	0.2682531	0.20911496
a_{12}	0.6190600	0.34076766
a_{13}	0.3063939	0.19650418
a_{14}	0.8349220	0.39199604
a_{15}	0.4974583	0.30408219
a_{16}	0.4885269	0.28519404
a_{17}	0.2184656	0.16335754
a_{18}	0.0985148	0.09014961
a_{19}	0.2174616	0.15950367
a_{20}	0.2785998	0.19236989

Como podemos observar nos resultados obtidos, as dificuldades estimadas dos itens pela TRI apresentam valores muito parecidos com as dificuldades obtidas pela TCT. Observamos, por exemplo, que os Itens 3, 6, 1 e 11 continuam sendo os mais difíceis; nesta ordem, apresentando dificuldades superiores a 1. Por outro lado, o Item 7 é o mais fácil, pois apresenta uma dificuldade negativa, apesar de não ser muito fácil. Observamos, ainda, que os Itens 2 e 5, assim como na análise clássica, possuem dificuldades mediana, pois seus valores são muito próximos de zero.

Para determinarmos a discriminação dos itens, usamos o comando abaixo:

```
> discriminacoes = items[seq(2, 40, by=2), ]
```

Com relação às discriminações, observamos que pelo modelo da TRI, os Itens 18, 9 e 1; em ordem crescente; são os que menos discriminam, indicando, portanto, algum pro-

blema com relação à eficácia em distinguir respondentes de diferentes habilidades. Ressaltamos que estes itens também foram identificados pela análise clássica de testes como sendo problemáticos quanto à discriminação.

Sendo mais criteriosos, observamos que a maioria dos itens não discriminam bem, pois os mesmos apresentam valores relativamente baixos, por exemplo, entre 0,20 e 0,60. O item 3 aparece como item mais difícil, com $b = 1,81$ (veja Tabela 6.3.1) assim como, é o que mais discrimina $a = 1,57$ (veja Tabela 6.3.1), pois um item difícil necessita de uma habilidade maior para ser respondido corretamente.

Comando para obter as estimativas médias dos respondentes.

```
# Extrai do objeto "estatisticas" as habilidades médias e
devios padrão
> habs = estatisticas[-c(1:40),1:2]
```

Após estimar as habilidades médias dos 80 alunos, observamos as seguintes medidas estatísticas descritivas:

```
> summary(habs[,1])
Mean
Min.    :-1.216508
1st Qu. :-0.481125
Median :-0.089506
Mean    : 0.002828
3rd Qu. : 0.499550
Max.    : 2.008155
```

A menor habilidade média foi estimada em $-1,21$ e maior foi $2,00$. Com relação a maior habilidade média estimada, identificamos que a mesma corresponde a um único aluno que obteve o maior escore calculado pela teoria clássica de testes, indicando, portanto, que o mesmo respondeu aos itens de forma coerente. Segue ainda que, em média, os alunos obtiveram habilidade $0,002$ o que está muito próximo do valor mediano $-0,089$, indicando, portanto, que a distribuição das habilidades estimadas comporta-se de forma razoavelmente simétrica em torno destes valores. Isto pode ser melhor visualizado no histograma (**Figura 6.5**) e no box-plot (**Figura 6.6**) apresentados abaixo.

O comando `hist(habs[,1])` gera o histograma das habilidades médias estimadas.

```
> hist(habs[,1], main = "Habilidades médias estimadas via TRI  
por método Bayesiano MCMC", ylab = "Freq. absoluta",  
xlab = expression( paste("Habilidades médias estimadas  
(", theta, ")") ) )
```

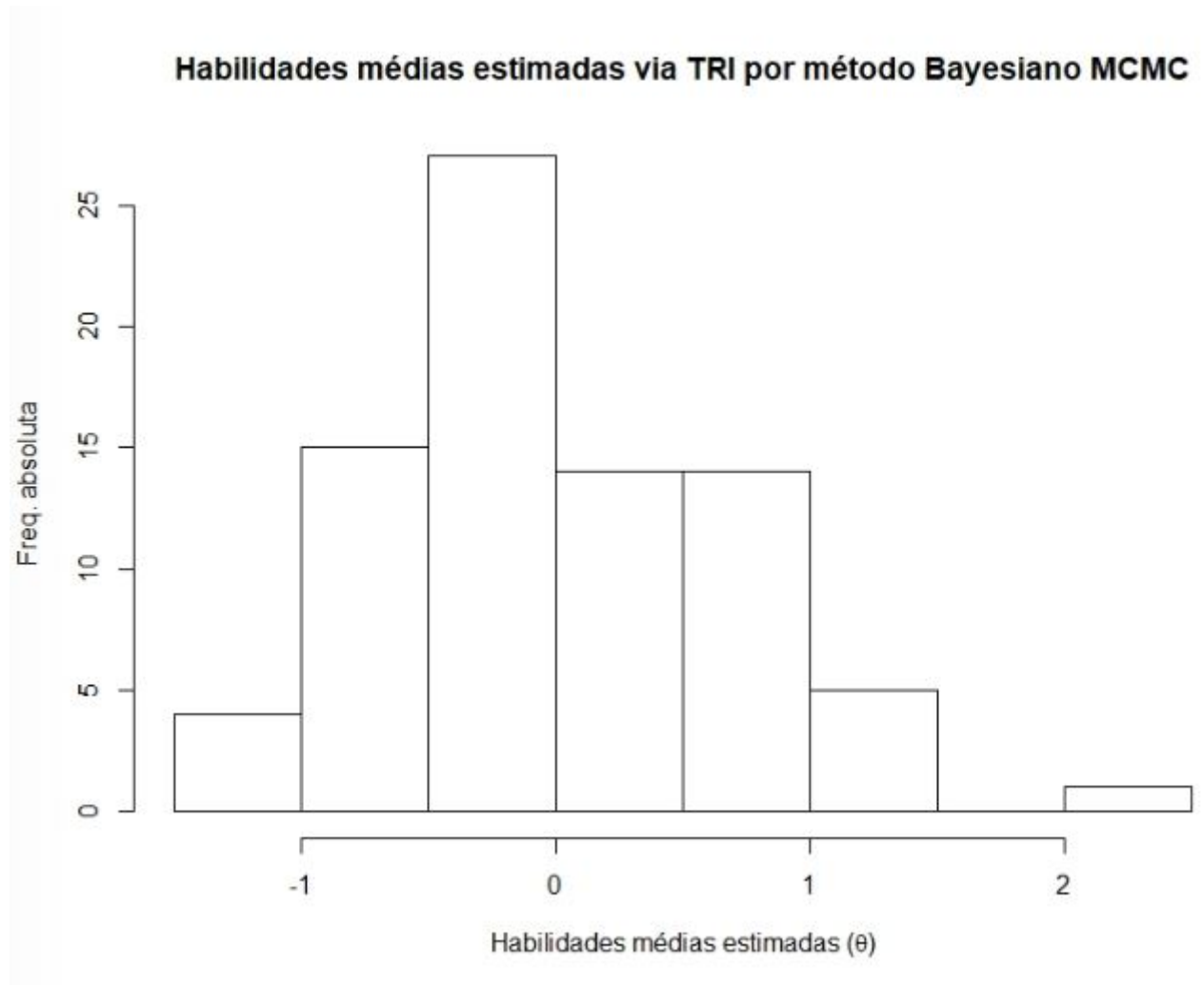


Figura 6.5: Histograma das habilidades médias estimadas via TRI por método Bayesiano

O comando `boxplot(habs[,1])` do gera o box-plot das habilidades médias estimadas.

```
>boxplot(habs[,1])
```

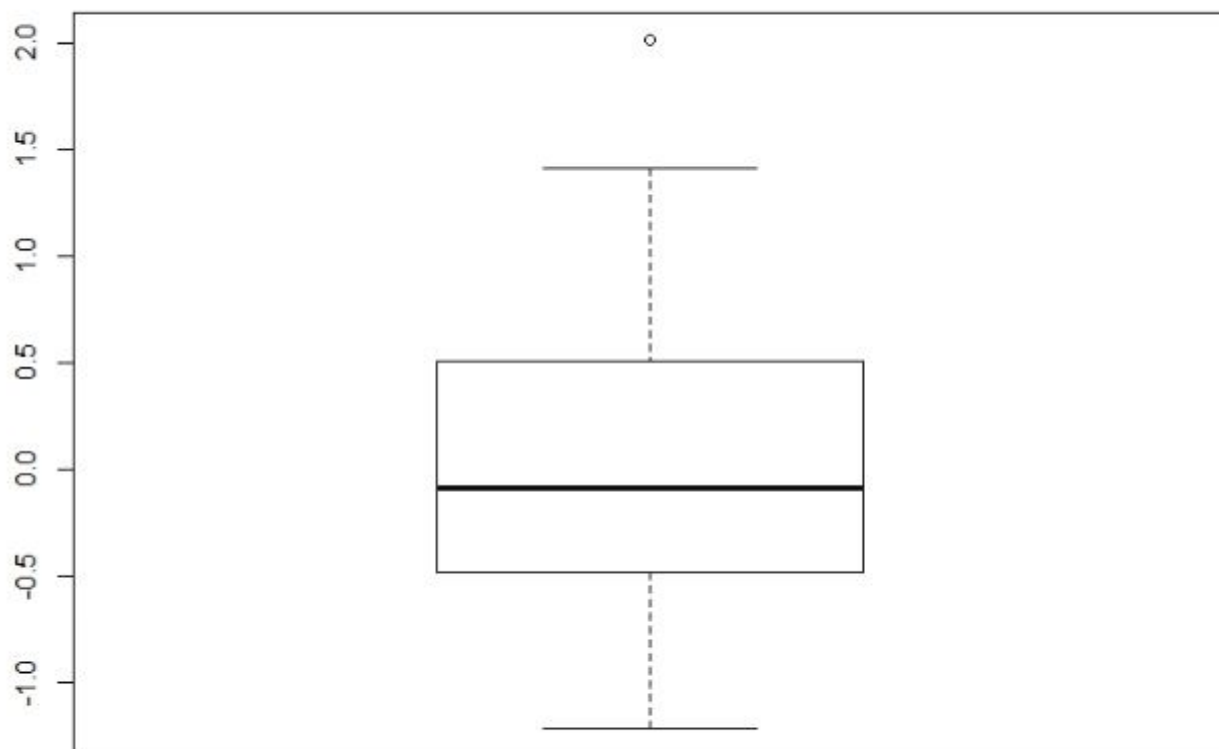



Figura 6.6: Box-plot das habilidades médias estimadas.

O comando *habs* determina as habilidades médias estimadas para cada um dos 80 alunos e também os seus desvios padrões (veja a Tabela 6.3.1. Estimadas no intervalo truncado $[-3, 3]$, pois o modelo utilizado prevê que 99% das habilidades computadas estejam nesse intervalo.

```
> habs
```

6.3.2 Comparando os resultados da TCT e da TRI

Como citado anteriormente, a TCT utiliza o escore total, que consiste na soma dos escores brutos para gerar uma nota em relação ao desempenho do respondente. Já na TRI, a nota depende da habilidade estimada para cada respondente. A fim de comparar os dados

Habilidades médias estimadas para cada respondente

	Habilidade	Desvio padrão		Habilidade	Desvio Padrão
θ_1	0,57562566	0,5185335	θ_{41}	0,37634696	0,5167296
θ_2	1,03503473	0,5144876	θ_{42}	0,21060865	0,5509677
θ_3	0,62018112	0,5227384	θ_{43}	0,65137220	0,5197342
θ_4	-0,04544166	0,5653331	θ_{44}	0,43981785	0,5385965
θ_5	1,01982407	0,5193563	θ_{45}	0,26795380	0,5268522
θ_6	-0,58882547	0,5811621	θ_{46}	0,80169322	0,5155544
θ_7	-0,03564736	0,5860941	θ_{47}	-0,06073779	0,5794531
θ_8	0,91741082	0,5194134	θ_{48}	2,00815531	0,5833923
θ_9	-0,22680445	0,5676407	θ_{49}	-0,77376389	0,6140527
θ_{10}	-0,43021591	0,5992308	θ_{50}	-1,08024409	0,6611433
θ_{11}	0,88948952	0,5418372	θ_{51}	1,40669952	0,5434281
θ_{12}	0,18436333	0,5263271	θ_{52}	0,78758876	0,5036802
θ_{13}	0,27900274	0,5016976	θ_{53}	0,67494342	0,5430667
θ_{14}	-0,26172678	0,5807644	θ_{54}	-0,32654435	0,5646892
θ_{15}	-0,41561776	0,5768841	θ_{55}	-0,63581952	0,6077006
θ_{16}	-0,19744743	0,5853387	θ_{56}	-1,04471325	0,6509982
θ_{17}	-1,21650753	0,6519101	θ_{57}	-0,12628692	0,5530869
θ_{18}	0,60865449	0,5236266	θ_{58}	-0,91271641	0,6500999
θ_{19}	0,48931035	0,5138393	θ_{59}	0,23772955	0,5282028
θ_{20}	-0,74009559	0,6128384	θ_{60}	-0,07910407	0,5572355
θ_{21}	0,77412379	0,4979895	θ_{61}	0,95424141	0,5514059
θ_{22}	-0,33965937	0,5883571	θ_{62}	-0,09990719	0,5543853
θ_{23}	-0,64273657	0,6338064	θ_{63}	-0,59120349	0,6073797
θ_{24}	1,02162676	0,5186148	θ_{64}	0,24041113	0,5462408
θ_{25}	0,53026865	0,4996463	θ_{65}	-0,58940281	0,5993432
θ_{26}	-0,78257115	0,6251809	θ_{66}	-1,06615126	0,6721907
θ_{27}	0,28294121	0,5263547	θ_{67}	-0,87829969	0,6272573
θ_{28}	-0,77154504	0,6245059	θ_{68}	-0,53733066	0,6178675
θ_{29}	-0,12650700	0,5587452	θ_{69}	-0,39527311	0,5740305
θ_{30}	-0,31679752	0,5668416	θ_{70}	-0,82900125	0,6465982
θ_{31}	-0,49093739	0,5940249	θ_{71}	-0,43454466	0,6114297
θ_{32}	0,62266705	0,5366112	θ_{72}	0,08159835	0,5354131
θ_{33}	-0,06804452	0,5536369	θ_{73}	-0,47785392	0,5893747
θ_{34}	-0,51627531	0,6067996	θ_{74}	1,18256200	0,5818576
θ_{35}	-0,46879915	0,6019411	θ_{75}	0,38805841	0,5170822
θ_{36}	-0,11902169	0,6861594	θ_{76}	-0,41903146	0,5915383
θ_{37}	-0,18351246	0,5513595	θ_{77}	-0,38777666	0,5819587
θ_{38}	0,07746531	0,5335999	θ_{78}	-0,01173859	0,5500615
θ_{39}	-0,54938070	0,6097967	θ_{79}	-0,27710965	0,5578791
θ_{40}	0,89086049	0,5016160	θ_{80}	0,26626574	0,5340290

obtidos na análise clássica e na TRI, os dados foram organizados nas tabelas apresentadas nas **Figuras 6.7; 6.8 e 6.9**).

Resultado de cada respondente do teste na TCT e na TRI				
Nº	Escore total	Habilidade	TCT	TRI
1	8	0,57562566	4,0	6,0
2	10	1,03503473	5,0	6,7
3	11	0,62018112	5,5	6,0
4	7	-0,0454417	3,5	4,9
5	10	1,01982407	5,0	6,7
6	5	-0,5888255	2,5	4,0
7	7	-0,0356474	3,5	4,9
8	9	0,91741082	4,5	6,5
9	4	-0,2268045	2,0	4,6
10	6	-0,4302159	3,0	4,3
11	10	0,88948952	5,0	6,5
12	7	0,18436333	3,5	5,3
13	7	0,27900274	3,5	5,5
14	6	-0,2617268	3,0	4,6
15	5	-0,4156178	2,5	4,3
16	5	-0,1974474	2,5	4,7
17	3	-1,21650753	1,5	3,0
18	11	0,60865449	5,5	6,0
19	8	0,48931035	4,0	5,8
20	4	-0,7400956	2,0	3,8
21	10	0,77412379	5,0	6,3
22	6	-0,3396594	3,0	4,4
23	4	-0,6427366	2,0	3,9
24	10	1,02162676	5,0	6,7
25	9	1,02162676	4,5	6,7
26	5	-0,7825712	2,5	3,7
27	9	0,28294121	4,5	5,5

Figura 6.7: Tabela com notas para cada respondente - autor

Pode-se observar que na coluna 2, apresenta os escores totais para cada respondente listado na coluna 1, esses escores encontram-se na extremidade direita da tabela representada nas **figuras 6.1 e 6.2**. A partir desses escores, na coluna 4 determina-se a nota para cada respondente, estabelecendo uma relação linear tal que um x de um total de 0 a 20 acertos, determina um valor y na escala $[0, 10]$, tal que $y = \frac{x}{2}$.

Resultado de cada respondente do teste na TCT e na TRI					
Nº	Escore total	Habilidade	TCT	TRI	
28	4	-0,771545	2,0	3,7	
29	5	-0,126507	2,5	4,8	
30	4	-0,3167975	2,0	4,5	
31	6	-0,4909374	3,0	4,2	
32	11	0,62266705	5,5	6,0	
33	6	-0,0680445	3,0	4,9	
34	4	-0,5162753	2,0	4,1	
35	5	-0,4687992	2,5	4,2	
36	3	-0,1190217	1,5	4,8	
37	5	-0,1835125	2,5	4,7	
38	7	0,07746531	3,5	5,1	
39	4	-0,5493807	2,0	4,1	
40	10	0,89086049	5,0	6,5	
41	7	0,37634696	3,5	5,6	
42	7	0,21060865	3,5	5,4	
43	8	0,6513722	4,0	6,1	
44	6	0,43981785	3,0	5,7	
45	7	0,2679538	3,5	5,4	
46	11	0,80169322	5,5	6,3	
47	6	-0,0607378	3,0	4,9	
48	15	2,00815531	7,5	8,3	
49	4	-0,7737639	2,0	3,7	
50	3	-1,08024409	1,5	3,2	
51	12	1,40669952	6,0	7,3	
52	9	0,78758876	4,5	6,3	
53	9	0,67494342	4,5	6,1	
54	5	-0,3265444	2,5	4,5	

Figura 6.8: Tabela com notas para cada respondente - autor

Na coluna 3, lista-se as habilidades médias para cada respondente, obtidas na análise TRI (última matriz de resultados) e estabelece uma função linear, afim de converter um valor x no intervalo truncado $[-3, 3]$, para um valor y na escala de $[0, 10]$, tal que $y = 5 \frac{(x+3)}{3}$ relacionado na coluna 5.

Resultado de cada respondente do teste na TCT e na TRI					
Nº	Escore total	Habilidade	TCT	TRI	
55	4	-0,3265444	2,0	4,5	
56	3	-1,04471325	1,5	3,3	
57	5	-0,1262869	2,5	4,8	
58	4	-0,9127164	2,0	3,5	
59	7	0,23772955	3,5	5,4	
60	6	-0,0791041	3,0	4,9	
61	9	0,95424141	4,5	6,6	
62	4	-0,0999072	2,0	4,8	
63	5	-0,5912035	2,5	4,0	
64	8	0,24041113	4,0	5,4	
65	3	-0,5894028	1,5	4,0	
66	2	-1,06615126	1,0	3,2	
67	3	-0,8782997	1,5	3,5	
68	3	-0,5373307	1,5	4,1	
69	5	-0,3952731	2,5	4,3	
70	5	-0,8290013	2,5	3,6	
71	5	-0,4345447	2,5	4,3	
72	5	0,08159835	2,5	5,1	
73	3	-0,4778539	1,5	4,2	
74	12	1,182562	6,0	7,0	
75	9	0,38805841	4,5	5,6	
76	4	-0,4190315	2,0	4,3	
77	5	-0,3877767	2,5	4,4	
78	7	-0,0117386	3,5	5,0	
79	5	-0,2771097	2,5	4,5	
80	7	0,26626574	3,5	5,4	

Figura 6.9: Tabela com notas para cada respondente - autor

É notório a variação dos resultados para o mesmo respondente nas duas teorias, assim como essa variação não é constante, ou seja, para cada respondente existe uma diferença entre as notas que é própria do respondente. Isso ocorre, por exemplo com os respondentes 15 e 16, ambos obtiveram nota 2,5 na TCT, porém na TRI as notas foram 4,3 e 4,7, respectivamente. Outra observação importante é que os valores são sempre superiores na TRI do que na TCT, o que reforça a importância da TRI, quanto a observação das habilidades e proficiências dos respondentes (Veja os Gráficos de dispersão apresentados nas **Figuras 6.10 e 6.11**). No gráfico de dispersão apresentado na **Figura 6.12**, é possível observar que a

correlação linear entre os resultados obtidos nas duas teorias é linear crescente.

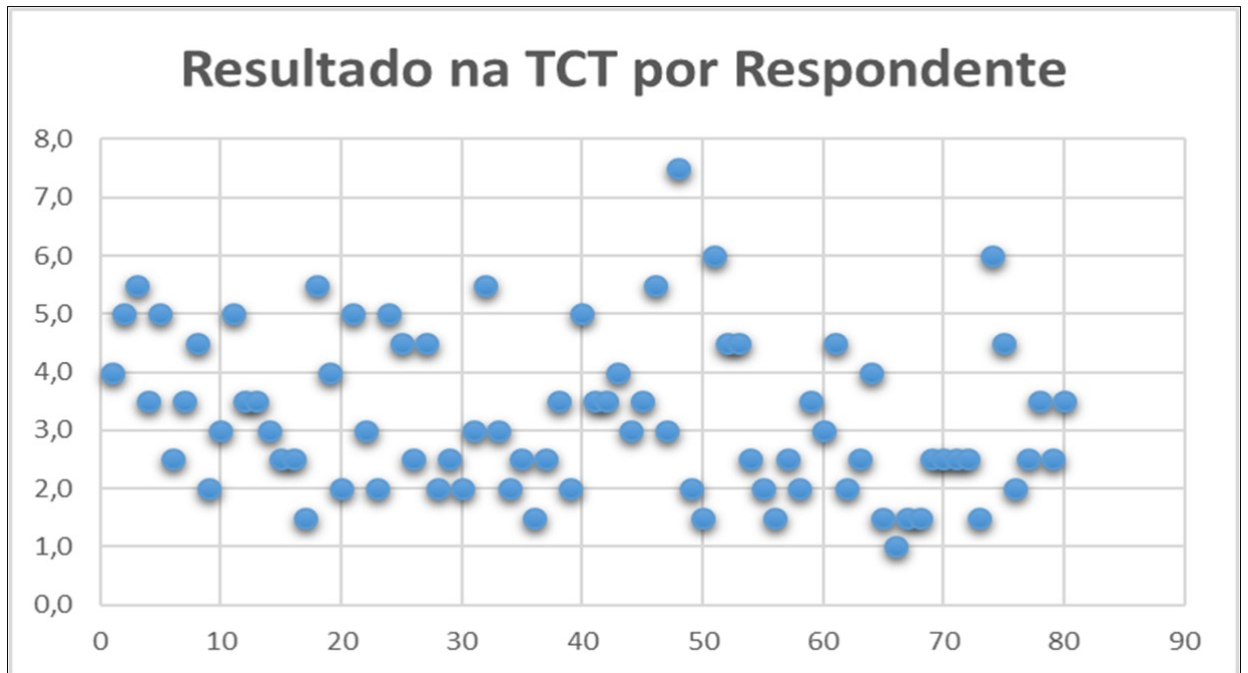


Figura 6.10: Gráfico das notas obtidas por cada respondente na TCT - autor

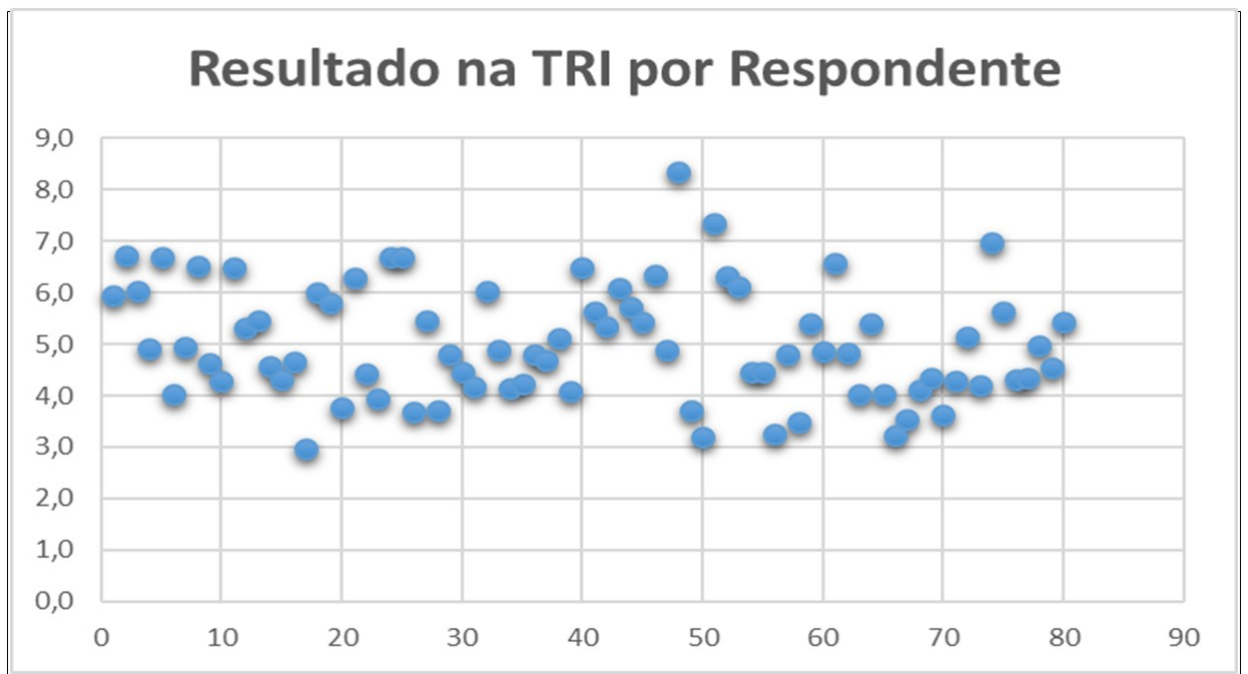


Figura 6.11: Gráfico das notas obtidas por cada respondente na TRI - autor

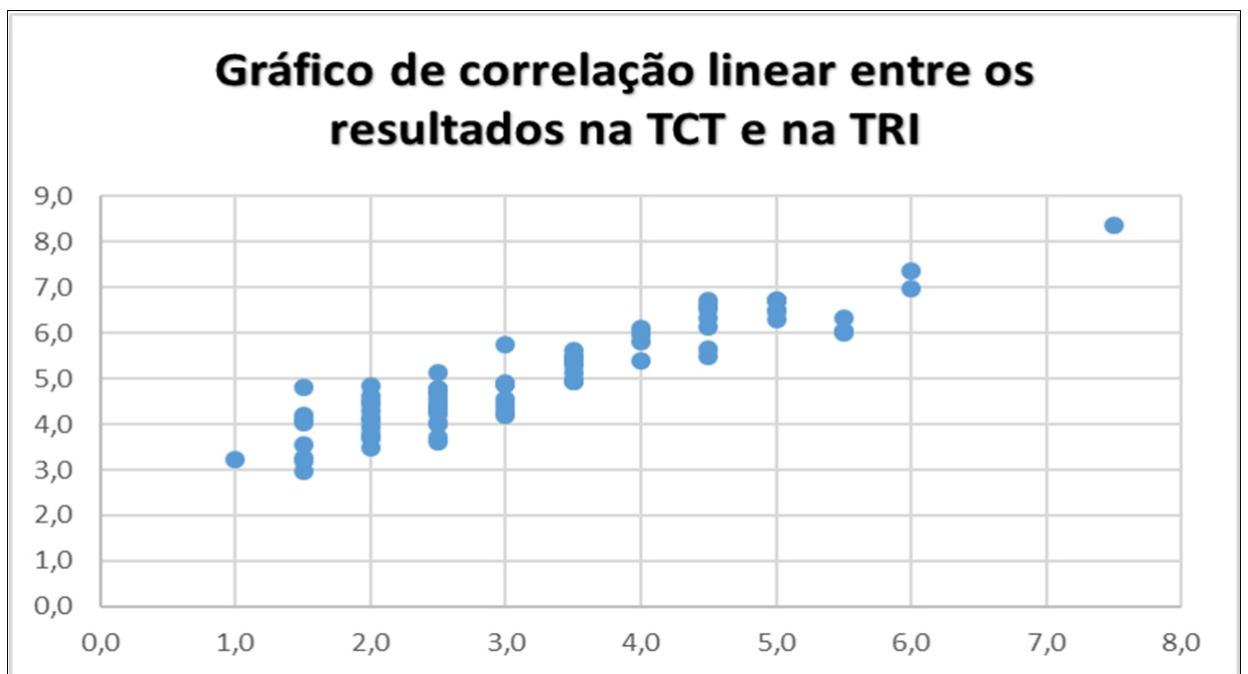


Figura 6.12: Gráfico da Correlação linear entre os resultados - autor

Capítulo 7

Conclusões

Ao promover esse estudo sobre a TRI, observa-se o quanto a educação se mantém inerte em meio a tanta evolução, prevalecendo uma avaliação tradicional, decorativa e punitiva.

Na busca de práticas avaliativas, realizamos a pesquisa bibliográfica destacando os pontos principais da TCT e da TRI, e nota-se que uma não substitui a outra, ou seja, a TRI não veio para substituir a TCT, ambas são importantes processos avaliativos.

Quando olhamos para o ENEM, observa-se itens bem elaborados, que estimulam os alunos a pensar, sugerindo uma situação problema intrigante a fim de que os respondentes desenvolvam um raciocínio construtivo. Pretende-se auxiliar na prática dos professores de ensino médio, a fim que desenvolvam a prática simulada, utilizando bons itens de um banco próprio de dados, e dessa forma estimulem o aluno a criar estratégias de resolução e adequação aos mecanismos da prova, como o tempo de resposta dos itens.

Destaca-se a TRI como importante ferramenta para os professores, pois possibilita diagnosticar e estimar os desempenhos individuais e coletivos dos alunos, favorecendo a compreensão das dificuldades dos itens e permitindo estabelecer relações de comparação entre as habilidades trabalhadas em relação ao coletivo, como individualmente. Ao se aplicar simulados periódicos, o professor poderá utilizar a TRI para identificar as dificuldades do grupo ou individualmente, o que permitirá planejar melhor suas ações, promovendo um ensino de qualidade e uma aprendizagem mais significativa.

Com a realização deste trabalho constatou-se que a utilização da TRI é uma proposta interessante para o Ensino Médio, porém é um tanto desafiadora. A ideia inicial deste estudo consistia em tornar a TRI mais acessível aos professores do ensino médio possibilitando a utilização desta teoria em sua prática pedagógica, mas durante a análise dos resultados ficou claro a complexidade da sua aplicação, necessitando de um profissional com conhecimento

no software *R*, capacitado na área da estatística para inferência dos dados.

Através deste estudo, pode-se estimular o desenvolvimento de vários trabalhos futuros, como a construção de um sistema computacional capaz de calcular o TRI, facilitando a sua utilização pelos professores; utilização da TRI como política pública, para melhoria da Educação Básica; Análise pedagógica dos itens interpretando os resultados obtidos nas teorias; entre outros.

Referências Bibliográficas

- [1] ANDRADE, D.F.; DOS ANJOS; D. *Teoria da Resposta ao Item com uso do R*, João Pessoa-PB (2012).
- [2] ANDRADE, D.F.; TAVARES, H. R.; VALLE, R. C; *Teoria de Resposta ao Item: conceito e aplicações*, SINAPE (2000).
- [3] INEP; *Guia de Elaboração e Revisão de Itens*, Volume I, Brasília-DF (2000).
- [4] LUCKESI, C.C.; *Avaliação da aprendizagem escolar*. 14 ed. São Paulo: Cortez (2002).
- [5] PASQUALI, L; *Teoria e métodos de medida em ciências do comportamento* Brasília-DF, Laboratório de Pesquisa em Avaliação e Medida/Instituto de Psicologia/UnB: INEP (1996).
- [6] RABELO, E.H.; *Avaliação: novos tempos, novas práticas*. Petrópolis: Vozes (1998).
- [7] STONE M.H.; WRIGHT B.D.; *Best Test Design*. Chicago (1979).

Apêndice A

Primeiro Apêndice

Neste Apêndice consta as respostas dos itens elaborados para o simulado aplicado nas turmas de 3º ano do ensino Médio.

Item 1

Após a análise da tabela, calculamos aproximadamente as razões entre o número de casos confirmados e casos descartados:

$$\text{Bahia: } \frac{170}{120} = 1,4$$

$$\text{Maranhão: } \frac{53}{31} = 1,7$$

$$\text{Piauí: } \frac{62}{31} = 2$$

$$\text{Rio Grande do Norte: } \frac{81}{35} = 2,3$$

$$\text{Sergipe: } \frac{26}{14} = 1,9$$

O gabarito é a alternativa **d**.

Item 2

As dimensões do lava-pés preenchido por água é: $2m, 1,75m$ e $20cm$ ($0,2m$). Portanto o volume de água é dado por:

$$v = 2 \cdot 1,75 \cdot 0,2$$

$$v = 0,7 \sqrt[3]{3}.$$

Como $1m^3$ equivale a $1000L$, então, $0,7m^3$ equivale a $700L$. Se o cloro residual deve ser conservado em $25mg/L$, temos que: $25 \cdot 700 = 17500mg$. Como cada $1mg$ equivale a $0,001g$, então, $17500mg$ equivale a $17,5g$.

O gabarito é a alternativa **b**.

Item 3

A partir da análise do gráfico de colunas, admitimos que o valor percentual representando a Paraíba está em torno do valor $5,5$. Daí, temos:

$$5,5\% \text{ de } 399 = 21,9.$$

O gabarito é a alternativa **c**.

Item 4

Cada pacote com 5 figurinhas custa R\$ 6,00 na Suíça. Portanto o valor a ser pago por cada figura é de R\$ 1,20. Como a figura está sendo vendida por um colecionador a um valor equivalente a 750% de seu valor original, então:

$$750\% \text{ de } 1,20 = 9,00$$

O gabarito é a alternativa **d**.

Item 5

A partir dos dados apresentados na tabela, calculamos a média de pessoas entrevistadas em cada sábado:

$$\frac{(198+75+104)+(207+51+79)+(153+63+132)+(136+46+76)}{4} =$$

$$\frac{377+337+348+258}{4} =$$

$$\frac{1320}{4} = 330$$

O gabarito é a alternativa **a**.

Item 6

No jogo de dominó existe 7 peças carroças. Logo a probabilidade de escolher ao acaso uma pedra, ser uma carroça, é de:

$$\frac{7}{28} = \frac{1}{4}$$

O gabarito é a alternativa **e**

Item 7

Para chegar ao tesouro, a equipe A precisa dar um total de 25 passos. Como cada passo equivale a 82cm, então:

$$25 \cdot 82cm = 6170cm.$$

Como cada 1cm equivale a 0,01m, temos 61,7m.

O gabarito é a alternativa **b**.

Item 8

Para determinarmos a mediana de todas as pontuações, precisamos ordenar os números:

280 302 316 320 360 405 420 445

Daí, calculamos a média aritmética dos dois termos centrais:

$$\frac{320+360}{2} = 340$$

Logo, as equipes 1C, 2A, 1B e 3B, receberam um ponto, e as equipes 2C, 3A, 2B e 1A, receberam 2 pontos.

O gabarito é a alternativa **c**.

Item 9

Um ponto pertence a uma região circular se as suas coordenadas ao serem substituídas na equação da circunferência resultarem em um valor menor que zero. Analisando o plano cartesiano identifica-se as coordenadas dos pontos são: A (4, -2), B (1, -1), C (6, 5), D (2, 4) e E (7, 1). Aplicando esses valores na expressão $x^2 + Y^2 - 8x - 4y + 11$, temos:

$$A \rightarrow 4^2 + (-2)^2 - 8.4 - 4.(-2) + 11 = 39 - 32 = 7, \text{ temos } 7 > 0.$$

$$B \rightarrow 1^2 + (-1)^2 - 8.1 - 4.(-1) + 11 = 17 - 8 = 9, \text{ temos } 9 > 0.$$

$$C \rightarrow 6^2 + 5^2 - 8.6 - 4.5 + 11 = 72 - 68 = 4, \text{ temos } 4 > 0.$$

$$D \rightarrow 2^2 + 4^2 - 8.2 - 4.4 + 11 = 31 - 32 = -1, \text{ temos } -1 < 0.$$

$$E \rightarrow 7^2 + 1^2 - 8.7 - 4.1 + 11 = 61 - 60 = 1, \text{ temos } 1 > 0.$$

O gabarito é a alternativa **d**.

Item 10

A forma escolhida por Sofia para poupar dinheiro representa uma progressão aritmética, cujo primeiro termo é igual a 1 e sua razão também é 1. Como ela começou a poupar no dia 1º de junho e parou em 25 de agosto, então ela poupou durante 85 dias, colocando no último dia o valor de R\$ 85,00. Aplicando a soma de termos de uma Progressão aritmética, calculamos o valor poupado durante esse período:

$$S_n = \frac{(a_1 + a_n)n}{2}$$

$$S_{85} = \frac{(1 + 85)85}{2}$$

$$S_{85} = \frac{(86.85)}{2} = 3655$$

Como Sofia já tinha R\$ 1 000,00 a ser acrescido no valor poupado, então o valor total é de R\$ 4 655,00.

O gabarito é a alternativa **d**.

Item 11

Analisando a figura, percebemos que cada degrau representa um triângulo retângulo em relação a rampa. Aplicando o teorema de Pitágoras calcula-se a medida da rampa equivalente a um degrau.

$$x^2 = 15^2 + 30^2$$

$$x^2 = 225 + 900$$

$$x^2 = 1125 \quad x = 15\sqrt{5}, \text{ como } \sqrt{5} = 2,2, \text{ então}$$

$$x = 15.2,2 = 33cm$$

O comprimento da rampa equivale a distância de 20 degraus, portanto $20.33 = 660cm$. Como cada 1cm equivale a 0,01m, então temos 6,60m de comprimento. Calculando a superfície a ser revestida, temos:

$$S = 6,60.1,20 = 7,92m.$$

O gabarito é a alternativa **a**.

Item 12

A distância equivalente a $6,5\text{cm}$ na escala $1 : 3\,300\,000$ é de:

$$6,5 \cdot 330000 = 21450000\text{cm}.$$

Como cada cm equivale a $0,00001\text{km}$, então a distância a ser percorrida é de $214,5\text{Km}$.

O gabarito é a alternativa **b**.

Item 13

Cada frasco possui 5 doses de vacina padrão, portanto para vacinar 4,7 milhões de pessoas, com a dose padrão, são necessários 940 000 frascos. O mesmo frasco pode imunizar até 25 pessoas com a dose fracionada, portanto para vacinar 15 milhões de pessoas, com a dose fracionada, são necessários 600 000 frascos. No mínimo devem ser usados 1 540 000 frascos.

O gabarito é a alternativa **c**.

Item 14

Se a seção da base da pirâmide foi realizada a $\frac{3}{8}$ da altura da pirâmide, podemos através do paralelismo, estabelecer uma relação de proporcionalidade, tal que:

$$\frac{3}{8} = \frac{x}{36}$$

$$8x = 108$$

$$x = 13,5\text{cm}.$$

O gabarito é a alternativa **b**.

Item 15

Pelas instruções de uso, temos que 3mL é o indicado para 24L de água, portanto, cada 1mL equivale a 8L de água. Assim como, 2 gotas é indicado para cada 1L de água. Logo estabelecemos a relação:

$$\text{Se } 1\text{mL} \text{ equivale a } 8\text{L}, \text{ então } 1\text{mL} = 8 \cdot 2 = 16 \text{ gotas}.$$

O gabarito é a alternativa **a**.

Item 16

Analisando os dados fornecidos na situação problema, observamos que:

A alternativa **a** está errada, pois estado com menor taxa de feminicídio é Roraima, com nenhum caso.

A alternativa **b** está errada, pois o número de feminicídio no Brasil aumentou em 134 casos, de 2016 a 2017.

A alternativa **c** está errada, pois o aumento de $6,5\%$ no número de homicídios dolosos em 2017, foi em relação a 2016.

A alternativa **d** está correta. Pois:

$$\frac{946}{4473} = 0,21, \text{ aproximadamente, o que equivale a } 21\%.$$

A alternativa **e** está errada, pois o número de feminicídio no estado do Mato Grosso é de 4,6 casos a cada 100 mil mulheres.

O gabarito é a alternativa **d**.

Item 17

Como a Serra da Boa Vista tem altitude de $1195m$, então pelo quadro sua pressão atmosférica é de $656mmHg$. Como o ar atmosférico tem em sua composição 21% de oxigênio, então:

21% de $656 = 137,76$ corresponde a pressão parcial de oxigênio nessa altitude. Logo, a variação da pressão parcial de oxigênio é:

$$159,6 - 137,76 = 21,84mmHg$$

O gabarito é a alternativa **c**.

Item 18

A inclinação da reta representada no gráfico pode ser determinada pela tangente do ângulo φ , representada por:

$$\tan\varphi = \frac{L_1 - L_0}{\theta_1 - \theta_0}.$$

Como $L_1 - L_0 = \delta L$ e $\theta_1 - \theta_0 = \delta\theta$, então podemos escrever:

$$\tan\varphi = \frac{\delta L}{\delta\theta}. \text{ Como } \Delta L = L_0 \cdot \alpha \cdot \Delta\theta, \text{ substituindo essa expressão na relação anterior, temos:}$$

$$\tan\varphi = \frac{L_0 \cdot \alpha \cdot \Delta\theta}{\delta\theta}. \text{ Logo,}$$

$$\tan\varphi = L_0 \cdot \alpha.$$

O gabarito é a alternativa **a**.

Item 19

O lucro (L) é a relação entre a quantidade (x) de sonhos vendidos a R\$ 3,00 cada um, subtraído do valor gasto na produção do doce, neste caso R\$ 50,00. Esta relação é representada pela expressão $L = 3x - 50$.

O gabarito é a alternativa **e**.

Item 20

As intersecções entre os segmentos citados na situação-problema define oito regiões triangulares congruentes, formando um octaedro.

O gabarito é a alternativa **c**.