



**Serviço Público Federal
Universidade Federal do Pará
Instituto de Ciências Exatas e Naturais
Programa de Pós-Graduação em Matemática e Estatística**

Métodos Estatísticos para Detecção de Fraudes em Testes

José Ivan da Luz Rodrigues

Belém - Pará
2016



**Serviço Público Federal
Universidade Federal do Pará
Instituto de Ciências Exatas e Naturais
Programa de Pós-Graduação em Matemática e Estatística**

Métodos Estatísticos para Detecção de Fraudes em Testes

José Ivan da Luz Rodrigues

Dissertação apresentada ao Programa de Pós-Graduação em Matemática e Estatística (PPGME) do Instituto de Ciências Exatas e Naturais (ICEN) da UNIVERSIDADE FEDERAL DO PARÁ (UFPA), como requisito para a obtenção do título de Mestre em Estatística, orientado pelo Prof. Dr. Héilton Ribeiro Tavares.

Belém - Pará
2016

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema de Bibliotecas da UFPA

Rodrigues, José Ivan da Luz , 1979-
Métodos estatísticos para detecção de fraudes em
testes / José Ivan da Luz Rodrigues. - 2016.

Orientador: Héilton Ribeiro Tavares.
Dissertação (Mestrado) - Universidade
Federal do Pará, Instituto de Ciências Exatas e
Naturais, Programa de Pós-Graduação em
Matemática e Estatística, Belém, 2016.

1. Estatística matemática. 2. Métodos
Estatísticos. 3. Testes e medidas
Educativas-Metodologia. 4. Fraude na
Educação-Testes-Detecção-Avaliação. 5.
CopyDetect. I. Título.

CDD 22. ed. 519.5



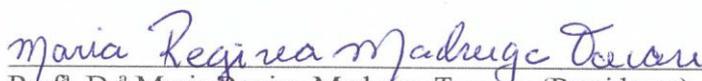
Serviço Público Federal
Universidade Federal do Pará
Instituto de Ciências Exatas e Naturais
Programa de Pós-Graduação em Matemática e Estatística

Métodos Estatísticos para Detecção de Fraudes em Testes

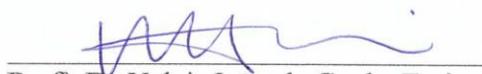
José Ivan da Luz Rodrigues

Orientador: Prof. Dr. Héilton Ribeiro Tavares – PPGME/ICEN/UFPA

Banca Examinadora:


Prof.^a Dr.^a Maria Regina Madruga Tavares (Presidente)


Prof.^a Dr.^a Marinalva Cardoso Maciel


Prof.^a Dr. Valcir João da Cunha Farias

Belém - Pará
2016

“A mente que se abre a uma nova idéia jamais voltará ao seu tamanho original”

Albert Einstein.

Agradecimentos

A presente dissertação efetiva mais um passo rumo a minha realização intelectual e profissional, sendo um caminho trilhado desde a infância marcada por inúmeras dificuldades em todos os aspectos, que foram gradativamente superadas, pela bondade de Deus ao me abençoar com uma família cheia de princípios que norteiam os mandamentos do evangelho (Bíblia), mostrando-me que poderia conquistar novos horizontes, de maneira honesta e digna. Agradeço o apoio incessante de meus pais: Iracema Rodrigues e José Benedito Rodrigues, meus irmãos: Lorena e Ivanilson, a minha esposa (Eliane) e o meu filho (Matheus), minha tia Rosa Rodrigues que me auxiliaram e me incentivaram a enfrentar as dificuldades para conquistar mais essa vitória, e alguns outros familiares que direta e indiretamente ajudaram-me a trilhar esta glória.

Agradeço aos meus grandes amigos do Comando de Policiamento da Região Metropolitana da Polícia Militar do Pará, Sgt PM Adriana, Sgt PM Miriam, Sgt PM Lustosa, aos Cabos PM: Valéria e Vanderlei e, aos Soldados PM: Ketlin, Joyce, Dutra, Breno e Talita, pelo o apoio em momentos difíceis. Agradeço em especial ao Ten's Cel's PM's: Leão e Salim e aos Majores PM's: Lima e Albuquerque, pela grandiosa consideração e oportunidade, os quais não mediram esforços dentro de suas competências em me apoiar e intervir no que fosse necessário para que eu pudesse realizar este sonho.

Agradeço à Universidade Federal do Pará - UFPA, à Faculdade de Estatística, ao PPGME (Programa de Pós-Graduação em Matemática e Estatística), aos professores: Terezinha Ferreira de Oliveira e Maria Regina Madruga Tavares, pelo incentivo, críticas e elogios no correr de nosso curso.

Agradeço a meu orientador Prof. Dr. Héilton Ribeiro Tavares por suas intervenções, críticas, elogios, e sugestões teórico-metodológicas, sendo de total importância no desenvolvimento desta pesquisa. Ressalto minha estima e admiração por este docente impar na área de conhecimento (Estatística), o qual me permitiu desfrutar de seu imenso intelecto para dar seguimento e conclusão a este trabalho.

MÉTODOS ESTATÍSTICOS PARA A DETECÇÃO DE FRAUDES EM TESTES.

José Ivan da Luz Rodrigues

RESUMO: A presente dissertação foi desenvolvida com o propósito de subsidiar estudos para a identificação de possíveis esquemas (fraudes) praticados em aplicações de testes avaliativos, considerando a similaridade de respostas entre examinandos do tipo fonte (sujeito suspeito de ter suas respostas copiadas (s)) e o copiador (suspeito de copiar as respostas da fonte (c)). Dentre mais de vinte índices propostos na literatura, procurou-se explicar alguns desses métodos, que foram desenvolvidos com base na Teoria Clássica dos Testes (TCT) ou na Teoria da Resposta ao Item (TRI). Usou-se para exemplificação o algoritmo desenvolvido por Zopluoglu (2013) utilizando o pacote estatístico CopyDetect para a geração de dados simulados, para esta implementação abordou-se alguns índices desenvolvidos com base na Teoria da Resposta ao Item ($\hat{\Omega}$, Teste Binomial Generalizado, K , K_1 e K_2 , S_1 e S_2). Assim, para efeito de estudo em um primeiro momento fixou-se dois indivíduos que chamamos de fonte (s) e copiador (c), para verificarmos a ocorrência de fraude, logo neste momento não se constatou evidências de fraudes. Posteriormente, adotou-se as respostas do indivíduo copiador para o fonte e vice-versa, com o intuito identificar dentre 100 examinando selecionados aleatoriamente 5% destes como possíveis burladores, desta forma conseguiu-se o resultado esperado, pois dentre os índices citados acima dois deles identificaram 5 fraudadores, o índice $\hat{\Omega}$ e Teste Binomial Generalizado, para um nível de significância de 5%.

PALAVRAS-CHAVE: Fraude, Teoria Clássica dos Testes, Teoria da Resposta ao Item, Métodos estatísticos, CopyDetect.

STATISTICAL METHODS FOR THE DETECTION OF FRAUD IN TEST.

José Ivan da Luz Rodrigues

ABSTRACT: This work was developed in order to support studies to identify possible schemes (fraud) practiced in evaluative testing applications, considering the similarity of responses among examinees type source (subject suspected of having your copied answers (s)) and the copier (suspected of copying the answers from the source (c)). Among more than twenty levels proposed in the literature, we tried to explain some of these methods, which were developed based on the Classical Theory of Tests (TCT) or Item Response Theory (IRT). It used to exemplify the algorithm developed by Zopluoglu (2013) using the statistical package CopyDetect to generate simulated data for this implementation we approached if some indices developed based on Item Response Theory (Omega, Test Binomial Generalized, K , K_1 and K_2 , S_1 and S_2). Thus, to study the effect at first stood two individuals who call the source (s) and Copier (c) to verify the occurrence of fraud, then at this point not found evidence of fraud. Later, we adopted the individual answers the copier to the source and vice versa, in order to identify among 100 examining selected randomly 5% of these as possible burladores in this way could be the expected result because from the indexes mentioned above two 5 of them identified fraudsters, Omega index and Test Binomial Generalized for a 5% significance level.

KEYWORDS: Fraud, Classical Theory of Tests, Item Response Theory, Statistical Methods, CopyDetect.

Sumário

CAPÍTULO I – INTRODUÇÃO	12
1.2 Motivação	14
1.3 Objetivos:.....	15
1.3.1 Geral	15
1.3.2 Específico.....	15
1.4 Estruturação do Trabalho:.....	15
CAPÍTULO II – O UNIVERSO DAS FRAUDES	17
2.1 Breve Histórico sobre Fraude.	17
2.2 Definindo Fraude	20
2.3 Intuitos do Golpista	21
2.4 Código Penal Brasileiro (CPB)	21
2.5 Concurso Público no Brasil – Previsão Legal.....	22
2.5.1 Constituição de 1934	22
2.5.2 A Constituição Federal do Brasil de 1988	23
2.5.3 Código Penal Brasileiro pela Lei nº 12. 550 de 15 de Dezembro de 2011 – Fraudes em Certames de Interesse Público	24
2.6 Fraudes em Concursos Públicos no Brasil.....	25
2.6.1 Alguns Relatos de Fraudes em Concursos Públicos no Brasil	26
2.7 Outras Variações de Fraudes	29
2.7.1 Esquema da Nigéria.....	29
2.7.2 Fraudes em Cartão de Crédito	29
2.7.3 Fraude Acadêmica	29
2.7.4 Fraudes em Testes Avaliativos.....	30
2.7.5 Fraudes de Engenharia Social	31
CAPÍTULO III – DA TEORIA CLÁSSICA DOS TESTES (TCT) A TEORIA DA RESPOSTA AO ITEM (TRI).....	32
3.1 Teoria Clássica dos Testes (TCT): Aporte Histórico	32
3.2 Modelos adotados pela Teoria Clássica dos Testes (TCT).....	34
3.2.1 Análise Fatorial	34
3.2.1.1 Hipóteses do Modelo Ortogonal.....	35
3.2.2 Modelo linear clássico	37
3.3 Teoria da Resposta ao Item (TRI).....	38
3.4 Modelos Matemáticos Adotados pela Teoria da Resposta ao Item.....	40
3.4.1 Aporte Conceitual.....	40
3.4.2 Modelos para Avaliação de Itens Dicotômicos	40

3.4.3 Definição da Métrica para o Traço Latente (Θ)	43
3.5. Hipóteses Relevantes para os Modelos de Resposta ao Item.	43
3.5.1 Unidimensionalidade do Teste	43
3.5.2 Independência Local.....	44
3.6 Alguns Modelos de Resposta ao Item para Avaliação não Dicotômica	44
3.6.1 Modelos de Resposta Nominal (MRN).....	44
3.6.2 Modelos de Resposta Gradual (MRG).....	45
3.7. Estimação de parâmetros: Itens e Habilidade	47
CAPÍTULO IV – MÉTODOS ESTATÍSTICOS PARA A DETECÇÃO DE FRAUDES EM TESTE.	
.....	50
4.1 Introdução.....	50
4.2 Modelos para Detecção de Fraudes em Testes	53
4.3 Alguns Modelos Baseados na Teoria Clássica dos Testes (TCT)	54
4.3.1 Índice de Crawford.	54
4.3.2 Índice de Dickenson	55
4.3.3 Índice de Anikeef	55
4.3.4 Índice de Regressão	56
4.3.5 Índices de Angoff	57
4.3.6 Estatística g	58
4.3.7 Índices PAR1 e PAR2.....	59
4.3.8 Índice de Análise de Similaridade de Erro (IASE).....	60
4.4 Alguns Modelos Baseados na Teoria da Resposta ao Item (TRI).....	60
4.4.1 Índice ω	60
4.4.2 Teste Binomial Generalizado	61
4.4.3 Índice K	63
4.4.3.1 Distribuição Empírica para o Índice K	64
4.4.3.2 Aproximação Teórica para Obtenção do Índice K	65
4.4.3.3 Obtenção de Variantes para do Índice K	66
4.4.4 Índice S ($S1$ e $S2$).....	67
4.4.4.1 Índice $S1$	67
4.4.4.1.1 Distribuição de M	69
4.4.4.1.2 Modelo para estimação de μ	69
4.4.4.1.3 Verificação do modelo proposto μ	70
4.4.4.2 Índice $S2$	71
4.4.4.3 Limitações do Índice S	76
CAPÍTULO V – IMPLEMENTAÇÃO COMPUTACIONAL.....	77

5.1 Introdução.....	77
5.1.1 Pacote Estatístico CopyDetect.....	77
5.1.2 Etapas exultadas pelo Algoritmo Implementado no Pacote Estatístico CopyDetect (CopyDetect1) para a Detecção de Fraude	79
5.1.3 Resultados	81
5.1.3.1 Índice ômega	83
5.1.3.2 Testes Binomial Generalizado.....	84
5.1.3.3 Índice K	86
5.1.3.4 Variantes do índice K ($K1, K2, S1$).....	87
5.1.3.5 Índice $S2$	89
5.1.4 Considerações Finais	93
APÊNDICE A	94
Algoritmo para Geração de Dados.....	94
REFERÊNCIAS BIBLIOGRÁFICAS	97

CAPÍTULO I – INTRODUÇÃO

Métodos avaliativos no seio das sociedades humanas são empregados desde o início dos tempos, mesmo de forma rústica, com o propósito de analisar as habilidades humanas em diversas áreas de conhecimento e subsidiar as propostas metodológicas para melhoramento das mesmas. No decorrer dos séculos a disponibilização e o aprimoramento da instituição escola possibilitaram novas oportunidades de inserção social, por meio da educação, possibilitando a ricos e pobres, em tese, um mesmo patamar no que tange a ascensão intelectual.

Testes avaliativos são de suma importância para quantificação educacional, pois através dos resultados obtidos é possível avaliar a qualidade destas instituições. Possibilitando a implementação de novas políticas e metodologias adotadas por parte do poder público e gestores escolares em busca de altos padrões de excelência. Porém, as demandas de acentuada competitividade que a sociedade moderna impõe, acabam por potencializar a busca de indivíduos por meios fraudulentos, como forma de almejar novas conquistas, prejudicando consideravelmente os instrumentos avaliativos, outros indivíduos, e, conseqüentemente a conquista de valorosos padrões de excelência através de suas práticas nocivas.

Nesse sentido, a desonestidade praticada por partes dos fraudadores interfere diretamente na validade e qualidade das inferências a partir de resultados de testes (BLISS, 2012). Segundo Cizek (1999) (apud BLISS, 2012, p.01) existem alguns danos causados por esta prática:

Uma criança pode ser julgada pronta para o jardim de infância quando ela não está; a consequência para o indivíduo pode ser de um ano ou mais de frustração ou sentimento de fracasso. O estudante pode ser admitido a um colégio para o qual não está devidamente preparado; ele também pode estar impedindo a admissão de um estudante mais qualificado.

Em consonância às práticas avaliativas surgiram métodos fraudulentos, onde pessoas com intuito de obter vantagens ilícitas tentam ludibriar as fiscalizações para lograr êxito frente à concorrência. Nesta vertente, nas últimas décadas a elaboração de técnicas para a detecção de fraude surgiu como grande avanço para tentar conter ou reprimir atos fraudulentos, subsidiando aplicação de testes avaliativos proporcionando mais lisura ao mesmo. Segundo Bolton e Hand (2002) (apud KOVACH, 2011, p. 16):

Detecção de fraude consiste em identificar atividades não autorizadas uma vez que a prevenção de fraudes falhou. Na prática a detecção de fraudes deve ser aplicada constantemente, pois o sistema não tem ciência de quando a prevenção falha.

Métodos burlatórios em meios avaliativos são cada vez mais frequentes em todos os âmbitos, os aperfeiçoamentos destas técnicas por pessoas mal-intencionadas desafiam fiscalizadores e pesquisadores a evitar este tipo de prática. A tendência do ser humano em almejar melhorias para si, e, conseqüentemente retorno financeiro, por vezes, faz com que este se submeta a práticas nocivas lesando outras pessoas. A *cola* seja através de papelzinhos, memorização de itens, comunicação eletrônica etc, é o suporte para a questão da fraude na avaliação. Segundo Pimenta e Pimenta (2011, p. 4),

Aos esforços para burlar a avaliação, juntam-se os sites, os blogs, o vídeo, o filme¹ explicando como e porque burlar. Destarte, esta disseminação, bem como as recorrentes queixas de professores nos diversos níveis de ensino, nos leva a crer que além de ser um fenômeno que se mantém em nossa sociedade, este vem admitindo novos contornos e novos sentidos que distanciam cada vez mais a avaliação enquanto prática pedagógica que faz o diagnóstico das aprendizagens dos alunos.

A prática da *cola* se tornou cada vez mais presente dentro das instituições de ensino, sendo extremamente nociva na avaliação dos discentes, pois os praticantes demonstram, em tese, altas proficiências que na realidade não possuem, prejudicando o processo avaliativo. “No dicionário de língua portuguesa colar significa copiar fraudulentamente. Sinônimo de pesca, a cola é a consulta a uma fonte não-autorizada numa avaliação.”² Assim, a cola pode ser vislumbrada dentre duas ópticas: uma como incentivadora para corrupção e outra como forma demonstrar grande intelecto evitando a exposição de alunos de pouca proficiência.

Neste contexto, podemos definir cada participante dentro do esquema fraudulento, ou seja, as relações de trocas de informações entre estudantes. Assim, na prática da *cola* temos o indivíduo caracterizado como ativo sendo este o que se beneficia diretamente da informação para a obtenção de vantagem e o agente passivo o que fornece a cola, sendo ambos, peças essenciais para a caracterização do ato ilícito (BUNN et al, 1992).

¹ O filme Quem Não Cola Não Sai da Escola – Cheaters, de Andrew Gurland (EUA, comédia, 2002), tem como tema central a cola. No YouTube, há vídeos que ensinam técnicas para colar (em How to cheat on any test); Chat explicando como colar em <http://br.answers.yahoo.com/question/index?qid=20060818190643AAenHDI>; e ainda em www.perguntascertinas.com.br/2008/04/22/como-colar-na-prova-manual-da-cola/. **Acessos em:** 27-04-2016.

² Cola na Escola: Breves reflexões sobre a problemática. **Disponível em:** <http://www.ufrgs.br/psicoeduc/wiki/Cola_na_Escola:_Breves_reflex%C3%B5es_sobre_a_problema%C3%A1tica>. **Acesso em:**03-05-2016.

Priorizando a lisura e eficácia dos testes, de modo a avaliar o examinando de forma mais contundente e não apenas na preocupação se obtiveram bom desempenho durante avaliações, surgiram como proposta a subsidiar esta expectativa, os métodos para a detecção de fraudes em teste, os quais possibilitam ao investigador sinalizar a ocorrência de prática ilícita na sua realização, através dos padrões de respostas dos indivíduos.

1.2 Motivação

As adversidades econômicas vividas em nossa nação no decorrer dos anos, principalmente nesta atual gestão brasileira, com diversas notícias de corrupções e fraudes, em especial, o caso da Petrobras através da “Operação Lava-Jato”, são um dos principais fatores de desestruturação ou falência de entidades públicas e conseqüentemente ocasionam enorme avanço do desemprego no Brasil.

A instabilidade do emprego no setor privado, a falta de oportunidades no mercado de trabalho, a busca por melhor qualificação etc. motivou durante os anos a grande procura por concursos públicos no Brasil, sendo para a obtenção de emprego ou almejando uma vaga em universidade. Porém, a busca por essas oportunidades proporcionou a algumas pessoas mal-intencionadas a possibilidade de atentar contra a administração pública e conseqüentemente contra aqueles que agem de forma honesta durante os certames, provocando grandes prejuízos a todos devido aos métodos fraudulentos praticados por estas pessoas que buscam obter vantagens por meios ilícitos, dilacerando expectativas e sonhos que são refletidos em frustrações e descrenças nos certames.

Neste contexto, nos debruçamos perante o tema proposto em nosso trabalho de forma a subsidiar novas metodologias para a detecção de fraudes em teste, em especial a concursos públicos no Brasil, proporcionando melhorias no combate a fraudes em certames para oferecer mais lisura ao mesmo. Desse modo, inseriremos uma nova forma de abordagem frente aos esquemas fraudulentos praticados por quadrilhas especializadas e assim dando melhor suporte aos organizadores dos certames, beneficiando aqueles que não trapaceiam.

Considerando o exposto acima, abordaremos dentre os vários índices estatísticos propostos na literatura desenvolvidos com base na Teoria Clássica dos Testes ou na Teoria da Resposta ao Item, explanaremos alguns destes. Porém, para a geração e apresentação de resultados de dados simulados através do software R, com a utilização do pacote estatístico CopyDetec, adotaremos somente os índices baseados na Teoria da Resposta ao Item, por ser

mais completa que a Teoria Clássica dos Testes, assim mostraremos aqueles considerados como grandes promissores na detecção de fraudes,

1.3 Objetivos:

1.3.1 Geral

- a) Apresentar a evolução ao longo do tempo de técnicas propostas para a detecção de fraudes em teste;
- b) Nortear e incentivar novas pesquisas referentes ao assunto.

1.3.2 Específico

- a) Apresentar alguns métodos baseados em índices estatísticos implementados no software R, através de seu pacote estatístico CopyDetect propostos para a detecção de fraudes em testes.

1.4 Estruturação do Trabalho:

Capítulo 1: Introdução: Este capítulo apresenta uma breve contextualização sobre a importância de métodos avaliativos e nocividade de atos burlatórios, assim como a motivação, objetivo e estruturação desta pesquisa.

Capítulo 2: Universo das Fraudes: Este capítulo apresenta uma contextualização sobre fraude, explanando seus primórdios, suas características (tipos) e sua previsão legal. Empenhamo-nos em mostrar algumas variantes (possíveis golpes) dando ênfase aos casos de fraudes a concursos públicos no Brasil, em um mundo globalizado e interligado pela expansão tecnológica, em especial a internet.

Capítulo 3: Da Teoria Clássica dos Testes (TCT) a Teoria da Resposta ao Item (TRI): Neste capítulo nos reportaremos a TCT e modelos adotados e a TRI, onde abordaremos alguns modelos utilizados pela TRI, tais como: modelo logístico, modelo de resposta nominal e o modelo de resposta gradual.

Capítulo 4: Métodos estatísticos para a Detecção de Fraudes em Testes: Neste capítulo nos reportaremos a alguns testes baseados na TRI ou não.

Capítulo 5: Implementação Computacional: Neste capítulo falaremos especificamente sobre o software R com a utilização de seu pacote estatístico CopyDetect, o qual será utilizado em nosso estudo para apresentar resultados de alguns índices estatísticos além de considerações pertinentes e sugestões de possíveis abordagens futuras.

E por fim, segue o Apêndice A com o algoritmo implementado no software R, através do CopyDetect.

CAPÍTULO II – O UNIVERSO DAS FRAUDES

2.1 Breve Histórico sobre Fraude.

Desde o advento das sociedades humanas, nota-se práticas sociais que em muito intencionavam burlar regras e leis sociais consuetudinárias, e mesmo, fundamentadas em códigos de leis instituídos, em especial à economia e nos negócios se perpetua entre os povos, através de indivíduos chamados de “golpistas”, pessoas que destinam literalmente suas vidas a executar práticas ilícitas como: fraudes, ciladas, esquemas, métodos, etc., com o intuito de gerar benefício próprio em decorrência de prejuízos alheios. Estudiosos recentemente descobriram que em meados de 500 A. C., egípcios praticavam fraudes, por meio de vendas de falsos gatos e outros animais sagrados embalsamados, a ricos e pobres, para utilização em cerimônia lutuosa.³

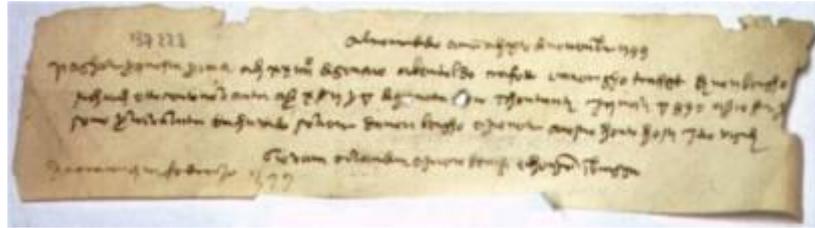
As trapaças ou fraudes se revelam também na mitologia Greco-Romana, através de Hercules ou Mercúrio, considerado Deus dos patifes e burladores, o qual fazia uso de práticas burlatorias enganando vários outros deuses, provocando constantes desentendimentos com Zeus (Deus superior). Nesta óptica podemos ressaltar outros deuses como: Loki do antigo nórdicos europeu, Eshu na África da mitologia Yoruba da qual originou o Candomblé brasileiro, na China Sun-Wukong, na Austrália Bamapana, na Índia Indra, etc.; Xenofonte (427-355 A.C.), em seus assentos referentes a conflitos (guerra) orientava seus líderes guerreiros a obter êxito (trapaças) em suas batalhas através da inocência de seus adversários. Vale ressaltar o grande Cícero (106-43 A.C.), o qual expressa seu pensar no livro “De Officiis”, capítulo 41, da seguinte forma: "Duas ainda são as maneiras com as quais pode-se fazer injustiça: a violência e a fraude; a fraude é própria da raposa e a violência do leão; ambas são adversas a natureza humana, mas a fraude desperta maior repulsão."⁴

Fraudes aplicadas em diversos meios para a obtenção de vantagens eram comuns, principalmente a pesos, medidas, alimentos e bebidas durante a idade média. O advento e conseqüentemente a expansão das letras de câmbio, em meados de 1100 D.C, iniciou-se outra forma de fraude chamada de “documental”.

³Monitor das Fraudes. Disponível em: <http://www.fraudes.org/showpage1.asp?pg=2> acesso: 25-02-2016.

⁴Idem

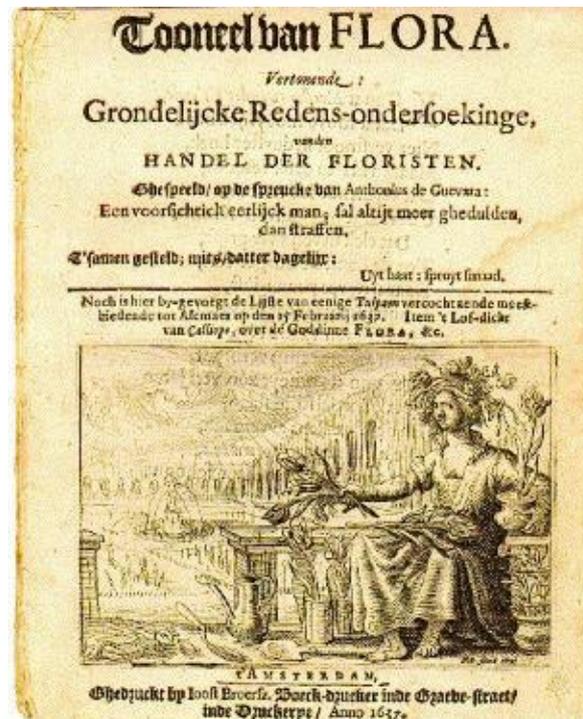
Imagem 2.1 – Letra de Câmbio do século XII.



Fonte: <http://www.fraudes.org/showpage1.asp?pg=2>.

Surgiu na Holanda, durante o período de 1637, a chamada “bolha” das Tulipas, fraude grupal que arruinou vários investidores que empregaram seus erários no mercadejo dos bulbos desta flor.

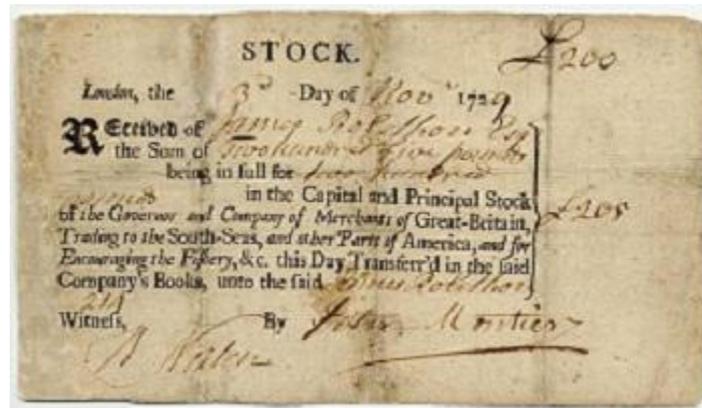
Imagem 2.2 – Panfleto da Tulipamania.



Fonte: <http://www.fraudes.org/showpage1.asp?pg=2>.

A chamada bolha da “South Sea Co.”, entidade inglesa destinada ao ramo de navegação e comércio, durante o século XVIII (1720), aplicava outro grande golpe (fraude) através de propagandas mentirosas, induzindo os investidores a adquirir mais ações e a valores mais altos, ocasionando grandes receitas para a instituição através de práticas ilícitas, porém em determinado momento a fraude veio à tona e consequentemente os investidores perderam todo seu investimento.

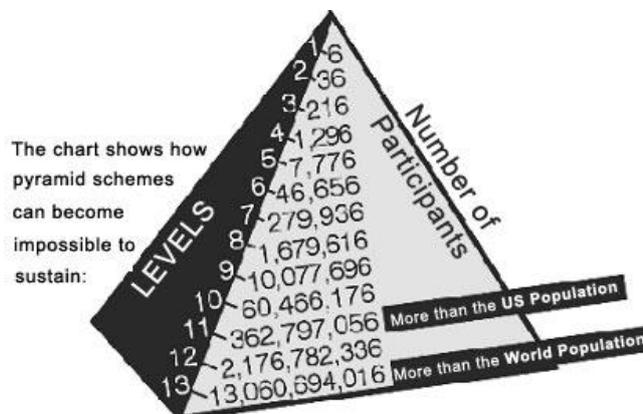
Imagem 2.3 – Ação da South Sea Company.



Fonte: <http://www.fraudes.org/showpage1.asp?pg=2>.

É plausível ressaltar o surgimento de outra relevante fraude durante o século XX (1920), o golpe denominado “Esquema de Ponzi”, aplicado nos EUA, causando enormes malefícios a milhares de pessoas que se deixaram iludir com o esquema. O “Esquema de Ponzi”, também conhecido como “Pirâmide” é constituído de diversos tipos de golpes, o mesmo pode ser definido como: “um encadeamento entre pessoas, cujo intuito é o recrutamento crescente de indivíduos para o esquema, o qual prega promessas que não podem ser cumpridas”⁵. Este tipo de fraude muitas vezes é encapotado com outro tipo de estrutura que estabelecem vendas cruzadas como marketing multiníveis, que são regulamentadas, neste tipo de esquema o indivíduo é induzido a efetuar um único pagamento, com a promessa de receber benefícios muitos superiores aos investidos, porém o teórico benefício não ocorre e quando ocorre é apenas para o arquiteto do esquema ou algumas pessoas que se beneficiam.

Imagem 2.4 – “Esquema de Ponzi”.



Fonte: https://pt.wikipedia.org/wiki/Esquema_em_pir%C3%A2mide.

⁵ Wikipédia, a enciclopédia livre. Disponível em: https://pt.wikipedia.org/wiki/Esquema_em_pir%C3%A2mide
Acesso: 26-02-2016

Conforme relatos acima explanados, observa-se que a prática de fraudes ocorre há bastante tempo. A evolução humana e tecnológica ocasionou o aperfeiçoamento dos golpes, pois os fraudadores são amoldáveis a novas situações e com isso se aproveitando de novas possibilidades.

Fraudes se tornaram globalizadas, sendo operados esquemas recorrentes, de forma adaptáveis em várias partes do mundo por integrantes de quadrilhas de diversas nacionalidades ou não. Assim, golpistas cada vez mais aperfeiçoados estão sempre em busca de novas vítimas com o único intuito de tirar proveitos da ingenuidade das pessoas.

2.2 Definindo Fraude

O termo fraude provem de tempos remotos e sua definição não é trivial, pois desde os primeiros relatos sobre o fato, são atrelados acontecimentos variados com causas interligadas, que envolvem vertentes éticas, legais, institucionais dentre outras. (BLANQUÉ, 2002; ROCKNESS e ROCKNESS, 2005) (apud FREITAS, 2012, p. 27)

O órgão americano encarregado de estipular condutas em auditorias a *American Institute of Certified Public Accountants-AICPA* apud FREITAS (2012, p. 27), descreve fraude da seguinte maneira:

A fraude é qualquer ato ou omissão intencional planejado para causar engano a terceiros. Ordinariamente, a fraude envolve a deturpação intencional, a ocultação deliberada de um fato relevante com o propósito de induzir outra pessoa a fazer ou deixar de fazer algo em detrimento dele ou dela.

Segundo a renomada *Association of Certified Fraud Examiners-ACFE* (2012, p. 6), considerada a maior autoridade em qualificação e organização de especialistas no combate a prática de atos fraudulentos, conceitua fraude como: “O uso de uma ocupação para fins pessoais, enriquecimento através do uso indevido intencional ou aplicação errada da organização empregadora de recursos ou bens.”

O site Wikipédia define fraude da seguinte forma:

Em direito penal, **fraude** é o crime ou ofensa de deliberadamente enganar outros com o propósito de prejudicá-los, usualmente para obter propriedade ou serviços dele ou dela injustamente. Fraude pode ser efetuada através de auxílio de objetos falsificados. No meio acadêmico, fraude pode se referir a fraude científica – a falsificação de descobertas científicas através de condutas inapropriadas – e, de uso comum, **fraude intelectual** significa a falsificação de uma posição assumida ou sugerida por um escritor ou

interlocutor, dentro de um livro, controvérsia ou debate, ou uma idéia apresentada enganosamente para esconder conhecidas fraquezas lógicas. Fraude jornalística implica numa noção similar, a falsificação de furos jornalísticos. Fraudes podem ser cometidas através de muitos métodos, incluindo fraude de correspondência, por meios de Tecnologia de Informações, fraude por telefone e fraude por internet.⁶

2.3 Intuitos do Golpista

Explicitaremos abaixo os intuitos relevantes do fraudador:

- Vantagem através de meios ilícitos;
- Esconder sua identidade real em meio aos esquemas ilícitos produzidos, com o intuito de dificultar sua autoria caso o golpe seja descoberto, ou seja, se tornar “invisível” aos olhos de um especialista na detecção da fraude;
- Mostrar que mesmo com o grande avanço de técnicas na prevenção de fraudes, pode ser possível burlar o sistema, por exemplo, “hackers”.

2.4 Código Penal Brasileiro (CPB)

O Decreto-Lei nº 2.848, de 7 de dezembro de 1940, regulamenta o referido código o qual versa sobre fraude no artigo 171, expresso em seu capítulo VI DO ESTELIONATO E OUTRAS FRAUDES, da seguinte forma:

Art. 171 - Obter, para si ou para outrem, vantagem ilícita, em prejuízo alheio, induzindo ou mantendo alguém em erro, mediante artifício, ardil, ou qualquer outro meio fraudulento:

Pena - reclusão, de um a cinco anos, e multa.

§ 1º - Se o criminoso é primário, e é de pequeno valor o prejuízo, o juiz pode aplicar a pena conforme o disposto no art. 155,

§ 2º - Nas mesmas penas incorre quem:

Disposição de coisa alheia como própria

I - vende, permuta, dá em pagamento, em locação ou em garantia coisa alheia como própria;

Alienação ou oneração fraudulenta de coisa própria

II - vende, permuta, dá em pagamento ou em garantia coisa própria inalienável, gravada de ônus ou litigiosa, ou imóvel que prometeu vender a terceiro, mediante pagamento em prestações, silenciando sobre qualquer dessas circunstâncias; Defraudação de penhor

III - defrauda, mediante alienação não consentida pelo credor ou por outro modo, a garantia pignoratícia, quando tem a posse do objeto empenhado; Fraude na entrega de coisa

⁶Wikipédia, a enciclopédia livre. Disponível em: <<https://pt.wikipedia.org/wiki/Fraude>> Acesso: 29-02-2016.

IV - defrauda substância, qualidade ou quantidade de coisa que deve entregar a alguém; Fraude para recebimento de indenização ou valor de seguro

V - destrói, total ou parcialmente, ou oculta coisa própria, ou lesa o próprio corpo ou a saúde, ou agrava as conseqüências da lesão ou doença, com o intuito de haver indenização ou valor de seguro; Fraude no pagamento por meio de cheque

VI - emite cheque, sem suficiente provisão de fundos em poder do sacado, ou lhe frustra o pagamento.

§ 3º - A pena aumenta-se de um terço, se o crime é cometido em detrimento de entidade de direito público ou de instituto de economia popular, assistência social ou beneficência.

2.5 Concurso Público no Brasil – Previsão Legal

2.5.1 Constituição de 1934

No século XIX durante o período Imperial no Brasil, ficava discricionária a autoridade imperial a nomeação ou exoneração para o exercício de cargos públicos, ou seja, era de praxe o imperador adotar as nomeações de funções públicas de forma direta ou indireta através de cessão sob a forma “em confiança”, desta maneira a autoridade poderia adotar medidas convenientes a si em relação a admissão ou exoneração dos agentes públicos (MELLO, 1990).

Durante a vigência da Constituição republicana do século XIX, o gestor da república poderia intervir de forma totalmente parcial na contratação ou exoneração de pessoas no serviço público, pois a mesma lhe garantia esta prerrogativa. Porém, a partir da promulgação da Constituição de 1934 foi possível nomear e exonerar imparcialmente pessoas no serviço público, configurando neste momento o surgimento do concurso público no Brasil, ou seja, a prática dos atos foi mais bem gerida, evitando a nomeação ou demissão pela pura vontade do gestor.

Dentro do contexto acima explanado, observa-se de forma bem clara a inserção legal do concurso público no Brasil, através da Constituição de 1934, onde expressa dentre outras coisas a situação de justiça e bem-estar social, deixando implícita a igualdade entre os brasileiros,

Nós, os representantes do povo brasileiro, pondo a nossa confiança em Deus, reunidos em Assembléia Nacional Constituinte para organizar um regime democrático, que assegure à Nação a unidade, a liberdade, a justiça e o bem-estar social e econômico, decretamos e promulgamos a seguinte.

Nesta vertente a referida Constituição trata de forma mais específica através do Título VII, a situação dos funcionários públicos, onde podemos destacar alguns artigos:

Art. 168 – Os cargos públicos são acessíveis a todos os brasileiros, sem distinção de sexo ou estado civil, observada as condições que a lei estatuir.

Art. 169 - Os funcionários públicos, depois de dois anos, quando nomeados em virtude de concurso de provas, e, em geral, depois de dez anos de efetivo exercício, só poderão ser destituídos em virtude de sentença judiciária ou mediante processo administrativo, regulado por lei, e, no qual lhes será assegurada plena defesa.

Parágrafo único - Os funcionários que contarem menos de dez anos de serviço efetivo não poderão ser destituídos dos seus cargos, senão por justa causa ou motivo de interesse público.

Art. 170 - O Poder Legislativo votará o Estatuto dos Funcionários Públicos, obedecendo às seguintes normas, desde já em vigor:

1º) o quadro dos funcionários públicos compreenderá todos os que exerçam cargos públicos, seja qual for a forma do pagamento;

2º) a primeira investidura nos postos de carreira das repartições administrativas, e nos demais que a lei determinar, efetuar-se-á depois de exame de sanidade e concurso de provas ou títulos;

Como explanado anteriormente, pode-se limitar a vontade do gestor em relação a constituição de cargos públicos, porém segundo o site IOB CONCURSOS,

Somente em 1967, com a sexta Constituição do Brasil, elaborada por juristas “de confiança” do regime militar, foi validada a obrigatoriedade do concurso público para ingresso em todos os cargos, excetos para os cargos em comissão (cargos de confiança) – norma mantida pela atual constituição.

Assim, seguindo esta vertente a Constituição de 1988, expressa acintosamente os requisitos para a investidura do cargo público no Brasil, adotando princípios basilares fundamentais para a justificação de seus atos, como o objetivo de evitar práticas ilícitas.

2.5.2 A Constituição Federal do Brasil de 1988

Em meio ao clamor do povo brasileiro pela democracia do país e conseqüentemente a extinção do governo militar, surge em 1988 a Carta Magna da federação, estabelecendo direitos e deveres aos cidadãos brasileiros. Por conseguinte, dentre as várias previsões legais estabelecidas pela Constituição Federal de 1988, encontram-se de forma contundente os pré-requisitos para a investidura ao cargo público, respeitando alguns princípios constitucionais fundamentais como: a eficiência, publicidade, a impessoalidade, moralidade e legalidade, em todos os atos praticados pela administração pública em qualquer esfera, conforme expressão **Capítulo VII-Seção I**, o qual faremos ressalvas ao Art. 37 e alguns de seus incisos:

Art. 37 - A administração pública direta e indireta de qualquer dos Poderes da União, dos Estados, do Distrito Federal e dos Municípios obedecerá aos princípios de legalidade, impessoalidade, moralidade, publicidade e eficiência e, também, ao seguinte: (Emendas Constitucionais: nº 18/98, nº 19/98, nº 20/98, nº 34/2001, nº 41/2003, nº 42/2003 e nº 47/2005).

I–os cargos, empregos e funções públicas são acessíveis aos brasileiros que preencham os requisitos estabelecidos em lei, assim como aos estrangeiros, na forma da lei;

II– a investidura em cargo ou emprego público depende de aprovação prévia em concurso público de provas ou de provas e títulos, de acordo com a natureza e a complexidade do cargo ou emprego, na forma prevista em lei, ressalvadas as nomeações para cargo em comissão declarado em lei de livre nomeação e exoneração;

III – o prazo de validade do concurso público será de até dois anos, prorrogável uma vez, por igual período;

IV– durante o prazo improrrogável previsto no edital de convocação, aquele aprovado em concurso público de provas ou de provas e títulos será convocado com prioridade sobre novos concursados para assumir cargo ou emprego, na carreira;

V – as funções de confiança, exercidas exclusivamente por servidores ocupantes de cargo efetivo, e os cargos em comissão, a serem preenchidos por servidores de carreira nos casos, condições e percentuais mínimos previstos em lei, destinam-se apenas às atribuições de direção, chefia e assessoramento;

VIII – a lei reservará percentual dos cargos e empregos públicos para as pessoas portadoras de deficiência e definirá os critérios de sua admissão;

IX – a lei estabelecerá os casos de contratação por tempo determinado para atender a necessidade temporária de excepcional interesse público.

Com esta Carta Magna, pode-se regular melhor o concurso público no Brasil, permitindo que qualquer cidadão dentro dos requisitos necessários para o cargo pudesse concorrer em igualdade e condições com outros, sem nenhum tipo de discriminação.

2.5.3 Código Penal Brasileiro pela Lei nº 12. 550 de 15 de Dezembro de 2011 – Fraudes em Certames de Interesse Público

Em nosso código penal o Capítulo V, o qual versa sobre fraudes em certame de interesse público, foi inserido no Título X (Dos Crimes Contra a Fé Pública), ressaltando o tipo penal segundo Art. 311-A, conforme explícito abaixo:

Art. 311-A. Utilizar ou divulgar, indevidamente, com o fim de beneficiar a si ou a outrem, ou de comprometer a credibilidade do certame, conteúdo sigiloso de:

I - concurso público;

II - avaliação ou exame público;

III - processo seletivo para ingresso no ensino superior; ou

IV - exame ou processo seletivo previstos em lei:

Pena - reclusão, de 1 (um) a 4 (quatro) anos, e multa.

§ 1º Nas mesmas penas incorre quem permite ou facilita, por qualquer meio, o acesso de pessoas não autorizadas às informações mencionadas no caput.

§ 2º Se da ação ou omissão resulta dano à administração pública:

Pena - reclusão, de 2 (dois) a 6 (seis) anos, e multa.

§ 3º Aumenta-se a pena de 1/3 (um terço) se o fato é cometido por funcionário público. “(NR)”

2.6 Fraudes em Concursos Públicos no Brasil

O concurso público se tornou uma “febre” com o decorrer dos anos, pois muitos almejam o sonhado cargo público principalmente pela sugerida estabilidade de emprego proporcionado, em meio à grande crise econômica atualmente vivida em nosso país ocasionando o aumento do desemprego, a procura por este vem se tornando uma enorme tendência como forma de superar a instabilidade do setor privado.

A prática da fraude a cada dia se torna mais presente em todos os ambientes, a busca por melhores faculdades, melhores empregos, melhores notas, etc, faz com que pessoas utilizem-se desta para efetivação de um sonho ou conquista, não se preocupando com a moral, ética ou a nocividade de seu ato a outrem para almejar seus objetivos.

Mesmo com todo o aparato legal previstos na Constituição Federal de 1988, no que tange a lisura dos atos praticados obedecendo aos princípios aqui já relatados e as conseqüências para aqueles que praticam atos fraudulentos privilegiando alguns, em concursos públicos, vem se tornando literalmente uma epidemia em nossa sociedade lesando a todos que participam de forma direta ou indireta dos certames, causando prejuízos não só financeiros, mas também psicológicos, pois para os que se dedicam fervorosamente aos estudos com o sonho de ingressar na administração pública, seja qual a esfera, se torna frustrante quando casos deste tipo são noticiados, ocasionando desestímulo e descrédito nos concursos públicos, tendo a impressão em um linguajar popular de “jogo de cartas marcadas.”

A invalidação do concurso público em casos de fraudes é pertinente, pois nesta situação é visível a intenção em privilegiar determinados candidatos durante a seleção. Segundo Carvalho Filho (2007) (apud SILVA, 2012, p. 11),

Nem sempre a Administração se tem havido com a devida legitimidade na realização de concursos públicos. Ao contrário, é comum ouvir-se reclamações de candidatos quanto a diversos aspectos dos concursos, como favorecimentos pessoais, regras de privilégio para alguns candidatos, critérios discriminatórios em editais, suspeitas de fraude, questões de prova mal formuladas etc. Por isso, têm sido significativa a descrença e a revolta, inteiramente justificáveis, aliás, de pessoas interessadas em ingressar no serviço público em relação à idoneidade e à verdadeira função seletiva de certos concursos.

[...]

Devemos distinguir, porém, a invalidação do concurso **antes** e **depois** da investidura dos aprovados.

Se a ilegalidade ocorre no curso do certame, a Administração pode invalidar o procedimento sem que esteja assegurado qualquer direito de defesa aos

participantes contra a anulação. O mesmo se passa se a ilegalidade é constatada após a sua realização, mas antes da investidura dos aprovados: a invalidação se legitima normalmente pela exclusiva atuação administrativa. A razão é que, como os candidatos e os aprovados têm mera expectativa em relação aos atos de investidura, é incoerente que lhes possa outorgar de opor-se ao desfazimento do certame.

Frente a esta situação as entidades provedoras de concursos públicos, devem fiscalizar de maneira mais eficaz a sua realização e sugerindo punições mais severas aqueles que tentam burlar o sistema.

2.6.1 Alguns Relatos de Fraudes em Concursos Públicos no Brasil

Estudantes de Medicina são Presos Suspeitos de Fraudar Vestibulares

A Polícia Federal (PF) no dia 17 de novembro de 2015, efetuou a prisão de quatro estudantes do curso de medicina da Pontifca Universidade Católica de Goiás (PUC-GO), por suspeita de prática de atos fraudulentos durante a realização de vestibulares de seis universidades públicas e privadas de Goiás, Minas Gerais e Distrito Federal. A prisão dos estudantes ocorreu em suas residências, localizadas na cidade de Goiânia-GO, local em que a Polícia Federal executou também mandados de busca e apreensão. Conforme investigações feitas pela instituição, os discentes apresentavam idade ente 19 e 23 anos e cursavam o 2º semestre de medicina na cidade onde residiam, posteriormente os discentes foram encaminhados a sede da Polícia Federal para prestarem esclarecimentos sobre o fato em tela.

Segundo o Delegado Regional de combate ao crime organizado da Polícia Federal, Jocenildo Cavalcante, “Levantou a suspeita durante o procedimento de fiscalização porque eles se comportaram de maneira estranha, saíram rápido. E um deles levou um recorte de uma das provas que já é uma burla do edital. Isso já levantou uma suspeita.”⁷

Ao ter conhecimento do fato a sr^a Olga Ronchi, vice-reitora da PUC-GO, determinou de imediato a instauração de uma comissão para a apuração do incidente. Conforme esta,

Foi detectado forte indício de que havia alguma anomalia que feria inclusive o edital. A PUC nomeou uma equipe que trabalhou a partir daquele momento fazendo o cruzamento de dados e chegamos a indícios bem concretos de que havia um grupo para, provavelmente, uma tentativa de fraude.⁸

⁷ Estudantes de medicina são presos suspeitos de fraudar vestibulares. **Disponível em:** <<http://g1.globo.com/goias/noticia/2015/11/estudantes-de-medicina-sao-presos-suspeitos-de-fraudar-vestibulares.html>>. **Acesso:10-05-2016.**

⁸ Idem

Imagem 2.5 – Estudantes estão Presos na Sede da Polícia Federal, em Goiânia (Foto: Paula Resende/G1).



Fonte: <http://g1.globo.com/goias/noticia/2015/11/estudantes-de-medicina-sao-presos-suspeitos-de-fraudar-vestibulares.html>.

A autoridade policial a frente do caso informou que, “Eles ficam durante o período mínimo necessário para poder sair com o conteúdo em mãos e depois repassam a outras pessoas que faziam as transmissões dos conteúdos sigilosos aos beneficiados.” Através de depoimentos coletados pela Polícia Federal, notou-se que os estudantes haviam se inscrito também em vestibulares em Brasília e Ouro Preto-MG.

Em caso de confirmação do ato, os agentes fraudadores serão indiciados por associação criminosa, com pena prevista de 1 a 3 anos de prisão e por fraude em certame de interesse público, com malefícios a administração pública, com pena de 2 a 6 anos de reclusão e multa. Além, das sanções penais relatadas os discentes responderão administrativamente pelo ato e sendo comprovado poderão ser penalizados com a expulsão da instituição (PUC-GO).

Grupo Quebrou Sigilo do Enem

No ano de 2014, o Exame Nacional do Ensino Médio (Enem) mais uma vez voltou a ganhar repercussão no país, pois novamente foi vítima da prática de fraude. Após investigação da Polícia Civil de Minas Gerais, onde desarticulou um esquema milionário de fraude articulada por uma quadrilha especializada em fraudar vestibulares de medicina de faculdades de Minas Gerais e Rio de Janeiro, constatou-se que um dos delinquentes havia corrompido um fiscal para obtenção da prova do Enem de 2013 e cobrado entre R\$ 70.000,00 a R\$

100.000,00 pelo repasse dos gabaritos aos interessados. Para José Cláudio de Oliveira, comprador do teste, “a facilidade do controle das salas em que as provas eram feitas facilitou a fraude”.

A quadrilha utilizava como meio de repasse das respostas ligações telefônicas. Perante indícios contundentes durante as investigações a Polícia constatou que cerca de 40 pessoas seriam beneficiadas pelo esquema, rendendo aos fraudadores um lucro entre R\$ 2.800.000,00e R\$ 4.000.000,00.

Em 3 dezembro de 2014, nos estados de Minas Gerais e Rio de Janeiro, a Polícia Civil prendeu 21 pessoas acusadas de participar da fraude, nas investigações constatou-se o envolvimento de duas na fraude no Enem, deste José Cláudio acadêmico de medicina da Faculdade de Barbacena é considerado mentor do ato, o qual possuía um padrão de vida privilegiado. Em interceptações de mensagens de texto realizadas durante as investigações ele afirmou ter repassado o valor de R\$ 10.000,00 para a obtenção de dois cadernos de provas do Enem de cor amarela. O segundo principal cooperador do esquema se tratava do Sr Quintino Ribeiro Neto de 63 anos de idade, funcionário público aposentado, o qual seu papel na fraude era aliciamento de pessoas.

Durante o monitoramento dos acusados, várias conversas entre José Cláudio e aliciados foram interceptadas, onde o mentor demonstra claramente que é possível burlar o sistema e valor cobrado pelo benefício concedido, conforme alguns trechos de conversas:

O Enem é bagunçado, não tem fiscalização. “É escolinha pública que aplica”. “É vaga de federal, pode ficar tranquilo”. “São essas velhinhas de colégio que tomam conta”. “É bagunçado, é a maior zona, não tem detector de metal”. “No Enem eu vou te cobrar R\$ 70 mil. Depois não vai pagar faculdade.”⁹

O Delegado da Delegacia de Caratinga, Fernando Lima, declarou: “Não tenho dúvida alguma que houve fraude no Enem”¹⁰. Os fraudadores chegaram a oferecer benefícios em outras faculdades, inclusive algumas renomadas como a UNB e PUC de Campinas, porém segundo averiguações não foram constatados problemas nestas instituições de ensino superior.

⁹ Grupo quebrou sigilo do Enem. Disponível em:http://patriciafrancisco.jusbrasil.com.br/noticias/112266909/grupo-quebrou-o-sigilo-do-enem?ref=topic_feed. Acesso em: 12-05-2016.

¹⁰ Idem

2.7 Outras Variações de Fraudes

Explanaremos alguns tipos de fraudes dentre vários existentes, com o intuito de educar, ou seja, esclarecer de forma mais contundentes determinados tipos de golpes empregados por burladores, que se utilizam da trapaça para conquistar vantagens.

2.7.1 Esquema da Nigéria

Sua origem decorre em meados de 1920, iniciou-se através do correio normal para o fax em seguida migrou para a internet. O golpe ocorre através de uma mensagem enviada a uma determinada vítima por alguém em algum lugar do mundo (geralmente africano), o qual solicita ajuda para retirar uma grande quantidade de dinheiro que estaria retido por algum motivo, logo o golpista promete uma valorosa recompensa por tal auxílio. Assim, a vítima atraída pela idéia de dinheiro fácil fornece seus dados bancários e é induzida a pagar algumas taxas para liberação dos valores. Após perder enormes quantias que a vítima se dá conta da fraude. Em sua grande maioria a mensagem recebida é constituída no idioma, porém no Brasil já houve relatos deste golpe aplicado com texto em português.¹¹

2.7.2 Fraudes em Cartão de Crédito

Com o intuito de adquirir informações pessoais, os fraudadores instalam equipamentos sobrepostos aos caixas automáticos ou terminais de pagamentos de lojistas, assim se apropriam de senhas de clientes possibilitando a realização de diversas operações fraudulentas.¹²

2.7.3 Fraude Acadêmica

Consiste em uma prática de apropriação de texto, compra, venda de trabalho, direitos autorais, etc.¹³.

Segundo Green (2004) (apud RAMOS, 2012, p. 18), há dois critérios básicos que definem o comportamento fraudulento:

¹¹Fraudes nos canais eletrônicos. **Disponível em:** <http://www.hsbc.com.br/1/2/br/informacoes/seguranca?WT.ac=HBBR_PVLINK03#/fraudes-nos-canais-eletronicos> **Acesso: 29-02-2016.**

¹²*Idem*

¹³Wikipédia, a enciclopédia livre. **Disponível em:** <<https://pt.wikipedia.org/wiki/Fraude>> **Acesso: 29-02-2016.**

O primeiro envolve a violação de uma regra prescritiva, compulsória, reguladora e orientadora de condutas, considerando que a regra deve ser justa e aplicada com justiça. O segundo critério exige que a regra, ao ser violada, promova alguma vantagem para o violador.

Academicamente o ato de fraudar, consiste em burlar os métodos avaliativos na criação individual e coletiva (REGO, 2010).

A notoriedade fraudulenta de um acadêmico pode ser expressa como falsos conhecimentos demonstrados em avaliações, ou seja, constituem “conhecimentos” através de práticas burladoras, aparentando alto grau de intelecto. (PIMENTA e PIMENTA, 2011)

2.7.4 Fraudes em Testes Avaliativos

Frente às grandes exigências que a sociedade moderna impõe as pessoas, fazendo com que busquem cada vez mais qualificações como forma de sobrevivência em um mundo extremamente competitivo, faz com que muitas pessoas se refugiem em meios ilegais para lograr êxito na vida, a falta de escrúpulo humana cada vez mais freqüente, não mede esforços para obter benefícios em detrimento a injusto prejuízo alheio, frustrando sonhos de quem realmente se dedica para conquistar seus objetivos de forma honesta. Nesta modalidade de fraude, o golpista pode se beneficiar de varias formas como:

- Tendo acesso aos itens que compõem o teste com antecedência;
- Adquirir informações sobre teste que foram aplicados a outros indivíduos e que serão reaproveitados;
- Através de compartilhamento de conhecimento durante o teste (colando).

Segundo Callahan (2004) (apud PIMENTA; PIMENTA, 2011, p. 05):

A justificativa dada pelos estudantes que praticam fraude é o medo de um fracasso escolar implicar em fracasso econômico e do descompromisso dos professores com uma atitude ética, quando não cobram lisura no processo de avaliação. Os comportamentos antiéticos resultantes das mudanças sociais e econômicas estão se difundindo e redefinindo os valores pessoais que estão associados à obsessão pelo dinheiro iniciada nos anos 1980 e 1990.

De acordo com Luckesi (2011) (apud RAMOS, 2012, p. 27), esta modalidade de fraude definir-se-ia, como:

O termo cola, no linguajar escolar, indica o ato pelo qual o estudante copia a resposta dada por outro colega a determinado item ou prepara uma cópia prévia para usá-la como base para sua resposta. Formas comuns de

manifestação da cola são os lembretes com o resumo do conteúdo escondidos com o objetivo de servirem de ferramenta para burlar um teste, exame ou prova. Também ocorre por meio da cópia de respostas que são obtidas, consensualmente ou não, de terceiro.

2.7.5 Fraudes de Engenharia Social

Neste tipo de golpe se desenvolve quando um funcionário de uma determinada instituição é convencido pelo golpista a fornecer informações peculiares para a fraude, ou a cometer a fraude. Esta situação ocorre sem o conhecimento real que a vítima está ajudando o fraudador (CRISTOFARO, 2006). Em sua maioria grandes fraudadores possuem enormes poderes de persuasão e são insistentes até conquista de seus objetivos. Há questionamentos, porém geralmente colaboram com o golpista. Para um fraudador ter conhecimento de algumas informações a priori como nome de colaboradores, diretores, chefes, coordenadores, etc., é de grande valia e lhe dá respaldo para lograr êxito em seu plano.

CAPÍTULO III – DA TEORIA CLÁSSICA DOS TESTES (TCT) A TEORIA DA RESPOSTA AO ITEM (TRI)

3.1 Teoria Clássica dos Testes (TCT): Aporte Histórico

A partir do final do século XIX, muitos pesquisadores dedicaram-se incansavelmente em estudar de forma mais eficaz as propriedades psicológicas humanas, com o intuito de melhor entender os estímulos, no que tange as reações ou ações dos indivíduos, logo estudiosos passaram a desenvolver modelos capazes de estimar estas propriedades, também conhecidas por traços latentes, que são variáveis que não podem ser observadas diretamente, tais como: proficiência em avaliação educacional, qualidade de vida, nível de depressão, etc...(CHACHAMOVICH, 2007 apud QUADROS, 2010).

Durante este período, psiquiatras franceses e alemães, passaram a estudar o quanto a deficiência mental poderia afetar o indivíduo em nível motor, sensorial, etc., Assim, como estudiosos ingleses que verificaram a acuidade referente a medidas de diferenças individuais, na área da genética.

A Teoria Clássica dos Testes (TCT) surgiu no começo do século XX, através de Charles Spearman (1904), que propôs uma nova metodologia e novos conceitos. No decorrer desse século Thurstone (1927/1928) apresentou colaborações relevantes para construção de medidas de traços latentes, principalmente para medidas de atitude. Sua proposta consistiu na elaboração de indicadores estatísticos, que foi denominado Lei dos Julgamentos Corporativos, o qual representou valorosa contribuição probabilística para o surgimento da Teoria da Resposta ao Item (TRI). Contudo, para melhor estudar o comportamento psicológico humano, Thurstone (1927/1928) implementou instrumento para a criação de escalas psicológicas, que posteriormente foram denominados dispositivos acumulativos e de desdobramento.

A Teoria Clássica dos Testes (TCT) apresentava graves problemas para expressar indicadores psicológicos, antes da década de 30, Thurstone (1928-1959) (apud PASQUALI; PRIMI, 2003, p. 02), ressaltou uma dessas falhas da seguinte forma:

Um instrumento de medida, na sua função de medir, não pode ser seriamente afetado pelo objeto de medida. Na extensão em que sua função de medir for assim afetada, a validade do instrumento é prejudicada ou limitada. Se um metro mede diferentemente pelo fato de estar medindo um tapete, uma pintura ou um pedaço de papel, então nesta mesma extensão a confiança neste metro como instrumento de medida é prejudicada. Dentro dos limites

de objetos para os quais o instrumento de medida foi produzido, sua função deve ser independente da medida do objeto.

Nos anos 50, a Psicometria atualmente denominada Teoria Clássica dos Testes (TCT), já se apresentava bastante consolidada, principalmente pelos estudos propostos por Guilford (1936/1954) e Gulliksen (1950), porém como anteriormente mencionado apresentava limitações consideráveis. Pois, os testes aplicados apresentavam enorme dependência entre o exame aplicado e os itens que o compõe. Logo, ao aferir o intelecto, poderíamos obter variações de respostas dependendo do mecanismo utilizado, ocasionando dúvidas referentes aos métodos de análise a ser utilizado.

Mesmo Thurstone, durante a década de 30, ter observado acintosamente as deficiências de medidas em Psicologia, não conseguiu formalizar adequadamente soluções para contornar estes problemas. Só após os anos 50, que Psicometristas descobriram inicialmente maneiras de resolver esta carência, através das relevantes contribuições de trabalhos desenvolvidos por Lazarsfeld (1959), o qual abordou em sua pesquisa a teoria do traço latente e Lord (1952) e Rasch (1960), sendo de total importância seus estudos para fins de subsídio a Teoria da Resposta ao Item (TRI), denominada como Teoria do Traço Latente, delineado por Lord durante a década de 50 e formalizada por Birnbaum (1968) e Lord (1980).

Apesar de falhas relevantes, já ressaltadas na década de 30 por Thurstone, e as primeiras formalizações desenvolvidas a partir dos anos 50, não foi possível incorporar as novas metodologias propostas pela Teoria da Resposta ao Item (TRI), ferramenta estatística que futuramente iria apresentar resultados mais consistentes que a Teoria Clássica dos Testes, devido à alta complexidade dos algoritmos matemáticos sugeridos e a inconsistência tecnológica da época, ineficaz para solucioná-los de forma contundente. Porém, com a evolução no campo da informática, com surgimento de microcomputadores e softwares eficientes, tornou-se possível solucionar o grande entrave que a Psicometria enfrentava, tornando possível a utilização da Teoria da Resposta ao Item (TRI). Sendo os primeiro softwares elaborados para a TRI, em 1979 com o BICAL de Wright, Mead e Bell, em seguida com o LOGIT (Wingersky, Barton, & Lord, 1982) e posteriormente BILOG (Mislevy & Bock, 1984).

Outras Deficiências da Teoria Clássica dos Testes (TCT)

Além da inconsistência anteriormente citada, ressalta-se outras também pertinentes, tais como:

1. Se mesmos indivíduos submetidos a teste diferentes com parâmetros de dificuldade e de inclinação diferentes, produziram resultados diferentes;
2. Fidedignidade do teste, esta exige que grupos submetidos ao mesmo exame devam produzir escores verdadeiros iguais e variância constante. Contudo, há de se considerar que alguns indivíduos respondem o teste de maneira mais eficiente que outro (HAMBLETON; SWAMINATHAN, 1985);
3. Embretson e Reise (2000) (apud SARTES et.al., 2013, p. 243), averiguaram mais limitações, com as seguintes suposições:

Testes longos são mais confiáveis do que testes curtos e que as propriedades de um intervalo da escala avaliada são obtidas somente se a distribuição de frequência dos escores for normal. Além disso, instrumentos que contém itens com formatos mistos, conduzem a um escore total dos indivíduos desbalanceado.

Devido às sérias limitações da TCT, tornou-se necessário o desenvolvimento de novas estratégias estatísticas para contorná-las, surgindo assim efetivamente a Teoria da Resposta ao Item (TRI), com novas metodologias e modelos matemáticos aprimorados que permitiram estudar cada item que compunha o instrumento de medida individualmente e também comparar indivíduos submetidos a testes diferentes, desde que houvesse itens em comuns em ambos os testes aplicados.

3.2 Modelos adotados pela Teoria Clássica dos Testes (TCT)

3.2.1 Análise Fatorial

O intuito desta poderosa ferramenta estatística dentro da Psicometria é identificar as habilidades latentes básicas, mensurando sua quantidade e formas de organização. Logo, esta se torna muito pertinente para avaliar estruturas em matriz de covariância ou correlação, simplificando-as em um menor quantitativo de variáveis possíveis. Segundo Johnson e Wichern (1992) (apud PRIMI, 2012, p. 298), “o propósito essencial da análise fatorial é descrever, se possível, as covariâncias entre variáveis em termos de um número menor de variáveis aleatórias subjacentes, mas inobserváveis, chamadas fatores”.

Com o intuito de se estudar variáveis latentes não observáveis e mensurá-las, Psicometristas se empenharam em aplicar vários testes de inteligência abordando múltiplos temas. Assim, partindo do princípio que diversos testes possuem altos inter-relacionamentos podendo ser representados por uma única variável latente responsável por isso, logo a partir desta inter-relação haveria possibilidade de melhor entendimento sobre sua estrutura. Desta forma, pode-se representar o modelo fatorial ortogonal composto por p -variáveis em uma determinada amostra da seguinte forma:

$$\begin{bmatrix} X_1 \\ \vdots \\ X_p \end{bmatrix}_{px1} = \begin{bmatrix} \bar{X}_1 \\ \vdots \\ \bar{X}_p \end{bmatrix}_{px1} + \begin{bmatrix} l_{11} & \cdot & l_{1m} \\ \vdots & & \vdots \\ l_{p1} & \cdot & l_{pm} \end{bmatrix}_{pxm} \begin{bmatrix} F_1 \\ \vdots \\ F_m \end{bmatrix}_{mx1} + \begin{bmatrix} e_1 \\ \vdots \\ e_m \end{bmatrix}_{mx1}, \quad (3.1)$$

onde:

- X_i representa o valor da i -ésima variável;
- \bar{X}_i representa a média da i -ésima variável;
- l_{ij} representa a carga fatorial i no fator j ;
- F_j representa o j -ésimo fator em comum;
- e_i representa o valor do i -ésimo fator específico;
- $m < p$ representa o número de fatores menores que o número de variáveis.

3.2.1.1 Hipóteses do Modelo Ortogonal

1. Média zero para os valores comuns e específicos;
2. Variância dos fatores igual a 1 e covariância entre elas 0, fazendo com que a matriz de covariância seja igual a identidade;
3. Covariância entre os fatores específicos seja igual a 0, logo esta matriz se tornará matriz diagonal.

Assim, para uma determinada variável i , supõe-se a seguinte equação para seu valor

$$X_i = \bar{X} + l_{i1}F_1 + \cdots + l_{im}F_m + e_i \quad (3.2)$$

Observa-se que:

- A variável observada i é dependente de um grupo de m variáveis latentes compostas por m fatores em comum (F_1, \dots, F_m) e fatores específicos (e_i)
- A variabilidade da variável i é explicada pela composição de variáveis comuns, isto é, relacionada a um grupo maior de p variáveis a qual também faz parte e a variância da parte específica que não está atrelada as outras p variáveis;
- A carga fatorial está intimamente relacionada a variação de um determinado fator j com a variação na variável i podendo ser representada por l_{ij} ;
- Relação linear entre os m fatores e a variável i .

Assim, para este modelo podemos representar a matriz de variância e covariância entre p variáveis, como segue asseguir

$$Cov(X) = LL' + \varphi = \begin{bmatrix} l_{11} & \cdot & l_{1m} \\ \vdots & & \vdots \\ l_{p1} & \cdot & l_{pm} \end{bmatrix}_{p \times m} \begin{bmatrix} l_{11} & \cdot & l_{1p} \\ \cdot & \cdot & \cdot \\ l_{m1} & \cdot & l_{mp} \end{bmatrix}_{m \times p} + \begin{bmatrix} \varphi_1 & 0 & 0 & 0 \\ 0 & \cdot & 0 & 0 \\ 0 & 0 & \cdot & 0 \\ 0 & 0 & 0 & \varphi_p \end{bmatrix}_{p \times p} \quad (3.3)$$

resultando em;

$$Var(X_i) = LL' + \varphi = l_{i1}^2 + \dots + l_{im}^2 + \varphi_i \quad (3.4)$$

$$Cov(X_i, X_k) = l_{i1}l_{k1} + \dots + l_{im}l_{km} \quad (3.5)$$

$$Cov(X_i, F_j) = l_{ij} \quad (3.6)$$

Desta forma, considerando h^2 como a variância da variável i , a qual é obtida pela soma dos quadrados dos m fatores mais a variância específica, teremos

$$h^2 = l_{i1}^2 + \dots + l_{im}^2 \quad (3.7)$$

Assim, será representada como

$$Var(X_i) = h^2 + \varphi_i, \quad (3.8)$$

onde

- A primeira parcela denota a variância da variável i relacionada à variação dos m fatores;

- A segunda parcela representa a parte da variância não compartilhada pelas demais variáveis, isto é, não correlacionadas aos m fatores, assim esta parcela se torna inerente a variável em estudo.

Com o apoio da análise fatorial foi possível representar a relação entre a pontuação de testes diferentes em função de um pequeno número de habilidades latentes (PRIMI, 2012). Logo, esta ferramenta estatística proporciona um modelo para justificar as diferenças entre examinandos com relação a pontuação dos testes através de um grupo de variáveis latentes menor (F_j).

A possibilidade de estudar empiricamente as variáveis não observáveis através da análise fatorial trouxe grande contribuição a Psicometria. Segundo Primi (2012, p. 299), “Isso ocorreu porque o pesquisador podia partir de um conjunto de variáveis observáveis e, por meio das inter-relações entre elas, investigar as possíveis dimensões subjacentes que seriam as causas desses comportamentos.”

Um dos grandes modelos adotado pela Teoria Clássica dos Testes (TCT) que possibilitou estudo mais aprofundado de características não observáveis dos indivíduos, como a inteligência e a personalidade, sugere-se para mais informações sobre este modelo a leitura de Mingotti (2005).

3.2.2 Modelo linear clássico

Modelo proposto por Spearman e apresentado de forma detalhada por Gulliksen o qual prevê que a pontuação observada de um determinado indivíduo pode ser representada pela reunião de duas parcelas, conforme abaixo:

$$X_i = T_i + e_i, \quad (3.9)$$

- T_i representando a pontuação verdadeira do indivíduo no teste em relação a variável medida;
- e_i representando a contagem de erro que ocorre em função das medidas psicológicas.

Segundo Primi (2012, p. 300), T_i pode ser obtido “pela média de um conjunto de infinitas medidas da mesma variável, no mesmo sujeito, quando estas são independentes, usando um instrumento imperfeito – com erros de medidas.” Através da seguinte equação:

$$T_i = \lim_{k \rightarrow \infty} \frac{\sum_{i=1}^k X_i}{K} \quad (3.10)$$

Em referência ao modelo segue algumas suposições:

- Se o erro é aleatório em um grande grupo de medida possuirá média 0;
- Para um conjunto grande de medida é plausível esperar que não haja correlação entre os escores verdadeiros e escores dos erros, logo teremo $\rho_{te} = 0$;
- Para diferentes testes que meçam a mesma variável a correlação entre eles será 0.

Considerando as equações (3.2) e (3.9) e assumindo que os escores verdadeiros sejam representados por m variáveis latentes, podemos escrever o referido modelo através do modelo fatorial com m , como segue,

$$T_i = \bar{X} + l_{i1}F_1 + \dots + l_{im}F_m + S_i \quad (3.11)$$

Reescrevendo (2.11) de forma a representar os escores observados, teremos

$$X_i = \bar{X} + l_{i1}F_1 + \dots + l_{im}F_m + e_i \quad (3.12)$$

Considerando o componente específico e o erro próximos de zero, conseqüentemente teremos altas cargas fatoriais nos fatores subjacentes que compõem esse construto psicológico. Logo, se o teste e a variável externa medem a mesma variável latente, apresentaram cargas fatoriais altas nos fatores que compõem esta variável. Logo, a correlação ou covariância entre duas variáveis estarão ligadas através das cargas dessas variáveis nos fatores comuns subjacentes a elas. Assim, a correlação é obtida

$$Cov(X_i, X_k) = l_{i1}l_{k1} + \dots + l_{im}l_{km} \quad (3.13)$$

3.3 Teoria da Resposta ao Item (TRI)

Segundo Hambleton e Swaminathan (1985) (apud PINHEIRO, 2007, p. 01): “A TRI é um conjunto de modelos matemático usados para fazer predições, estimativas ou inferências sobre habilidades (ou competências) medidas em um teste.” A Teoria da Resposta ao Item (TRI) surge como uma nova metodologia para avaliação de teste, com o intuito de suprir as lacunas geradas pela Teoria Clássica dos Testes (TCT), objetivando analisar de forma mais

contundente os itens que compõem o instrumento de medida e não apenas o teste como um todo.

Nesta óptica, esta ferramenta estatística desde os anos 40, já vinha sendo desenvolvida através de trabalhos propostos por George Rasch, vislumbrando medidas de traços latentes para modelos dicotômicos, denominado modelo de Rasch (1960). Porém, Frederic Lord (1952, 1953) foi o pioneiro a elaborar o modelo unidimensional de 2 parâmetros para análise de itens dicotômicos, ou seja, verificados com certo ou errado e também formas de estimação para estes parâmetros utilizando a função ogiva normal. Contudo, Lord observou pertinente a inserção de mais um parâmetro ao modelo, denominado parâmetro de acerto casual, constituindo assim uma expansão para um modelo de dois parâmetros (PASQUALI; PRIMI, 2003).

Birnbaum (1968) inseriu uma nova metodologia ao modelo desenvolvido por Lord, adotando a função logística em substituição a ogiva normal, por ser simples de trabalhar, por ser uma função explícita dos parâmetros dos itens e de habilidade não sendo necessária integração. Ressalta-se que Rasch (1960) elaborou um modelo unidimensional de 1 parâmetro utilizando a função ogiva normal, o qual foi adaptado a função logística por Wright (1968).

Posteriormente aos trabalhos acima enunciados, demandou-se a necessidade de elaborar modelos que avaliassem itens não apenas de forma dicotômica, logo foram desenvolvidos os modelos de resposta gradual por Samegima (1969), com o intuito de adquirir mais informações das respostas dos examinados (ANDRADE et al, 2000). Assim, como modelos para mais de uma categoria de resposta propostos por Bock (1972), Masters (1982) e Muraki (1992), com diferentes estruturas entre as categorias. Além destes, Bock e Zimowski (1997), propuseram os modelos logísticos de 1, 2 e 3 parâmetros para mais de uma população de respondentes, isso viabilizou a comparação de populações submetidas a distintos exames, porém com itens em comuns, o qual foi também debatido por Hedges e Vevea (1997) e Andrade (1999).

A disseminação da TRI pelo mundo nas últimas décadas, como uma das técnicas estatísticas mais consistentes em avaliação de teste, proporcionou pela primeira vez ao Brasil em 1995, a inserção desta ferramenta em avaliação educacional, oferecendo a possibilidade de analisar dados do Sistema Nacional de Ensino Básico – SAEB. Com a implementação deste

método foi possível avaliar desempenho de alunos de séries diferentes do ensino fundamental, comparando e colocando-os em uma mesma escala de conhecimento (ANDRADE et al., 2000).

Conforme Pasquali e Primi (2003, p. 100), uma das maiores contribuições a TRI nos dias atuais, foi o avanço da informática a partir da década de 80, através de microcomputadores com grande capacidade de processamento e surgimento de softwares apropriados, pois devido à alta complexidade dos modelos matemáticos adotados sem o aparato informático adequado seria bastante dificultoso alcançar grandes conquistas científicas.

3.4 Modelos Matemáticos Adotados pela Teoria da Resposta ao Item

3.4.1 Aporte Conceitual

A Teoria da Resposta ao Item (TRI) possui em sua gênese modelos matemáticos que consistem em estabelecer uma relação entre a probabilidade de um indivíduo explicitar uma determinada resposta como função dos parâmetros dos itens e de seu traço latente (habilidade, proficiência, etc...), ou seja, conforme aumenta o intelecto do indivíduo maior a propensão de resposta correta ao item.

De acordo com Pinheiro (2007, p. 04), os modelos adotados pela TRI consideram alguns fatores relevantes, tais como:

- A composição do item (dicotômico ou não);
- Quantidade de populações a ser examinadas;
- Quantitativo de traços latentes a ser mensurada;
- Número de parâmetros dos itens.

3.4.2 Modelos para Avaliação de Itens Dicotômicos

Segundo Silva (2013, p. 10): “A Teoria de Resposta ao Item consiste na construção de modelos, munidos de parâmetros que facilitam a interpretação do item. Existem modelos da TRI para diferentes tipos de itens: dicotômicos, politômicos ou para respostas contínuas.”

Os modelos logísticos para itens analisados de forma dicotômica são os mais utilizados, apresentando três variações os quais se distinguem pelo número de parâmetros que

descrevem o item. Tais funções denominadas logísticas de 1, 2 e 3 parâmetros, que consideram, respectivamente (ANDRADE et al., 2000):

- Somente a dificuldade do item;
- A dificuldade e a discriminação do item;
- A dificuldade, a discriminação e a propensão um indivíduo baixa habilidade acertar um item.

Dentre os modelos logísticos acima explanados, nos reportaremos ao modelo logístico de 3 parâmetros por ser atualmente o mais utilizado e completo, sendo os demais possíveis de ser obtidos a partir dele, o qual possui a seguinte configuração:

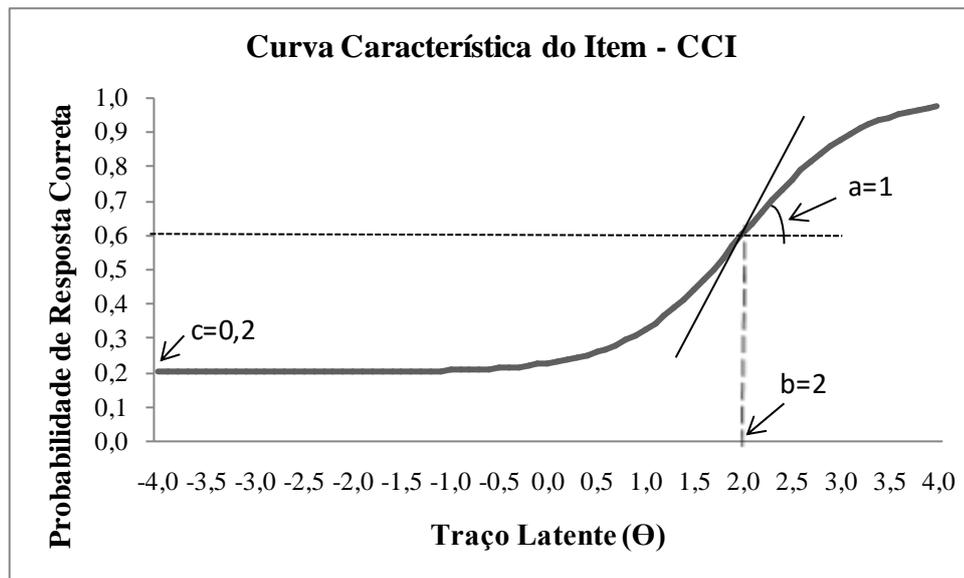
$$P(U_{ij} = 1|\theta_j) = c_i + (1 - c_i) \frac{1}{1 + e^{-D a_i(\theta_j - b_i)}}, \quad (3.14)$$

Sendo $i = 1, 2, \dots, I$, são os itens que compõem o teste e $j = 1, 2, \dots, n$, os elementos que compõem a amostra, podendo ser indivíduos, empresas e etc..., onde:

- U_{ij} é uma variável dicotômica que assume valor 1 quando o indivíduo j responde corretamente, concorda ou satisfaz as condições do item i , ou 0 caso contrário;
- θ_j é a habilidade do respondente (traço latente) do j -ésimo indivíduo;
- $P(U_{ij} = 1|\theta_j)$ é a probabilidade do respondente j , condicionada seu traço latente θ_j , responder corretamente, satisfazer ou concordar com as condições do item i , a qual é conhecida na TRI como Função de Resposta ao Item – FRI;
- b_i é o parâmetro de dificuldade (ou de posição) do item i , o qual é medido na mesma escala do traço latente;
- a_i é o parâmetro de inclinação ou discriminação do item i , com valor proporcional a inclinação da curva característica do item no ponto b_i ;
- c_i é o parâmetro de acerto casual;
- D é uma constante de valor 1, porém utiliza-se o valor 1,7 quando queremos aproximar os valores da função logística com os da ogiva normal.

A função 3.14, pressupõe que $P(U_{ij} = 1|\theta_j)$ é a proporção de respostas correta ao item i dentre todos os examinados da população com traço latente θ_j .

Figura 3.1 Exemplo de Curva Característica do Item (CCI).



Fonte: Próprio Autor do trabalho.

Considerações com relação a Figura 3.1:

A presente figura mostra claramente que a relação proposta pela função logística não é linear;

1. Observa-se que a probabilidade de resposta correta ao item aumenta com o crescimento do traço latente (θ_j) do indivíduo;
2. A curva apresenta deslocamento e inclinação na mesma escala do traço latente em análise definidos pelos parâmetros dos itens;
3. O parâmetro c_i , mostra que examinados com baixo traço latente, podem responder corretamente o item;
4. O parâmetro a_i , informa que quanto maior o seu valor maior será a possibilidade de discriminar indivíduos com alto, médio ou baixo traço latente para afirmativa positiva a um item proposto. Ressalta-se que este pode assumir valores entre $\pm\infty$, porém na prática adota-se valores entre 0 e 3, pois valores negativos, mostra que o aumento de habilidade do examinado diminui a probabilidades de resposta certa ao item;
5. O parâmetro de dificuldade (b_i), é medido na mesma escala do traço latente, admitindo valores entre $\pm\infty$, porém normalmente é adotado entre ± 4 , este designa o nível de traço latente para a propensão de êxito em um item seja $(1 + c) / 2$, assim um valor elevado deste parâmetro requer um alto traço latente para resposta correta.

3.4.3 Definição da Métrica para o Traço Latente (Θ)

Diversos autores tratam da métrica, Segundo Andrade et al (2000) (apud MOREIRA JUNIOR, 2013, p.146):

Os valores da escala do traço latente (ou habilidade) podem assumir teoricamente qualquer valor real entre $-\infty$ e $+\infty$, diferentemente da TCT, em que a escala geralmente varia entre 0 e a quantidade total de questões do teste. No entanto, o importante nessa escala não é a sua magnitude, mas as relações de ordem existentes. Dessa forma, é preciso estabelecer valores para a média e para o desvio padrão que representem a escala de habilidade dos indivíduos na população.

Na prática por conveniência e facilidade computacional, primeiramente estima-se os parâmetros dos itens dentro de uma escala proposta, geralmente adota-se a $\mu=0$ e desvio padrão $\sigma=1$, posteriormente é proposta a escala para habilidade dos indivíduos dentro da mesma métrica dos parâmetros dos itens.

A métrica do traço latente possibilita averiguar ordenadamente do menor para o maior a performance dos indivíduos, ou seja, o indivíduo situado em determinado nível da escala apresenta habilidade necessária para este nível e outros inferiores a ela (FONTANIVE et al., 2007).

3.5. Hipóteses Relevantes para os Modelos de Resposta ao Item.

A TRI apresenta em sua gênese pressupostos relevantes, dentre os quais podemos explicar em especial: a unidimensionalidade e a independência local. Como estas duas suposições não são demonstradas formalmente, então aceita-se ou não as mesmas, porém a averiguação de viabilidade de um modelo matemático, consiste em analisar suas hipóteses e seus dados empíricos. Contudo, mesmo sem a formalidade necessária elas podem ser verificadas de forma indireta, analisando se o descumprimento delas colabora para a produção de resultados inconsistentes, assim os resultados práticos determinarão sua adequabilidade (PASQUALI; PRIMI, 2003).

3.5.1 Unidimensionalidade do Teste

Este pressuposto informa que existe um conjunto de itens homogêneos proposto para mensurar um único traço latente, isto é, pressupõe-se que deve haver uma única habilidade pela realização dos itens de um teste. Porém, é notório que o comportamento humano é multideterminado ou multimotivado, desta forma fica evidente que há mais de um traço

latente influenciando a realização de alguma tarefa. Todavia, para subsidiar esta hipótese se faz necessário entender que haja apenas um fator dominante (traço latente) em análise perante um conjunto de itens propostos, o qual presume-se está sendo mensurado pelo exame (ANDRADE et al, 2000).

3.5.2 Independência Local

Outra suposição relevante dentro da TRI é a independência local, a qual sugere que não havendo variação de aptidão em um teste, com exceção do traço latente dominante, as respostas dos examinados submetidos aos itens são respondidas de forma independente, isto é, sua performance em um item proposto não influenciará em afirmativa positiva ao item posterior. Ressalta-se a nequice desta suposição, se o item apresenta em seu teor propensão de influenciar a resposta correta do examinado ou subsidiar a resposta a um item posterior. Assim, ocorrerão situações de indivíduos perceberem esta influência e outros não, logo se deve atentar para esta situação e a fim de validar esta suposição a coerência na elaboração de itens de um teste é de total importância, pois sendo verdadeira a independência local, pode-se obter a propensão de respostas do indivíduo a um conjunto de itens através do produto das probabilidades individuais de cada item (Andrade et al, 2000).

3.6 Alguns Modelos de Resposta ao Item para Avaliação não Dicotômica

Estes modelos subsidiam duas maneiras de análise, sendo uma para itens abertos, isto é, respondidos de forma livre e outra para itens de múltipla escolha avaliados de forma graduada, ou seja, itens implementados ou corrigidos com intuito de apresentar uma ou mais categorias intermediárias ordenadas no âmbito das categorias certo ou erradas. Para este item, considera-se não somente a afirmativa certa ou não pelo examinado, mas também sua opção de escolha, isto é, possibilita verificar o que influenciou na sua decisão para a escolha de uma determinada opção de resposta, podendo assim avaliar de forma mais detalhada a habilidade do indivíduo. Perante os modelos propostos pela TRI, abordaremos especificamente apenas dois deles, conforme abaixo.

3.6.1 Modelos de Resposta Nominal (MRN)

Este modelo propõe a obtenção de estimativas mais consistentes para avaliação de traços latentes de indivíduos sujeitos a testes de múltipla escolha. O qual considera de forma individual a probabilidade de resposta para cada alternativa de um item em função tanto dos

parâmetros dos itens quanto do traço latente, o que possibilita uma análise mais minuciosa em relação aos modelos dicotômicos que verificam apenas se os indivíduos responderam o item de forma certa ou errada (BOCK, 1972 apud COUTINHO, 2013, p. 44).

Bock considerou que a probabilidade com que um indivíduo j selecionaria uma determinada opção k (dentre as m_i opções possíveis) do item i poderia ser expressa da seguinte forma (Andrade et al, 2000):

$$P_{i,k}(\theta_j) = \frac{e^{a_{i,k}^+(\theta_j - b_{i,k}^+)}}{\sum_{t=1}^{m_i} e^{a_{i,t}^+(\theta_j - b_{i,t}^+)}} \quad (3.15)$$

com $i = 1, 2, \dots, I$, $j = 1, 2, \dots, n$ e $k = 1, 2, \dots, m_i$. Sendo $\sum_{k=1}^{m_i} P_{i,k}(\theta_j) = 1$, representando a soma das probabilidades sobre as m_i opções em cada θ_j (traço latente do individuo j) e $(a_{i,k}^+, b_{i,k}^+)$ os parâmetros dos itens para a k -ésima opção. Ressalta-se que para este modelo não há a exigência a priori de ordenação das opções de resposta.

Segundo (Bock, 1972; De Ayala e Sava-Bolesta, 1999) (apud Coutinho, 2013, p. 44) “mostraram que o aumento na precisão das estimativas dos traços latentes pode ser obtido quando as informações sobre o construto contidas nas categorias incorretas são incluídas no modelo”. Assim, com a inclusão de categorias incorretas ao modelo permitirá precisar melhor as estimativas para os traços latentes e averiguar o desempenho do indivíduo. Vale destacar que a opção de escolha por determinada categoria de resposta está intimamente ligada ao traço latente do sujeito.

3.6.2 Modelos de Resposta Gradual (MRG)

Modelo proposto por Samejima (1969), o qual pressupõe a ordenação entre si a priori das categorias de respostas do item. Assim, como o modelo de resposta nominal este tenta obter o máximo de informação das respostas dos indivíduos e não somente averiguar se responderam corretamente ou não o item proposto. Logo, se ordenarmos de forma crescente os escores das categorias de um item i e denotamos por $k = 0, 1, 2, \dots, m_i$, sendo $(m_i + 1)$ o quantitativo de categorias do i -ésimo item. Desta forma a propensão de um indivíduo j escolher uma determinada categoria ou mais alta do item pode ser expressa por uma expansão do modelo logístico de 2 parâmetros, conforme abaixo (ANDRADE et al, 2000):

$$P_{i,k}^+(\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_{i,k})}}, \quad (3.16)$$

Com $i = 1, 2, \dots, I$, $j = 1, 2, \dots, n$ e $k = 1, 2, \dots, m_i$. Sendo $b_{i,k}$ o parâmetro de dificuldade da k-ésima categoria do item.

Os outros parâmetros são denominados semelhantes aos definidos anteriormente. Destaca-se que o parâmetro de inclinação do item pode ser nomeado como discriminação em modelos para itens dicotômicos. Porém, em modelos de resposta ao item não dicotômico, o parâmetro de discriminação de uma determinada categoria de resposta é dependente dos parâmetros de inclinação comuns a todas as categorias do item, assim como do espaçamento constantes das categorias de dificuldade adjacentes.

Por definição este modelo deve apresenta a seguinte configuração, para as categorias de dificuldades do item:

$$b_{i,1} \leq b_{i,2} \leq b_{i,3} \dots \leq b_{i,m_i},$$

Para tal como foi discorrido anteriormente é necessário a ordenação entre o nível das categorias de um item, conforme a classificação de seus escores.

Pode-se expressar a propensão de um indivíduo j auferir um escore k no item i , conforme equação abaixo:

$$P_{i,k}(\theta_j) = P_{i,k}^+(\theta_j) - P_{i,k+1}^+(\theta_j).$$

Samejima define $P_{i,0}^+(\theta_j)$ e $P_{i,m_i+1}^+(\theta_j)$ de modo que:

$$P_{i,0}^+(\theta_j) = 1$$

e

$$P_{i,m_i+1}^+(\theta_j) = 0$$

Assim,

$$P_{i,0}(\theta_j) = P_{i,0}^+(\theta_j) - P_{i,1}^+(\theta_j) = 1 - P_{i,1}^+(\theta_j).$$

e

$$P_{i,m}(\theta_j) = P_{i,m}^+(\theta_j) - P_{i,m_{i+1}}^+(\theta_j) = P_{i,m}^+(\theta_j).$$

Teremos:

$$P_{i,k}(\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_{i,k})}} - \frac{1}{1 + e^{-Da_i(\theta_j - b_{i,k+1})}}, \quad (3.17)$$

Vale considerar que para um item com $(m_i + 1)$ categorias, precisarão ser estimados m_i valores de dificuldade, além do parâmetro de inclinação do item. Logo, a estimação do número de parâmetros de cada item, será feita conforme o quantitativo de categorias de resposta. Ressalta-se para esse modelo, não há necessidade dos itens possuírem o mesmo quantitativo de categorias de respostas.

Ressalta-se que existem outros modelos propostos na literatura que envolvem estudos abordando mais de uma população (ver Andrade et al. (2000)).

3.7. Estimação de parâmetros: Itens e Habilidade

Etapas de relevância na TRI, são a estimação de parâmetros de itens e habilidades dos indivíduos. Para tal considera-se que as respostas dos indivíduos aos itens dependem somente dos parâmetros dos itens e habilidade, porém normalmente ambos são desconhecidos, sendo apenas as respostas aos itens conhecidas.

Perante as etapas ao norte citadas, pode-se estimar os parâmetros em três etapas:

- Quando já conhecemos os parâmetros dos itens e devemos estimar as habilidades;
- Quando conhecemos as habilidades e devemos estimar os parâmetros dos itens;
- Estimação simultânea dos parâmetros.

Para a estimação em ambos os casos, normalmente adota-se o Método da Máxima Verossimilhança com a utilização de métodos numéricos, tais como algoritmo de Newton-Raphson (ver Isaac e Keller (1996), por exemplo) e “Scoring” de Fisher (ver Rao (1973), por exemplo), podendo também ser utilizados com frequência métodos Bayesianos (ver Mislevy (1986), por exemplo).

Para a estimação conjunta dos parâmetros dos itens e habilidade, adota-se geralmente duas abordagens: estima-se primeiramente os parâmetros dos itens e posteriormente as habilidades. Porém, para este caso segundo Andrade et al. (2000, p. 28) “o número de parâmetros a serem estimados simultaneamente pode ser extremamente grande $(3I + n)$, para

o ML3), levando a uma enorme exigência computacional que envolve a inversão de matrizes dessa ordem”. De forma a contornar esta dificuldade Birnbaum (1968), sugeriu a adoção de escores padronizados, por exemplo, com estimativas iniciais para as habilidades, podendo assim estimar os parâmetros dos itens com as habilidades conhecidas e em seguida estimando as habilidades com os parâmetros dos itens conhecidos, processo denominado vai e volta. Logo, o método é reproduzido até que seja alcançado um critério de parada, desta maneira considerando a Independência Local é possível obter as estimativas dos itens individualmente. De maneira análoga, porém considerando as respostas aos itens provenientes de indivíduos distintos como independentes, pode-se estimar as habilidades individualmente. Todavia, esta situação apresenta segundo Andrade et al (2000, p. 28) um grande problema,

Sabe-se que, para os parâmetros dos itens conhecidos, os Estimadores de Máxima Verossimilhança (EMV) das habilidades convergem (ver Sen & Singer (1993), por exemplo) para os seus verdadeiros valores quando o número de itens cresce; com as habilidades conhecidas, os EMV dos parâmetros dos itens, ζ_i , convergem para os seus verdadeiros valores quando o número de indivíduos cresce. Na estimação conjunta, as habilidades são denominadas de parâmetros incidentais, pois o número destes parâmetros (θ_j) cresce com o número de indivíduos; os parâmetros dos itens são denominados de parâmetros estruturais, e o número desses parâmetros não se altera quando a amostra cresce.

Fato sinalizado por Neyman e Scott (1948), os quais em uma metodologia distinta ao da TRI, observaram que o EMV dos parâmetros dos itens podem ser influenciados pelos parâmetros incidentais ocasionando sua tendenciosidade assintótica, ou seja, mostrando inconsistência. Todavia, Segundo Lord (1968) (apud ANDRADE et al, 2000, p. 28), “quando o número de itens e o número de indivíduos crescem, os EMV dos parâmetros dos itens e das habilidades podem ser não-viciados.”

Bock & Lieberman (1970), observando a inconsistência dos estimadores em uma fase da estimação, propuseram um método para obtenção dos estimadores em duas etapas, considerando a existência de uma distribuição vinculada a proficiência dos examinados da população em análise (ver Andersen (1980) para maiores detalhes), desta forma fornecendo suporte para a utilização do Método da Máxima Verossimilhança Marginal para a estimação dos itens, o qual através da distribuição adotada para a proficiência dos examinados da população, cuja a função densidade de probabilidade (fdp) denominada $g(\theta|\boldsymbol{\eta})$, sendo $\boldsymbol{\eta}$ considerado o conjunto de parâmetros da referida população em estudo e integrando a função de verossimilhança em relação a θ . Posteriormente a esta etapa, pode-se estimar as habilidades dos indivíduos individualmente através de um dos três métodos: máxima

verossimilhança, moda ou média da distribuição condicional de θ_j dado o vetor de resposta do indivíduo j , $\mathbf{u}_j = (u_{j1}, \dots, u_{jI})$, com $j = 1, \dots, n$ e o vetor de parâmetros de itens, $\boldsymbol{\zeta}_i = (a_i, b_i, c_i)$, com $i = 1, \dots, I$ conhecidos (ver Andrade et al. (2000) para maiores detalhes) .

Para efeito de exemplificação dos modelos até aqui explanados serão gerados dados simulados para possível constatação de fraudes no capítulo V. Porém, dentre os modelos propostos para a detecção de fraudes, serão abordados apenas aqueles elaborados com base na Teoria da Resposta ao Item.

CAPÍTULO IV – MÉTODOS ESTATÍSTICOS PARA A DETECÇÃO DE FRAUDES EM TESTE.

4.1 Introdução

A preocupação em obtenção de conceitos inerentes aos indivíduos com emprego de teste origina-se nos primórdios dos tempos, mais precisamente durante o século XVI por parte de jesuítas, pois os mesmos objetivavam o combate às práticas heréticas em especial ao protestantismo, com orientação escolar através de ritual de exames e provas.

Em contrapartida aos métodos avaliativos dos jesuítas, indivíduos em algumas comunidades sociais chamadas de tribo, eram considerados maduros, após serem aptos em testes de conhecimentos e costumes (SOEIRO; AVELINE, 1982).

Em comunidades remotas, a avaliação de seus integrantes consistia em cerimônias de torturas, cujo intuito era averiguar o condicionamento físico de seus membros e declarar uma crença social (ARANHA, 1996). Sociedades chinesas e gregas, também aplicavam métodos avaliativos em busca de indivíduos com perfis adequados para assumir determinados cargos relacionados ao trabalho.

Durante muito tempo o acesso a educação era privilégio de poucos, em especial das classes dominantes, nesse contexto a escola surge como forma de atender os desfavorecidos, haja vista que apenas a elite desfrutava deste benefício (ARANHA, 1996).

A instituição escola no decorrer da história sofreu consideráveis modificações. Sendo inicialmente idealizada para suprir as necessidades da burguesia, a qual no final do século XVIII se solidificou (SACRISTAN, 2007). A partir de meados do século XIX, verifica-se acentuadas discrepâncias educacionais entre a classe rica e os desfavorecidos durante o segundo momento da Revolução Industrial, onde ocorre o ápice tecnológico do século, com o intuito de cessar com enormes privilégios da nobreza e clero feudal, essas classes formaram uma aliança para se tornar mais forte (LUCKESI, 1991). Todavia, a ascensão ao poder por parte da burguesia, lhes possibilitou consolidar benefícios econômicos e sociais conquistados anteriormente.

É salutar ressaltar que o surgimento da escola, possibilitou no decorrer dos séculos novas formas de inclusão social através da educação e com isso houve a necessidade de

mensuração do conhecimento através de técnicas avaliativas (estatísticas) com aplicação de testes, os quais dariam um norte em relação ao ganho de conhecimento do indivíduo ao longo do tempo. Logo, é plausível destacar que assim como o estudante acumula conhecimento no decorrer do tempo, o mesmo aprende diversas formas de tentar demonstrar falsa aprendizagem, tentando burlar ou fraudar os meios fiscalizatórios para conseguir vantagem pessoal.

Exames avaliativos na educação de hoje são de suma importância, pois, através dos resultados obtidos, é possível ter uma visão mais contundente em relação à qualidade da educação oferecida pelas instituições de ensino. Estes se destacam como ferramenta primordial para educadores e consequentemente subsidiando decisões políticas. Para o corpo pedagógico, estes resultados lhe dão um direcionamento na aplicação de novas metodologias para o melhoramento do ensino e avaliar programas e escolas. É notório que a avaliação educacional é de grande relevância para a obtenção de excelentes parâmetros de forma a corroborar planejamento e execução de projetos no que tange alcançar a excelência educacional (SOTARIDONA, 2003).

As práticas de atos fraudulentos em testes padronizados prejudicam a consistência da avaliação educacional, discentes burladores trazem pra si grandes prejuízos que refletiriam na qualidade de sua formação, contribuindo assim para uma possível decadência na trajetória de suas carreiras. Estudantes íntegros que não utilizam meios burladores são extremamente prejudicados, principalmente quando os testes avaliativos são utilizados para fins de seleção para bolsas de estudo, emprego e etc. (ALLEN, 2012).

A utilização de testes padronizados na detecção de fraudes possibilita identificar carências acadêmicas do discente, projetos adequados ao nível de realização do aluno e a mensuração do desempenho escolar. Segundo Angoff (1974) (apud ALLEN 2012, p. 01), “Duas formas de engano por estudantes têm sido objeto da maior parte da pesquisa sobre métodos de detecção para observância de trapaças: cópia e representação”.

Segundo Sotaridona (2003, p. 01):

Há uma variedade de maneiras de fazer trapaça o problema pode ser dirigido, por exemplo, através do desenvolvimento de uma atitude correta entre os estudantes, através da prevenção, ou através da detecção. Embora a prevenção e desenvolver a atitude certa entre os alunos podem reduzir consideravelmente trapaças, detecção de fraude é importante para justificar a tomada de teste justo.

Métodos estatísticos ou observacionais podem ser empregados em testes de múltipla escolha para detectar cópia de resposta. No observacional, usa-se um indivíduo na figura de fiscal durante um exame para detectar fraude por algum artifício usado pelo examinando, já os métodos estatísticos averiguam possíveis fraudes através da modelagem das probabilidades de resposta dos examinandos, considerando dois requisitos:

- a) Não fazer trapaça;
- b) Averiguar padrões de respostas semelhantes entre os examinandos que são inverossímeis sob o modelo.

Semelhanças ou idênticas respostas aos itens por parte dos examinandos submetidos a exames de múltipla escolha se dá através de métodos burlatórios. Nesta vertente, usa-se na maioria das vezes para avaliar par de alunos a constância de respostas incorretas idênticas de forma a subsidiar medidas de similaridades de respostas (ANGOFF, 1974; BELLEZZA e BELLEZZA, 1989; VAN DER LINDEN e SOTARIDONA, 2004).

Ressalta-se, que o quantitativo de acertos por aluno em um teste influencia consideravelmente a inconstância de verossímeis respostas incorretas idênticas. Por conseguinte, é plausível salientar que medidas de tendência central e de dispersão, possibilitam explorar outras peculiaridades da distribuição como a frequência de respostas incorretas idênticas, pois são averiguados por contagem bruta. Logo, deve-se adotar os escores brutos, direta ou indiretamente para auxiliar as medidas de respostas com base na similaridade de respostas incorretas idênticas (ALLEN, 2012).

As coincidências de respostas semelhantes entre examinandos, vão além dos resultados brutos. Pois, se indivíduos recebem mesmos treinamentos com mesmo professor durante sua formação educacional, é plausível que ocorra correspondência de respostas incorretas aos mesmos itens, assim como escolha pelas as mesmas opções de respostas incorretas. Logo, se indivíduos apresentam laços de experiências compartilhadas, tenderiam a apresentar níveis médios elevados de respostas semelhantes (ALLEN, 2012).

Nesta óptica, surgiram novas propostas para aperfeiçoamento dos métodos avaliativos com o intuito de robustecer os anteriormente empregados, para tanto técnicas destinadas a identificação de fraudes em testes vem se tornando uma tendência nos dias atuais, como forma de avaliação mais rica possibilitando relatar esquemas e evitar malefícios a outrem.

4.2 Modelos para Detecção de Fraudes em Testes

Os primeiros métodos elaborados para a detecção de fraudes em teste surgiram em 1927, sendo Bird (1927-1929), Crawford (1930), Dickenson (1945), Anikeef (1954) e Saupe (1960) os pioneiros nesta empreitada. Porém, apesar de esforços para a formulação de técnicas sofisticadas, não houve possibilidade de empregá-las na época, devido a estruturação computacional exigida e a carência de softwares estatísticos adequados, inviabilizando a efetivação das metodologias propostas por estes pesquisadores (KHALID et. al, 2011).

Após mais de uma década Angoff (1974), elaborou uma gama de índices estatísticos, inserindo em seu estudo a mesma metodologia adotada por Saupe (1960), o qual considerou diversas variáveis com o intuito de detectar ocorrências de cópia, logo observou que o seu método é bem mais eficiente através da inserção de variáveis que envolvam a contagem de respostas certas e erradas. A partir de então, outros pesquisadores dedicaram-se na formulação de outros índices, como (Frary et al., 1977; Schumacher, 1980; Cody, 1985; Roberts, 1987; Hanson et al., 1987; Belleza e Belleza, 1989; Scrutiny!, 1993; Harpp et al., 1996; Holland, 1996; Wollack, 1997; Kadane, 1999; Ercole, 2002; Sotaridona e Meijer, 2002; Van Der Linden e Sotaridona (2004); Sotaridona et al., 2006; Vander Linden e Sotaridona (2006); Belov e Armstrong (2010)).

As metodologias propostas para subsidiar a detecção fraudes, são estruturalmente fundamentadas considerando duas teorias: a Teoria Clássica dos Testes (TCT) e Teoria da Resposta ao Item (TRI). Estas ferramentas estatísticas averiguam entre examinandos a similaridades de padrão de respostas, comparando-as com a frequência esperada de similaridades de respostas. A TCT apresenta como característica a dependência entre os itens que compõem o teste e o grau de proficiência do aluno. Assim, para cada examinando em um grupo compara-se seu padrão de resposta com os padrões de resposta dos demais indivíduos submetidos ao teste, incluindo pessoas que não praticaram cópia. Desta forma, obtém-se para um par de examinando um número esperado de similaridade de respostas através de estimativas tendenciosas.

A TRI se destaca como ferramenta estatística que pode ser utilizada para a detecção de fraude, logo a mesma apresenta em sua metodologia uma gama de modelos que podem ser utilizados conforme o tipo de teste proposto. O modelo de resposta nominal é sugerido para análise em teste de múltipla escolha. A fundamentação teórica da TRI propõe que para uma

dada habilidade de um examinando a propensão de responder corretamente um item é independente de outros examinandos submetidos ao teste, logo observa-se a grande diferença entre a TRI e a TCT, ou seja, a TRI não é dependente dos itens que compõem o teste. Os modelos propostos pela TRI consideram os parâmetros dos itens: níveis de dificuldades e discriminação das alternativas ou testes. Os modelos elaborados permitem a comparação de padrões de respostas considerando proficiências semelhantes entre examinandos.

- Taxa de Erro Tipo I: ocorre quando um índice caracteriza erroneamente um fato como fraude quando não é;
- Taxa de Erro Tipo II: ocorre quando um índice não consegue detectar um caso de fraude que aconteceu.

4.3 Alguns Modelos Baseados na Teoria Clássica dos Testes (TCT)

4.3.1 Índice de Crawford.

Crawford em 1930 formulou seu índice considerando para o par de alunos o percentual de respostas incorretas similares. Tal método é concebido pela diferenças entre proporções, conforme é mostrado a seguir:

$$Z_c = \frac{P - \bar{P}}{S_{P-\bar{P}}}, \quad (4.1)$$

$$P = \frac{Q}{W_A + W_B}, \quad (4.2)$$

onde

- \bar{P} é a média de P referente a todos os pares de estudantes que não praticaram cópia;
- $S_{P-\bar{P}}$ erro padrão entre as duas proporções;
- Q número de itens para os quais os alunos escolhem a mesma opção incorreta;
- W_A número de respostas erradas do estudante A;
- W_B número de respostas erradas do estudante B;
- $W_A + W_B$ constitui a soma das respostas erradas dos estudantes A e B, o termo Z_c implica em uma estatística normalmente distribuída, para obtenção de seus valores usa-se a tábua normal padrão.

Compara-se o valor de Z_c (calculado) com Z_t (valores encontrados na tábua da distribuição normal padrão) a um nível de significância determinado, caso $Z_c > Z_t$, conclui-se que houve trapaça durante o teste.

4.3.2 Índice de Dickenson

O referido índice foi formulado para averiguar a probabilidade de ocorrência de respostas erradas idênticas ao acaso. Seu objetivo é possibilitar a comparação entre o número de respostas incorretas idênticas com o número esperado de respostas possíveis para cada item. Desta forma, assume-se que a probabilidade do aluno responder incorretamente cada item que compõem o teste é a mesma. Nesta óptica o índice é caracterizado conforme a equação abaixo:

$$IE = \frac{C - 1}{C^2} \quad (4.3)$$

- C é o número de alternativas do item (supondo que todos os itens tem o mesmo número de alternativa);
- IE a proporção esperada de respostas incorretas idênticas.

Conclui-se que houve fraude durante o teste, se a proporção de respostas incorretas observadas for maior que duas vezes a proporção esperada (SAUPE, 1960).

4.3.3 Índice de Anikeef

O índice utiliza em sua estrutura um modelo de probabilidade, adotando a distribuição binomial para determinar a probabilidade observada do número de respostas incorretas idênticas entre dois testes. Neste método verifica-se o quantitativo de respostas incorretas idênticas entre um par de testes usando a distribuição binomial, cuja média é N e desvio-padrão $\sqrt{N_p(1-p)}$, sendo N o número total de respostas erradas do aluno suspeito de fraudar e p é o inverso do número de opções de respostas por item.

Uma baixa probabilidade observada de correspondência no número de respostas incorretas demonstra a ocorrência de fraude (THIESSEN, 2008).

Ressalta-se que seu método não é totalmente eficaz na detecção de cópia. Apresentando utilidade apenas em situações em que o aluno pratica cópia em mais de 16% das respostas originárias de outro aluno (THIESSEN, 2008).

4.3.4 Índice de Regressão

Método desenvolvido por Saupe (1960), com o intuito de identificar possíveis suspeitos de copiar respostas, foi desenvolvido através de regressão linear. Nesta formulação considera-se tanto as respostas corretas e incorretas idênticas, originários de dois índices de detecção para o número de similaridade de respostas corretas e incorretas. Estes índices são descritos abaixo:

1. Considera-se o número de resposta **corretas idênticas** a um determinado par de alunos, através da seguinte equação:

$$R_{AB} = a_1 + b_1(R_A + R_B)W_{AB} + \epsilon \quad (4.4)$$

2. Considera-se o número de resposta **erradas idênticas** a um determinado par de alunos, através da seguinte equação:

$$Q = a_2 + b_2(R_A + R_B)W_{AB} + \epsilon \quad (4.5)$$

Onde

- R_{AB} é o número de respostas certas entre dois alunos;
- $R_A + R_B$ é a soma das respostas corretas para os alunos;
- R_A é o número de respostas corretas do aluno A;
- R_B é o número de respostas corretas do aluno B;
- W_{AB} é o quantitativo de itens respondidos incorretamente pelos alunos, sem considerar a opção de escolha;
- Q é o número de itens que ambos os alunos optaram pela mesma opção incorreta.

Processo para Detecção de Fraude

- Calcula-se o erro padrão de estimativa (EP);
- Divide-se a regressão para cada par residual pelo EP, desta forma padronizando os resíduos (Z);
- Caracteriza-se como suspeitos de cola, pares de alunos com Z positivo acima de um determinado valor crítico ou nível de significância.

4.3.5 Índices de Angoff

Desenvolvido na década de 1970, Angoff formulou oito índices considerando uma variedade de variáveis para detectar possíveis esquemas em testes com base no modelo proposto por SAUPE (1960). Angoff sugeriu os índices A, B e H como promissores na detecção de fraudes.

O processo utilizado por este método considera:

- O candidato I responde corretamente R_i itens em um teste;
- O candidato J responde corretamente R_j itens em um teste e
- R_{ij} sendo o número de respostas corretas compartilhadas entre dois alunos.

Ressalta-se que a medida R_{ij} não é significativa para mensurar a similaridade entre alunos, pois conforme aumenta o nível de habilidade do aluno o número de respostas semelhantes tende a aumentar.

Avalia-se R_{ij} em relação à R_i e R_j , através da projeção de R_{ij} pelo cálculo residual de R_{ij} após a regressão sobre $\sqrt{R_i R_j}$ e $R_i R_j$. Os resíduos são distribuídos normalmente e expressa como probabilidades. O índice B pode ser obtido de forma similar, através de cálculo residual de Q_{ij} , após a regressão em $\sqrt{W_i W_j}$ e $W_i W_j$. Já o índice H utiliza seu cálculo respostas incorretas idênticas em uma sequência extensa de itens e número de itens respondidos incorretamente.

As equações de regressão são expressas como segue abaixo:

$$R_{ij} = \sqrt{R_i R_j} + R_i + R_j + \epsilon, \quad (4.6)$$

$$Q_{ij} = \sqrt{W_i W_j} + W_i + W_j + \epsilon, \quad (4.7)$$

$$K_{ij} = \sqrt{S_i} + S_i + \epsilon, \quad (4.8)$$

onde

- Q_{ij} é o quantitativo de itens respondidos incorretamente pelos indivíduos i e j;
- $W_i W_j$ é o produto entre o quantitativo de itens respondidos incorretamente pelos indivíduos i e j;

- S_i é o número de itens respondidos incorretamente por i ou j, mais o número de itens omitidos por examinando;
- K_{ij} é o número de itens respondidos incorretamente de forma idêntica e não respondidos (omissos) por i e j.

A caracterização do ato fraudulento ocorre se pares de resíduos padronizados positivos apresentarem valores superiores a um determinado valor crítico, conclui-se então pela ocorrência de fraude no teste.

4.3.6 Estatística g

Estatística para detecção de fraude, desenvolvida por Farry et al (1993-1997) e Cody (1985), a qual considera em seu cálculo uma das estatísticas: o número de respostas idênticas ou número de respostas incorretas idênticas. Logo, assumi-se para o modelo uma determinada distribuição de probabilidade para o número de respostas idênticas ou o número de respostas incorretas idênticas considerando um par de alunos. O modelo assume que cada indivíduo apresenta uma probabilidade de responder a cada alternativa de um determinado item e considera independentes as respostas de indivíduos diferentes a todos os itens do teste. Para obtenção da pontuação total das probabilidades de resposta ao item, propuseram duas funções:

- Resposta j correta ao item i

$$\hat{P}(U_{ia} = j) = P_{ij} \frac{X_a}{\bar{X}}, \quad 0 \leq X_a \leq \bar{X} \quad (4.9)$$

$$\hat{P}(U_{ia} = j) = 1 - (1 - P_{ij}) \left[\frac{N - X_a}{N - \bar{X}} \right], \quad \bar{X} \leq X_a \leq N \quad (4.10)$$

- Resposta j incorreta ao item i

$$\hat{P}(U_{ia} = j) = P_{ij} \left[\frac{1 - \frac{P_{ic} X_a}{\bar{X}}}{1 - P_{ic}} \right], \quad 0 \leq X_a \leq \bar{X} \quad (4.11)$$

$$\hat{P}(U_{ia} = j) = P_{ij} \left[\frac{N - X_a}{N - \bar{X}} \right], \quad \bar{X} \leq X_a \leq N \quad (4.12)$$

onde

- U_{ia} variável aleatória que contém as cópias de respostas suspeitas ao item i;

- P_{ij} proporção de alunos que dão resposta j ao item i ;
- \bar{X} média observada da pontuação total no teste;
- X_a pontuação do indivíduo a suspeito de copiar as respostas no teste;
- N número total de itens no teste;
- P_{ic} proporção de alunos que responderam corretamente ao item i .

Segundo Khalid et al (2011, p. 53), Farry et al (1993-1997) utilizou a média e o desvio-padrão desta distribuição para subsidiar a padronização da estatística g , onde a média é subtraída do número observado de respostas idênticas e dividido pelo desvio-padrão e assim comparando-a a um nível de significância determinado na tábua da distribuição normal padrão, concluí-se que houve fraude se $g > Z_t$ (valor obtido na tábua da distribuição normal padrão, considerando um determinado nível de significância).

Ressalta-se que o índice proposto por Cody (1985), utiliza praticamente a mesma metodologia, porém com um diferencial adota em seu modelo apenas a similaridade de respostas erradas.

4.3.7 Índices PAR1 e PAR2

Para estes métodos foram propostos duas medidas descritivas com intuito de detectar cópias de respostas, com a utilização de uma distribuição bivariada, ambos elaborados por HANSON et al (1987).

Para o índice PAR1 as seguintes medidas são utilizadas em sua estrutura:

- A sequência mais extensa de respostas idênticas ou omissas para pares de indivíduos;
- O número de respostas incorretas idênticas.

Para o índice PAR2 as seguintes medidas são utilizadas em sua estrutura:

- O número de respostas incorretas na sequência mais extensa de respostas idênticas;
- A função estatística PJ , expressa da seguinte forma:

$$PJ = 100 \frac{JI1I2}{NITENS - (TJOINT - JI1I2)}, \quad (4.13)$$

onde

- $JI1I2$ número de respostas incorretas idênticas;

- *NITENS* número total de itens;
- *TJOINT* número total de itens respondidos de forma idêntica;

Segundo Khalid et al (2011, p. 54) se os índices PAIR1 e PAIR2, apresentarem valores elevados há fortes indícios de fraude no teste.

4.3.8 Índice de Análise de Similaridade de Erro (IASE)

Índice obtido pela distribuição binomial, proposto por Belezza e Belezza (1989), considerando as similaridades de erros através de probabilidades. O modelo sugere que para a obtenção das probabilidades, deve-se considerar par de examinandos que apresentam determinado número de respostas incorretas idênticas nos itens de teste.

Para a obtenção do índice, considerar-se P como elemento essencial para a produção de um valor de p associado a um determinado número de erros comuns, que pode ser expressa através da equação abaixo:

$$p = \frac{N!}{K!(N-k)!} P^k (1-P)^{N-K}, \quad (4.14)$$

onde

- N número de itens com mesma resposta incorreta, considerando a opção de escolha de resposta;
- K o número de itens comuns que ambos os estudantes responderam incorretamente, independentemente da opção escolhida;
- P probabilidade subjetiva determinada pelo pesquisador.

Segundo Khalid et al (2011, p. 55) se par de examinandos apresentar um valor probabilístico muito baixo, há fortes evidências de fraude no teste.

4.4 Alguns Modelos Baseados na Teoria da Resposta ao Item (TRI)

4.4.1 Índice ω

Método elaborado tendo como base um dos modelos da Teoria da Resposta ao Item (TRI), o Modelo de Resposta Nominal (MRN) proposto por BOCK (1972), tem o intuito de estimar a probabilidade de uma resposta de examinandos a uma das categorias de resposta ao item $v = [I, \dots, h, \dots, V]$, Assim, a probabilidade do respondente j com determinado nível de

habilidade θ_j responder a opção h do item i com parâmetros de discriminação e dificuldade ξ_{ih} e λ_{ih} , pode ser representada conforme a equação abaixo:

$$P_{ih}(\theta_j) = \frac{e^{(\xi_{ih} + \lambda_{ih}\theta_j)}}{\sum_{v=1}^V e^{(\xi_{iv} + \lambda_{iv}\theta_j)}}. \quad (4.15)$$

Sendo h_{cs} o número de itens respondidos de forma idêntica por s (fonte suspeita) e c (suspeito de copiar as respostas de s) e $E(h_{cs}|\theta_c, \mathbf{U}_s, \boldsymbol{\zeta})$, a esperança de h_{cs} , condicionada a habilidade θ_c , vetor de respostas ao item da fonte $\mathbf{U}_s = (U_{1s}, \dots, U_{Is})$ onde U_{is} é a resposta ao item i e ao vetor de parâmetros dos itens $\boldsymbol{\zeta} = (\zeta_1, \dots, \zeta_I)$ com $\zeta_i = (\zeta_{i1}, \dots, \zeta_{iV}, \lambda_{i1}, \dots, \lambda_{iV})$. Considerando $\sigma_{h_{cs}}$ como desvio-padrão de h_{cs} , o índice ômega pode ser obtido da seguinte forma:

$$\omega = \frac{h_{cs} - E(h_{cs}|\theta_c, \mathbf{U}_s, \boldsymbol{\zeta})}{\sigma_{h_{cs}}}, \quad (4.16)$$

onde a esperança $E(h_{cs}|\theta_c, \mathbf{U}_s, \boldsymbol{\zeta})$ pode ser obtida através da somatória das probabilidades das respostas do indivíduo j , como segue abaixo:

$$E(h_{cs}|\theta_c, \mathbf{U}_s, \boldsymbol{\zeta}) = \sum_{i=1}^I P(u_{ic} = u_{is}|\theta_c, \mathbf{U}_s, \boldsymbol{\zeta}). \quad (4.17)$$

Para a determinação do desvio-padrão $\sigma_{h_{cs}}$, utiliza-se a equação abaixo:

$$\sigma_{h_{cs}} = \sqrt{\sum_{i=1}^I P(u_{ic} = u_{is}|\theta_c, \mathbf{U}_s, \boldsymbol{\zeta}) \left(1 - \sum_{i=1}^I P(u_{ic} = u_{is}|\theta_c, \mathbf{U}_s, \boldsymbol{\zeta})\right)} \quad (4.18)$$

Segundo Sotaridona (2003, p. 30), altos valores de ômega ou $p - \text{valor} < \alpha$ (nível de significância), indica evidência de fraude.

4.4.2 Teste Binomial Generalizado

Baseado no modelo dicotômico da TRI, considerando o número de respostas idênticas entre dois vetores de respostas, calcula-se a distribuição de probabilidade exata. Seja P_i a probabilidade de correspondência no item i , o qual pode ser expresso conforme abaixo:

$$P_i = [P(x_{ic} = 1|\hat{\theta}_c, \hat{\zeta}_i) * P(x_{is} = 1|\hat{\theta}_s, \hat{\zeta}_i)] + [P(x_{ic} = 0|\hat{\theta}_c, \hat{\zeta}_i) * P(x_{is} = 0|\hat{\theta}_s, \hat{\zeta}_i)], \quad (4.19)$$

onde

- x_{ic} e x_{is} são realizações de respostas observadas para o examinando (copiador) suspeito e o examinando fonte (suspeita) do item I , respectivamente;
- $\hat{\theta}_c$ e $\hat{\theta}_s$ são as estimativa de habilidade do examinando (copiador) suspeito e o examinando fonte (suspeita), respectivamente;
- $\hat{\zeta}_i$ é a estimativa dos parâmetros do modelo para o item i .

Desta forma, a probabilidades de se observar exatamente m correspondência em n itens entre dois vetores de respostas, será igual

$$f_n(m) = \sum \prod_{i=1}^n P_i^t Q_i^{1-t}, \quad (4.20)$$

sendo

- $Q_i = 1 - P_i$;
- $t = 1$ se examinandos (origem e copiadora) apresentarem respostas idênticas ao item i e zero caso contrário;
- O somatório corresponde a todas as combinações possíveis de m correspondência em n itens.

Assim, se caso haja duas correspondência de respostas em três itens, teremos $f_3(2) = Q_1P_2P_3 + P_1Q_2P_3 + P_1P_2Q_3$. Por fim, a probabilidade de se observar O_{cs} ou mais correspondências em n itens é igual

$$\sum_{j=O_{cs}}^n f_n(j) \quad (4.21)$$

com O_{cs} o número observado de respostas idênticas entre dois vetores de respostas

Para caracterização de fraude, compara-se P_i com um nível de significância α , caso $P_i < \alpha$, considera-se como ocorrência de fraude.

4.4.3 Índice K

Esta estatística avalia a similaridade de respostas incorretas em um teste de múltipla escolha entre dois examinandos considerando um como fonte, que chamaremos de (s) e o outro como a copiadora (c). Neste índice suspeita-se do indivíduo (c) ter copiado as respostas a partir da (s). Esta estatística considera apenas as respostas incorretas entre os examinandos (HOLLAND, 1996).

Notação

Consideremos:

- s examinando definido como o fonte;
- c examinando definido como suspeito de copiar as respostas de s ;
- $j = (1, \dots, J)$ representa examinados;
- $i = (1, \dots, I)$ representa itens;
- $v = (1, \dots, V)$ representa as categorias resposta ao item;
- w_j o número de respostas erradas do examinando j ;
- M com a realização m denotando o número de respostas similares erradas entre examinando j e s ;
- $r = (1, \dots, c', \dots, R)$ representa subgrupo de examinandos, onde cada grupo tem um número distinto de respostas erradas e c' é o grupo onde examinando c pertence;
- $j' = (1, \dots, n_r)$ representa um examinando no subgrupo r , onde cada subgrupo tem pelo menos um examinando;
- $\sum_{r=1}^R n_r = J - 1$; $M_r = (M_{r1}, \dots, M_{rj'}, M_{rn_r})$ representam um vetor do número de correspondentes respostas erradas em um determinado subgrupo r ;
- $M_{c'} = (M_{c'1}, \dots, M_{c'n_{c'}})$ representa um vetor do número de correspondentes respostas erradas de examinandos $n_{c'}$ no subgrupo c' onde c' consiste no subgrupo de examinados que obtiveram a mesma pontuação em relação ao mesmo número de respostas incorretas da copiadora;
- $Q_r = \frac{w_r}{I}$, indica a porcentagem de respostas erradas do subgrupo r onde I é o número total de itens no teste.

4.4.3.1 Distribuição Empírica para o Índice K

Para a obtenção desta distribuição, podem-se considerar dados empíricos através de J indivíduos terem respondido I itens. Para isso, elaboram-se dois grupos:

- Examinandos com o mesmo número de pontuação incorreta assim como c (subgrupo c') (G_1);
- Para cada um destes examinandos no subgrupo c' é determinado o número de itens que coincidem com as respostas incorretas da fonte (G_2) (este é o vetor $M_{c'}$).

Consideremos ainda:

- A distribuição de $M_{c'}$ corresponde a distribuição empírica de similaridade de respostas;
- Para o indivíduo c , denotemos $m_{c'c}$ especificamente como o número de respostas erradas correspondentes entre c e s ;
- Representemos a variável aleatória M_{rj} por M , se não é necessário para identificar os membros do grupo de j .

Considerando examinandos que possuem mesmo número de pontuação incorreta como do indivíduo copiador (c) e o número de correspondência de pontuação incorreta no item com o indivíduo fonte (s) e que seja pelo menos tão grande quanto $m_{c'c}$. Para $j' = (1, \dots, n_{c'})$, adotemos $I_{c'j'}$ como função indicadora com valor 1 para $m_{c'j'} \geq m_{c'c}$ e 0, caso contrário, assim o referido índice pode ser expresso conforme equação abaixo:

$$K = \frac{\sum_{j'=1}^{n_{c'}} I_{c'j'}}{n_{c'}}, \quad (4.22)$$

Ressalta-se que o nível de habilidade dos examinandos irá influenciar diretamente na contagem de correspondência de respostas incorretas, ou seja, o número de respostas incorretas similares é necessariamente pequena quando s ou c ou ambos têm muitas pontuações corretas (grande habilidade) e vice-versa. Devido esta situação, obtêm-se o índice K , condicionalmente sobre o número de contagens incorretas da copiadora suspeita. Com isso o índice é obtido através apenas de uma pequena quantidade de examinandos (subgrupo c'). É salutar salientar que a precisão do índice será influenciada pela quantidade de indivíduos no subgrupo c' .

Para pequenas amostras $J \leq 100$, indica-se utilizar uma aproximação teórica para a distribuição de correspondência empírica.

4.4.3.2 Aproximação Teórica para Obtenção do Índice K

É de salutar relevância explicar sobre o erro do Tipo I (α) antes da obtenção do índice K , o qual se reporta sobre a probabilidade de mal classificação de um indivíduo como copiadora quando na verdade ele não é. Neste sentido a grande preocupação é obter uma estatística não liberal, ou seja, uma estatística para a qual as taxas de erros Tipo I empírica e nominal sejam o máximo possível equivalentes, pois uma má classificação de indivíduos íntegros como copiadora pode ser muito grave a nível individual.

Segundo Seaman et al (1991) (apud SOTARIDONA, 2003, p. 09)

Índices de cópia que não conseguem manter a taxa de erro Tipo I nominal deve ser considerado inaceitável. Por outro lado, um índice de cópia não deve ser excessivamente conservador; Caso contrário, seu poder de detectar verdadeiros examinandos como copiadores será muito baixo.

A pequena quantidade de valores disponibilizados pela variável aleatória M , demonstra a pouca relevância de utilização da distribuição discreta empírica para pequenas amostras. Ocasionalmente em muitos casos a impossibilidade de se obter um erro Tipo I pré-definido, por exemplo, 0.01 (AGRESTI, 1996).

A possibilidade de utilizar uma distribuição aproximada para M , através de uma distribuição binomial, proposta por Holland (1996), ou seja, considera $M \sim Bin(w_s, p)$ sendo w_s o número de respostas erradas conhecidas de s , com p desconhecido. Então, com o objetivo de obter p , este pesquisador propôs duas alternativas, para a primeira possibilidade adota-se a mesma média para ambas as distribuições de M , binomial e empírica e considerando $\bar{m}_{c'}$ como a média da distribuição empírica, obtém-se p da seguinte maneira:

$$\bar{m}_{c'} = \frac{\sum_{j'=1}^{n_{c'}} m_{c'j'}}{n_{c'}} \quad (4.23)$$

Assim, tem-se uma estimativa para p através de $p_{c'}^*$, como segue abaixo

$$p_{c'}^* = \frac{\bar{m}_{c'}}{w_s} \quad (4.24)$$

Denotando o índice K por K^* e adotando $p_{c'}^*$, para sua obtenção teremos:

$$K^* = P(M \geq m_{c'c}) = \sum_{g=m_{c'c}}^{w_s} \binom{w_s}{g} (p_{c'}^*)^g (1 - p_{c'}^*)^{w_s - g}, \quad (4.25)$$

Com a utilização de $p_{c'}^*$ na distribuição binomial para obtenção do índice K , permitiu a Holland (1996) em seu estudo obter uma estimativa "conservadora" para a distribuição empírica, porém utilizável apenas para grandes conjuntos de dados empíricos (SOTARIDONA, 2003).

4.4.3.3 Obtenção de Variantes para do Índice K

A utilização dos padrões de respostas dos examinados no subgrupo c' é de suma importância para a obtenção de $p_{c'}^*$. Porém, ressalta-se que o tamanho da amostra é de extrema relevância na a produção de uma estimativa confiável de $p_{c'}^*$. De modo a contornar o problema, Holland (1996) propôs a utilização de regressão linear através da proporção de respostas erradas (Q_r) de cada examinando em cada número incorreto da pontuação do subgrupo r como regressores, com o intuito de obter um valor aproximado de $p_{c'}^*$. Utilizando amostra representativa de dados da Educational Testing Service (ETS), maior empresa de teste dos E. U. A., Hollanda (1996) demonstrou a relação linear entre $p_{c'}^*$ e Q_r empiricamente, considerando a obtenção de p_r^* da mesma forma como proposto para $p_{c'}^*$. De forma a obter a estimativa para p_r^* , adota-se \hat{p}_r usando Q_r no cálculo de probabilidade binominal (Sotaridona, 2003). A expressão para \hat{p}_r é obtida através de função linear por parte com parâmetros a e b , caracterizando o intercepto e o declive, respectivamente, conforme condições expressas abaixo:

$$\hat{p}_r = \begin{cases} a + bQ_r & \text{se } 0 < Q_r \leq 0.3 \\ [a + 0.3b] + 0.4b[Q_r - 0.3] & \text{se } 0.3 < Q_r \leq 1 \end{cases} \quad (4.26)$$

Segundo Sotaridona (2003, p. 10),

a e b tem que ser especificado, a fim de estimar \hat{p}_r em (4.26). Holland (1996) utilizou $a = 0.085$ e para diferentes valores de b dependendo do teste particular que foi utilizado. No entanto, a partir de seu estudo não ficou claro como se obtiveram estes valores. Além disso, eles podem variar entre os diferentes testes.

Com base no exposto acima Sotaridona (2003), propôs duas variantes do índice K , os denominados K_1 e K_2 baseado na estimativa de p^* , onde utilizou regressão linear e quadrática para estimar \hat{p}_r^* através de \hat{p}_1^* e \hat{p}_2^* , sendo que para \hat{p}_2^* é obtido conforme equação $\hat{p}_2^* = E(\beta_0 + \beta_1 + \beta_2 Q_r^2 + \epsilon_r)$, com $\beta_0, \beta_1, \beta_2$ coeficientes de regressão e ϵ_r é o erro que assume distribuição normal com média zero e variância σ^2 , conforme o exposto as equações para K_1 e K_2 são:

$$K_1 = P(M \geq m_{c'c}) = \sum_{g=m_{c'c}}^{w_s} \binom{w_s}{g} (\hat{p}_1^*)^g (1 - \hat{p}_1^*)^{w_s-g}, \quad (4.27)$$

$$K_2 = P(M \geq m_{c'c}) = \sum_{g=m_{c'c}}^{w_s} \binom{w_s}{g} (\hat{p}_2^*)^g (1 - \hat{p}_2^*)^{w_s-g}, \quad (4.28)$$

Para a estimação de p através de $p_{c'}^*$, usa-se somente examinandos do subgrupo c' . Ressalta-se que \hat{p}_1^* e \hat{p}_2^* , utilizam informações pertinentes com base em R subgrupos. Logo, espera-se que \hat{p}_1^* e \hat{p}_2^* forneçam estimativas para p melhores que $p_{c'}^*$ (SOTARIDONA, 2003).

Segundo Sotaridona (2003, p. 31), valores de $(K, K_1 \text{ e } K_2) < \alpha$, conclui-se que houve ocorrência de fraude no teste.

4.4.4 Índice S (S_1 e S_2)

4.4.4.1 Índice S_1

Índice desenvolvido por Sotaridona e Meijer (2003), o índice S_1 foi proposto de forma similar ao índice K_2 , o qual se baseia na variável aleatória M . Porém, existem diferenças entre os dois índices, conforme abaixo:

- No índice K_2 , a distribuição proposta para a variável aleatória M segue uma distribuição binomial e para S_1 , M é modelado através da distribuição de Poisson;
- O parâmetro μ da distribuição de Poisson é estimado através de um modelo log-linear e para o índice K_2 o parâmetro p da distribuição binomial é estimado através de um modelo de regressão linear.

Antes de prosseguirmos adotemos as seguintes notações abaixo que serão utilizadas no decorrer das explicações:

- s representa o examinando fonte (origem da cópia);

- c representa o examinando suspeito de copiar de s ;
- w representa o número de respostas erradas;
- w_s representa o número de respostas erradas conhecidas de s ;
- w_r representa o número de respostas erradas do grupo r ;
- M com a realização m denotando o número de respostas similares erradas entre examinando j e s ;
- $M_{c'c}$ com a realização $m_{c'c}$ denotando o número de respostas similares erradas entre examinandos c e s ;
- $r = (1, \dots, c', \dots, R)$ representa subgrupo de examinandos, onde cada grupo tem um número distinto de respostas erradas e c' é o grupo onde examinando c pertence;
- $j = (1, \dots, J_r)$ apresenta o mesmo número de respostas erradas de examinados;
- J_r representa o número de examinandos com relação ao número de respostas incorretas no grupo r ;
- J_c representa o número de examinados com o mesmo número de respostas erradas quanto do examinado c ;
- rj é o índice que representará o examinando j do grupo r com relação ao número de respostas incorretas;
- U_{irj} será a resposta do examinando rj ao item i ;
- W_s será o conjunto de itens de tamanho w_s , respondido incorretamente por s ;
- A_{irj} é a variável indicadora para cada examinando rj , sendo 1 quando $U_{irj} = U_{is}$ e 0 caso contrário. Note que a resposta ao item com índice s está indicando que s não pertence a qualquer grupo com número incorreta de resposta;
- $M_{rj} = \sum_{i \in W_s} A_{irj}$ representa o número de correspondentes respostas erradas entre rj e s . Para um determinado par s e c , M_{rj} é observado para cada examinando rj . Representemos a variável aleatória M_{rj} por M , quando não é necessário identificar o examinando;
- $\hat{\mu}_{c'}$ representa o número de respostas incorretas no grupo c' onde examinando c pertence;

Considerando que a estimativa para μ seja representada por $\hat{\mu}_{c'}$, a equação para a expressão de S_1 , segue conforme abaixo

$$S_1 = \sum_{w=m_{c'c}}^{w_s} \frac{e^{\hat{\mu}_{c'}} \hat{\mu}_{c'}^w}{w!}, \quad (4.29)$$

$$S_1 = P(M_{c'c} \geq m_{c'c}) = \sum_{w=m_{c'c}}^{w_s} \frac{e^{\hat{\mu}_{c'}} \hat{\mu}_{c'}^w}{w!}, \quad (4.30)$$

sendo $P(M_{c'c} \geq m_{c'c})$ a probabilidade de w_s ser maior ou igual a $m_{c'c}$.

4.4.4.1.1 Distribuição de M

Sotaridona e Meijer (2003) elaboraram a composição da distribuição de M , considerando as respostas dadas pela fonte a um conjunto de itens fixos de teste e o número de respostas erradas da copiadora que coincide com a da fonte. Assim, devido à baixa taxa de detecção que a distribuição binomial apresentou para K e K_2 (LEWIS e THAYER, 1998; SOTARIDONA e MEIJER, 2003), adotou-se o índice S_1 , pois o mesmo utiliza a distribuição de Poisson como forma de aproximação viável para a distribuição de M . Logo, S_1 pode ser entendido como taxa de monitoramento ou número de respostas equivalentes pelas respostas incorretas dadas pela fonte.

4.4.4.1.2 Modelo para estimação de μ

Na equação proposta para S_1 , tem-se conhecimento apenas de w_s e $m_{c'c}$, já o parâmetro μ deve ser estimado. Ressalta-se, que ocorre variação na média de M para níveis de habilidades diferentes dos examinandos. Logo, baixos valores de M são mostrados quando a maioria dos examinandos apresenta elevados níveis de proficiência, pois existe baixa coincidência entre a quantidade de itens respondidos erroneamente pelo indivíduo suspeito de praticar cópia e os itens respondidos errados pela fonte suspeita. Por conseguinte, se examinandos apresentarem baixa proficiência ocorrerá o inverso, ou seja, apresentará grande correspondência de respostas incorretas entre o copiador e a fonte. Informação de grande relevância para estimação de μ , pois durante o processo de estratificação dos examinandos considera-se o número de correspondência de respostas erradas obtidas por eles (SOTARIDONA e MEIJER, 2003).

Considerando a distribuição de Poisson para M , normalmente se utiliza o modelo log-linear para modelar o log da média de M (AGRESTI, 1996, p. 73). Com a adoção deste modelo, evita-se que μ esteja linearmente relacionado com a variável preditora, que neste caso é o número de respostas erradas (SOTARIDONA e MEIJER, 2003). Com M assumindo uma distribuição binomial composta o modelo log-linear é plausível para a modelagem de M (HANSON, 1994).

Para estimar μ , consideramos o número de respostas erradas e o número médio de correspondências de pontuações incorretas para cada número de resposta incorreta do grupo r . Adotando μ_r como valor esperado da Poisson para a variável M_{rj} . O modelo log-linear apresenta a seguinte forma

$$\log(\mu_r) = \beta_0 + \beta_1 w_r, \forall r, \quad (4.31)$$

onde

- β_0 representa o intercepto, o qual é o logaritmo da média da população em todos os grupos R relativo ao número incorreto de respostas;
- β_1 representa o parâmetro de inclinação;
- w_r representa o número de respostas erradas do grupo r .

Para maiores detalhes sobre a estimação de β_0 e β_1 , sugere-se a leitura de (AGRESTI, 1996, p. 93).

Para obtenção de S_1 , necessita-se determinar a média ajustada para o grupo através do número de resposta incorreta, o qual pertence a copiadora. Logo, a média ajustada será expressa da seguinte forma

$$\hat{\mu}_c' = e^{(\hat{\beta}_0 + \hat{\beta}_1 w_c')} \quad (4.32)$$

4.4.4.1.3 Verificação do modelo proposto μ

Para averiguação da bondade do ajuste do modelo proposto em (4.31), considera-se a estatística G^2 , a qual utiliza a razão de verossimilhança em seu teor (AGRESTI, 1996, p. 89). Através da estatística G^2 , testam-se as hipóteses: H_0 : o modelo se adéqua aos dados vs H_1 : o modelo não se adéqua aos dados. Considerando $\hat{\mu}_r$ como o número médio ajustado de

respostas incorretas correspondentes entre as respostas incorretas no grupo r , obtemos G^2 conforme expressão abaixo

$$G^2 = 2 \sum_{r=1}^R \mu_r \log \left(\frac{\hat{\mu}_r}{\mu_r} \right). \quad (4.33)$$

Se há adequação do modelo aos dados, temos $\hat{\mu}_r = \mu_r$. Logo, o termo $\log \left(\frac{\hat{\mu}_r}{\mu_r} \right) = 0$ e conseqüentemente $G^2 = 0$. Assim, G^2 apresenta distribuição aproximadamente chi-quadrado com R graus de liberdade, menos o número de parâmetros do modelo. Para o modelo sugerido na equação (4.31) temos 2 parâmetros. O p-valor averigua a hipótese nula através da probabilidade localizada na cauda à direita. Ressalta-se, que valores altos de G^2 ou p-valores ínfimos, demonstram mau ajuste do modelo (AGRESTI, 1996, p. 89). Logo, se há mau ajuste do modelo aos dados, não seria viável a utilização da expressão (4.29) para averiguação de cópia de resposta.

4.4.4.2 Índice S_2

Índices desenvolvidos unicamente para fins de constatação de correspondência de cópia de respostas incorretas, tais como os índices K e K_2 , ocasionando perda de informações consideráveis que podem ser obtidas através de correspondências de cópias de respostas corretas. Logo, desconsiderando o número de correspondências de cópias de respostas corretas, supõe-se que c tem conhecimento necessário exigido para responder o item i sempre c e s respondem corretamente o item i . Todavia, a resposta correta dada pelo examinando suspeito de praticar cópia a um determinado item, pode ocorrer de forma aleatória, ou seja, devido ao acaso como: copiando ou adivinhando (SOTARIDONA e MEIJER, 2003).

Segundo Sotaridona e Meijer (2003, p. 34),

Os índices K e K_2 não são sensíveis a uma copiadora que está copiando apenas as respostas corretas da fonte. Este pode ser o caso quando c e s são amigos e s compartilha suas respostas com c nos itens onde ele ou ela tem quase certeza das respostas corretas. Outro exemplo é um exame de alto risco em que c pode subornar s para compartilhar seus itens corretamente respondidos para c .

Perante o contexto ao norte explícito Sotaridona e Meijer (2003), propuseram um novo índice que incorporasse em seus cálculos informações não somente de correspondências de respostas incorretas apresentadas pelos examinandos. Assim, de forma a contornar esta

limitação surgiu o índice o índice S_2 , o qual em seu teor adota também as correspondências de respostas corretas. Logo, adotando o valor 1 para indícios de cópia de resposta quando c e s escolhem a mesma opção errada para um item, e 0 se elas são ambas corretas ou sua resposta a um item não coincidem, ou seja, S_2 assumirá valor 1 se c e s optarem pela mesma opção errada para um item, δ caso a fonte e a copiadora estiverem corretas, e 0 caso contrário. Por conseguinte, os pesquisadores nomearam δ como a variável que demonstraria o número de cópia de resposta correta para um item para um determinado par (fonte e copiadora).

Adotando i^* , o qual representa a resposta correta dada pelo indivíduo s a um determinado item e U_{i^*rj} a resposta do examinando rj ao item i^* . Considerando, δ_{i^*rj} como a estimativa que contém informações de presença de cópia sobre o item i^* pelo examinando rj . Assim, δ_{i^*rj} será expresso conforme abaixo

$$0 \leq \delta_{i^*rj} \leq 1,$$

onde $\delta_{i^*rj} = 0$ caso rj sabe a resposta correta ao item i^* e $\delta_{i^*rj} = 1$ se rj não tem conhecimento sobre a resposta correta ao item i^* . Nesta situação há dificuldade de mensurar a proficiência de rj sobre o item i^* , logo há necessidade de se obter esta probabilidade. Todavia, segundo Sotaridona e Meijer (2003) (apud CAMACHO, 2011, p. 77) o nível de proficiência dos examinando pode ser obtido da seguinte forma: “a partir do número de pessoas que respondem corretamente o item e tem o mesmo número de respostas erradas”. Esta abordagem é deficiente, pois é nítido a dependência da estimativa em relação a população submetida ao teste, logo poderá ocorrer duas situações:

- Para examinandos com alta proficiência a estimativa produzida seria baixa;
- Para examinandos com baixa proficiência a estimativa seria alta.

Assim, de forma a contornar esta situação Sotaridona e Meijer (2003), acharam mais viável condicionar o nível de habilidade da copiadora suspeita. Para efeito de nomenclatura adotou-se j como um examinando pertencente ao certo número no grupo com respostas corretas.

Seja P_{i^*rj} a probabilidade de rj receber resposta a correta ao item i^* e A_{i^*rj} é a variável indicadora que possui valor 1 quando $U_{i^*rj} = u_{i^*s}$ e 0 caso contrário. Considera-se P_{i^*rj} como a probabilidade condicionada e não ser entendida como uma probabilidade

conjunta em que s e rj apresentam mesma resposta ao item i^* . Através de U_{i^*s} se expressa esta probabilidade da seguinte forma

$$P_{i^*rj} = Pr(U_{i^*rj} = u_{i^*s} | U_{i^*s}), \quad (4.34)$$

Logo a estimativa de probabilidade para P_{i^*rj} será

$$\hat{P}_{i^*rj} = \frac{\sum_{j=1}^{J_r} A_{i^*rj}}{J_r}. \quad (4.35)$$

Adotando δ_{i^*rj} como a estimativa transformada de P_{i^*rj} . Logo a função de transformação $f(P_{i^*rj})$ deve satisfazer as seguintes condições:

- $f(P_{i^*rj})$ se aproxima de 0 quando P_{i^*rj} aproxima de 1, ou seja, há baixo indícios de cópia de respostas quando P_{i^*rj} se aproxima de 1;
- $f(P_{i^*rj})$ se aproxima de 1 quando P_{i^*rj} se aproxima de 0, ou seja, é o indício de cópia de respostas quando se aproxima de 1 caso o indivíduo suspeito de copiar a resposta correta a um item mesmo com baixa probabilidade de receber a resposta correta ao item;
- Adotam-se diferentes funções peso quando temos testes com número diferente de opções de respostas. Considerando f e f' duas funções de pesos distintos e i^* e i'^* itens retirados de testes com diferentes opções de V e V' , com $V < V'$. Então, temos que ele afirma que $f(P_{i^*rj}) > f'(P_{i'^*rj})$ sempre que $P_{i^*rj} = P_{i'^*rj}$;

Segundo Sotaridona e Meijer (2003, p. 36)

A última condição surge da idéia de que os testes de múltipla escolha, com diferentes opções devem apresentar diferentes funções de transformação que diferem por um fator que é função do número de opções. Logo, há necessidade de uma função que represente a probabilidade de adivinhação a um item como um fator de escala.

Designando g para representar a probabilidade de se obter a resposta correta ao item i por adivinhação. Assim, considerando um teste com 5 opções, tem-se $g = 0,2$, em contra partida para um teste de com 4 opções teremos $g = 0,25$. Logo, a função abaixo satisfaz as condições acima

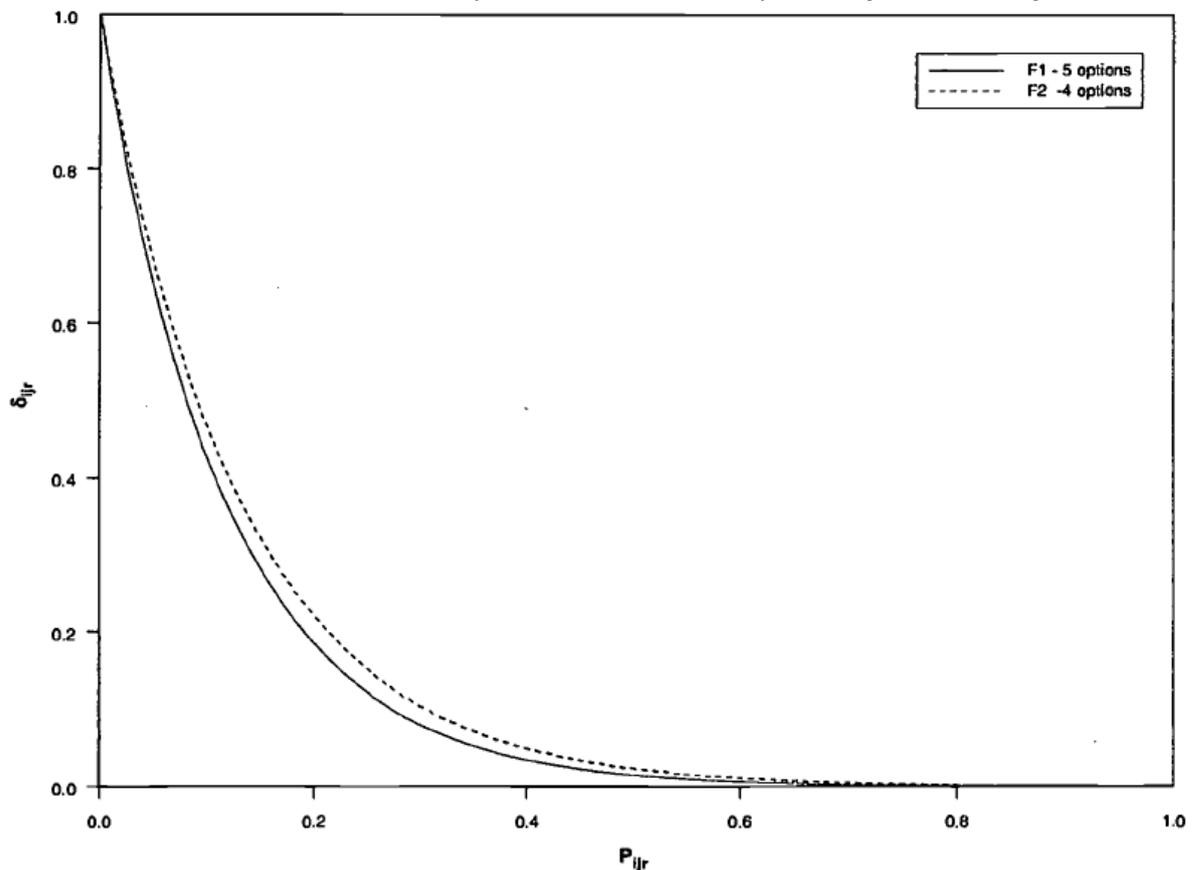
$$\delta_{i^*rj} = f(P_{i^*rj}) = (d_1 e)^{d_2 P_{i^*rj}}, \quad (4.36)$$

onde

$$d_2 = -\left(\frac{1+g}{g}\right) \text{ e } d_1 = \left(\frac{1+g}{1-g}\right)^{d_2 P_{i^*rj}}.$$

A Figura 4.1 demonstra o comportamento da equação (4.36) para $g = 0,2$ representado por (F_1) e $g = 0,25$ representado por (F_2). Nota-se que δ_{i^*rj} se aproxima de 0 com P_{i^*rj} se aproximando de 1 e P_{i^*rj} se aproxima de 1 com δ_{i^*rj} se aproximando de 0. Ressalta-se que $F_1(P_{i^*rj}) < F_2(P_{i^*rj})$ para $P_{i^*rj} \in [0, 1]$.

Figura 4.1 Comportamento de δ_{i^*rj} como função de P_{i^*rj} com $p_g = 0,20$ e $p_g = 0,25$.



Fonte: <http://onlinelibrary.wiley.com/doi/10.1111/j.1745-3984.2003.tb01096.x/epdf?r3_referer=wol&tracking_action=preview_click&show_checkout=1&purchase_referer=onlinelibrary.wiley.com&purchase_site_license=LICENSE_DENIED>.

Considerando M_{rj}^* , como a soma do número de correspondência de respostas erradas e a correspondência ponderada de respostas corretas entre os examinandos rj e s , segue que

$$M_{rj}^* = M_{rj} + \sum_{i^*} \delta_{i^*rj}, \quad (4.37)$$

onde a valoração de M_{rj}^* a cada item se dá conforme abaixo:

- 0 se as respostas de rj não corresponderem com as de s ;
- 1 se as respostas erradas de rj corresponderem com as de s ;
- δ_{i^*rj} se as respostas corretas de rj corresponderem com as de s .

Assim com base no exposto acima teremos:

- Altos valores de M_{rj}^* ocorreriam se grande parte das respostas erradas de rj corresponderem com as de s ;
- Grande coincidência de respostas corretas entre rj e s , apresentaria baixos valores de P_{i^*rj} ;
- A presença de altos valores de M_{rj}^* indica grande evidência de fraude, ou seja, cópia de respostas.

Observa-se para a não correspondência de respostas corretas entre rj e s , a segunda parcela de (4.37) seria 0, logo $M_{rj}^* = M_{rj}$. Por conseguinte, caso não haja correspondência de itens respondidos incorretamente, mas haja correspondência de itens respondidos corretamente, teremos $M_{rj} = 0$ e $M_{rj}^* = \sum_i \delta_{i^*rj}$. Desta forma temos que M_{rj} para a detecção de cópia de resposta incorreta e M_{rj}^* é sensível tanto para a detecção de cópias respostas corretas e incorretas.

Ressalta-se, que a variável aleatória M_{rj}^* possui valores reais positivos, onde M_{rj}^* é arredondado para o inteiro mais próximo. Mesmo a presença de pequenos erros neste procedimento, espera-se que ocorram baixas interferências na eficácia do valor final obtido. Assim como para M_{rj} adotou a distribuição de Poisson e o modelo log-linear para estimar a média o mesmo procedimento seguiram para M_{rj}^* (SOTARIDONA e MEIJER, 2003). Logo, exploraram empiricamente a adequação da distribuição de Poisson para o modelo M_{rj}^* utilizando a estatística G^2 .

Desta forma segue que o índice S_2 é definido como

$$S_2 = \sum_{w=m_{c'c}^*}^I \frac{e^{\hat{\mu}_{c'}} \hat{\mu}_{c'}^w}{w!}, \quad (4.38)$$

$$S_2 = P(M_{c'c}^* \geq m_{c'c}^*) = \sum_{w=m_{c'c}^*}^I \frac{e^{\hat{\mu}_{c'}} \hat{\mu}_{c'}^w}{w!}, \quad (4.39)$$

onde $M_{c'c}^*$ com realização $m_{c'c}^*$ é a soma do número de correspondências incorretas e a correspondência ponderada das respostas corretas entre c e s .

4.4.4.3 Limitações do Índice S

Segundo Zerpa e Barneveld (2013, p. 01), existem limitações deste índice, tais como:

Uma das limitações do índice S_1 , no entanto, é que ele não consegue detectar cópia de resposta quando as respostas dos examinandos são baseadas em similaridades de resposta correta. Em ordem para superar essa limitação, Sotaridona e Meijer (2003) propuseram o índice S_2 . O índice S_2 é uma melhoria na taxa de detecção de cópia de resposta sobre o S_1 porque incorpora informações de respostas corretas e incorretas entre dois examinandos, a fonte e o copiadora.

O problema com os índices estatísticos S_1 e S_2 (Sotaridona e Meijer, 2003) é que eles são dependentes do tamanho da população. Por exemplo, se um par de examinandos é suspeito de cópia de resposta, pode ser possível que estes examinandos podem obter um valor diferente para as estatísticas desta cópia se o seu par de vetores de resposta é incluído em uma população diferente. Além disso, quando a amostra é pequena os parâmetros não podem ser estimados com segurança.

Ressalta-se, que a equação para o índice S_2 é similar a S_1 , sendo que S_2 além de considerar as respostas cópia de incorretas também incorporam informações sobre a cópia que está contida nas correspondentes respostas corretas.

Para a constatação de fraude compara-se S_1 e S_2 com um nível de significância α , caso estes sejam inferiores a α há evidências de fraude.

CAPÍTULO V – IMPLEMENTAÇÃO COMPUTACIONAL.

5.1 Introdução

O software R é um programa livre que pode ser utilizados para diversos fins, por exemplo, oferecendo suporte para a utilização de uma gama de ferramentas estatísticas, cálculos, geração de diversos gráficos e etc. Software projetado para utilização em diversas plataformas como: UNIX, Windows e MacOS.

5.1.1 Pacote Estatístico CopyDetect

Dentre os mais variados pacotes estatísticos incorporados ao R, temos o CopyDetect, o qual serve de grande suporte para investigação de casos de suspeita cópia (cola/fraude) em testes avaliativos de múltipla escolha, baseado em índices estatísticos provenientes da Teoria Clássica dos Testes (TCT) ou da Teoria da Resposta ao Item (TRI). Logo, dentre os mais de vinte índices propostos na literatura para a detecção de cópia de respostas, este apresenta em seu teor índices estatísticos considerados mais eficazes e confiáveis com base nos estudos de simulação (Zopluoglu, 2016), para efeito de informação o pacote disponibiliza também o modelo logístico proposto na TRI para simulações (ver Andrade et al. (2000), para maiores detalhes).

Segundo Zopluoglu (2016), os índices estatísticos fornecidos pelo CopyDetect são $\hat{\Omega}$ (Wollack, 1996), Teste Binomial Generalizado (GBT) (Van der Linden e Sotaridona, 2006), o índice K (Holanda, 1996) e seus variantes K_1 e K_2 (Sotaridona e Meijer, 2002) e S_1 e S_2 (Sotaridona e Meijer, 2003).

O CopyDetect1 adota respostas do tipo dicotômica em seu teor, considerando:

- Resposta (0,0) representa a combinação entre dois vetores de resposta e demonstra resposta incorreta idêntica;
- Resposta (1,1) representa a combinação entre dois vetores de resposta e demonstra resposta correta idêntica;
- Resposta (NA, NA) representa a combinação entre dois vetores de resposta e demonstra resposta incorreta idêntica;

- Tipos de respostas como: (0, 1), (1, 0), (0, NA), (1, NA) entre dois vetores de resposta não são considerados como respostas idênticas. Considera-se as respostas (NA) como valor faltante, o qual é contada como uma resposta incorreta na composição para o cálculo de número de pontuação incorreta/número correta ou estimar os parâmetros de habilidade da TRI.

Segundo Zopluoglu (2013), “para os índices GBT e (ω), CopyDetect usa modelos da TRI dicotômicos para estimar a probabilidade de uma resposta correta dado os parâmetros de habilidade e de item, $P(x_i = 1|\hat{\theta}, \hat{\xi}_i)$.”

Ressalta-se, que para a geração dos índices o pacote necessita de uma amostra mínima de 100 examinandos.

O referido algoritimo utilizado para simulação de dados, encontra-se no apêndice A para efeito de consulta.

Emprego do CopyDetect1(data, item.par=NULL,pair)

Definição dos argumentos:

- **data**

Neste argumente é gerado uma composição de dados através de N linhas e n colunas, indicando o números de sujeitos e itens, respectivamente. Logo, há exigência de que as respostas aos itens sejam de forma dicotômica, onde 0 representa resposta errada e 1 resposta certa. Todas as variáveis (colunas) deve ser "numérico". Valores faltantes são possíveis de inserir através da denominação (NA).

- **item.par**

Este apresentará uma matriz de dados composta por n linhas e três colunas, com n representando o número de itens e a primeira, segunda e terceira colunas representando os parâmetros de dificuldade do item, discriminação do item, e o parâmetro adivinhação do item ou acerto ao acaso, respectivamente. Caso os parâmetros dos itens não sejam inseridos pelo elaborador, o pacote irtoys e mecanismo ltm, são automaticamente utilizados pelo pacote com o intuito de estimar os parâmetros dos itens através do modelo logístico de 2 parâmetros da TRI. Ressalta-se, que a disposição das linhas da matriz que contém os parâmetros dos itens tem de estar na mesma ordem que as colunas nos dados de resposta.

- **pair**

Este argumento produz um vetor de comprimento 2 para localizar os números de linhas para o par de examinados suspeitos, onde o primeiro e segundo elementos representam o número da linha do examinado copiadora (bulardo/suspeito) e o número de linha do examinado fonte (suspeito).

5.1.2 Etapas excultadas pelo Algoritmo Implementado no Pacote Estatístico CopyDetect (CopyDetect1) para a Detecção de Fraude

Adotou-se para efeito de exemplificação o referido algoritmo elaborado por Zopluoglu (2016) com o intuito de demonstrar a detecção de possíveis esquemas fraudulentos.

Para a simulação dos dados, considerou-se os seguintes procedimentos através de duas etapas:

1ª Etapa:

1. Simulação composta por 20 itens;
2. Adotou-se amostra de tamanho 250;
3. Utilizou-se o modelo logístico de 2 parâmetros;
4. Adotou-se apenas os índices elaborados com base na Teoria da Resposta ao Item:
 - $\hat{\Omega}$;
 - Teste Binomial Generalizado;
 - K e seus variantes (K_1 e K_2);
 - S_1 e S_2
5. Considerou-se inicialmente os indivíduos 30 (copiador) e 70 (fonte) como suspeitos, para efeitos de constatação de fraude;
6. Conforme item 2, procurou-se buscar 100 pares de indivíduos aleatórios para formação dos grupos:

- $r_1 = (1, \dots, c', \dots, R)$ composto de subgrupo de examinandos, onde cada grupo tem um número distinto de respostas erradas e c' é o grupo onde examinando c pertence;
 - $r_2 = (1, \dots, c', \dots, R)$ composto de subgrupo de examinandos, onde cada grupo tem um número distinto de respostas corretas e c' é o grupo onde examinando c pertence. Sendo que este é incluído para obtenção de S_2 , pois o mesmo adota em seu teor também as respostas corretas;
 - Posteriormente, obtém-se as probabilidades dos índices.
7. Com o intuito de verificar inicialmente se o indivíduo 30 haveria realmente praticado cópia de respostas do indivíduo 70, comparou-se os referidos índices propostos com diversos níveis de significância de 0,1%, 1% e 5%;
8. Em seguida é explanado cada índice individualmente, de forma a verificar a constatação de fraude.

2ª Etapa

9. Posteriormente a primeira etapa, adotou-se os procedimentos abaixo para obtenção desses índices para 5 pares de cópia de resposta aleatórias com o intuito de constatar a hipótese de fraude por outros indivíduos:
- Inicialmente, escolheu-se aleatoriamente um examinando suspeito de cópia;
 - Em seguida, escolhe-se um examinando fonte suspeito de ceder à cópia;
 - Finalmente seleciona-se aleatoriamente 10 itens (50% cópia) desta forma, substitui-se o vetor de resposta do examinando fonte no vetor de resposta do examinando copiador, adotando a idéia que o examinando copiador olha para o teste do examinando fonte e copia 10 dos 20 itens do teste;
 - Posteriormente, calcula-se os índices para os 5 pares selecionados aleatoriamente considerando a cópia simulada para os vetores de resposta;
 - Ressalta-se que os índices com indexador 1, foram gerados a partir das respostas simuladas originais, ou seja, sem a nova recodificação anteriormente explanada e

para os índices com indexador 2 foram adotadas as novas respostas com base na situação anteriormente citadas;

- A partir destes novos procedimentos pretende-se identificar em torno de 5% dos 100 indivíduos selecionados ao acaso como possíveis fraudadores.

5.1.3 Resultados

Durante o decorrer das análises nos reportaremos em um primeiro momento a suspeita que o examinando 30 tenha copiado as respostas do examinando fonte 70, considerando as seguintes hipóteses abaixo como forma de constatação ou não de fraude:

H_0 : Não há indícios de fraude vs H_1 : Há indícios de fraude

Tabela 5.1 - Parâmetros dos itens estimados pelo modelo logístico de 2 parâmetros.

Itens	Parâmetros		
	a_i	b_i	c_i
V1	0,8286654	-0,1793232	0
V2	0,9594891	-1,7927352	0
V3	3,0309051	0,5923723	0
V4	1,8326076	-0,2645729	0
V5	0,4309311	0,7815555	0
V6	0,4588047	-1,1984209	0
V7	0,7016700	1,3658200	0
V8	1,2802693	1,2372012	0
V9	2,1149210	-0,4933065	0
V10	0,4028021	-1,7098655	0
V11	0,6100396	1,5363445	0
V12	0,7767587	1,2032576	0
V13	1,9782962	-1,3736947	0
V14	0,5662357	-0,6134300	0
V15	0,5549593	2,2561695	0
V16	1,1237916	-1,2742621	0
V17	1,8362429	0,9157626	0
V18	1,0029354	-0,6325426	0
V19	0,5272478	-0,3574449	0
V20	1,6130340	-0,3401262	0

Fonte: CopyDetect – dados simulados.

Nas Tabelas 5.2 e 5.4, observa-se que entre os examinandos 30 e 70, ocorreu um total de 11 respostas idênticas relativas aos vetores: 1 e 2, 4 a 6, 12 a 16 e 20, das quais 3 foram

respostas incorretas idênticas referentes aos vetores: 5, 12 e 15 e 8 respostas corretas idênticas correspondentes aos vetores: 1, 2, 4, 6, 13, 14, 16 e 20. Logo, os examinandos 30 e 70 apresentaram um total de 8 e 17 respostas corretas e consequentemente 12 e 3 respostas erradas, respectivamente. Demonstrando níveis bem diferentes de proficiência, pois os valores estimados para suas habilidades foram -0,3753 e 2,5832, respectivamente, como explícito na Tabela 5.3. Assim, observa-se que, em tese, a propensão do copiador acertar o item proposto sem apoio da fonte é pequena.

Tabela 5.2 - Respostas dos examinandos 30 (suspeito de cópia) e 70 (fonte) em relação aos itens.

Itens	Examinandos	
	30 (suspeito de cópia)	70 (fonte)
V1	1	1
V2	1	1
V3	0	1
V4	1	1
V5	0	0
V6	1	1
V7	0	1
V8	0	1
V9	0	1
V10	0	1
V11	0	1
V12	0	0
V13	1	1
V14	1	1
V15	0	0
V16	1	1
V17	0	1
V18	0	1
V19	0	1
V20	1	1

Fonte: CopyDetect – dados simulados.

Tabela 5.3 - Habilidades estimadas e respostas corretas para os examinandos 30 (copiador) e 70 (fonte).

Examinandos	Habilidades estimadas	Nº de respostas corretas
30 (suspeito de cópia)	-0,3753	8
70 (fonte)	2,5832	17

Fonte: CopyDetect1 – dados simulados.

Tabela 5.4 – Nº de respostas entre os examinandos 30 (copiador) e 70 (fonte).

Respostas	Quantidade
Corretas idênticas	8
Incorretas idênticas	3
Idênticas	11

Fonte: CopyDetect – dados simulados.

Inicialmente considerando os indivíduos 30 e 70 como copiador e fonte, respectivamente. Testou-se a capacidade dos índices em detectar possível fraude com base em diversos níveis de significância (α).

Logo, nota-se na Tabela 5.5 que inicialmente os índices propostos não apresentaram evidências de fraude com base em diversos níveis de significância (α) adotados.

Tabela 5.5 – Avaliação de índices através de nível de significância α .

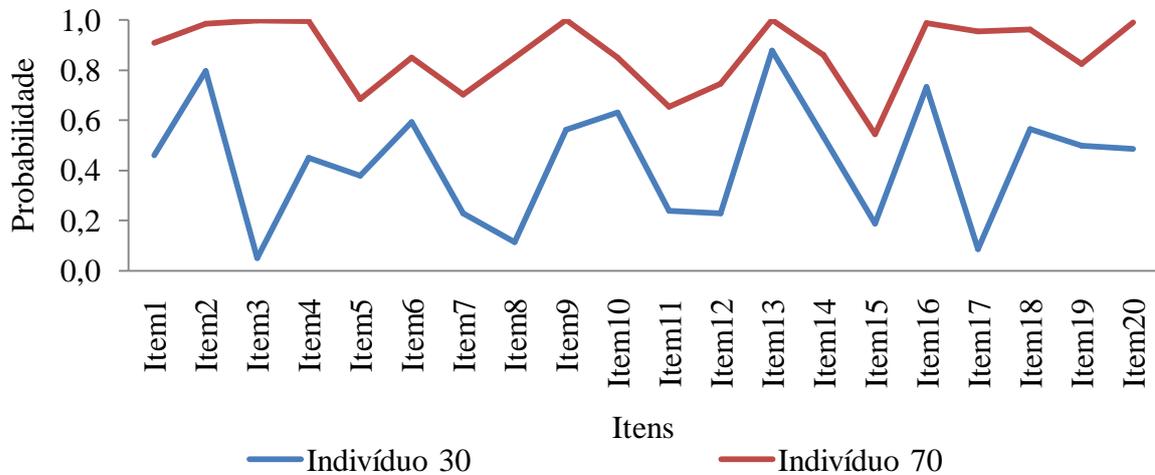
Índices	α		
	0.05	0.01	0.001
ω	Sem fraude	Sem fraude	Sem fraude
GBT	Sem fraude	Sem fraude	Sem fraude
K	Sem fraude	Sem fraude	Sem fraude
K_1	Sem fraude	Sem fraude	Sem fraude
K_2	Sem fraude	Sem fraude	Sem fraude
S_1	Sem fraude	Sem fraude	Sem fraude
S_2	Sem fraude	Sem fraude	Sem fraude

Fonte: CopyDetect – dados simulados.

5.1.3.1 Índice ômega

Nota-se na Figura 5.1, uma grande discrepância na probabilidade de o indivíduo fonte acertar cada item com relação ao indivíduo suspeito, fato este também constatado pela heterogeneidade de habilidades entre eles relatada acima, ocasionando obviamente maior possibilidade de resposta correta pela fonte devido sua alta proficiência.

Figura 5.1 - Probabilidades estimadas para respostas corretas para os indivíduos copiador (30) e fonte (70).



Fonte: CopyDetect – dados simulados.

Considerando os resultados expressos na Tabela 5.6, observa-se a aceitação da hipótese H_0 , ao compararmos a probabilidade de correspondência (p -valor) = 0,32347 com um nível de significância $\alpha = 5\%$. Caso já sinalizado anteriormente quando se comparou o referido índice com diversos níveis de significâncias apresentados na Tabela 5. 5.

Tabela 5.6 – Valores obtidos para o índice ω .

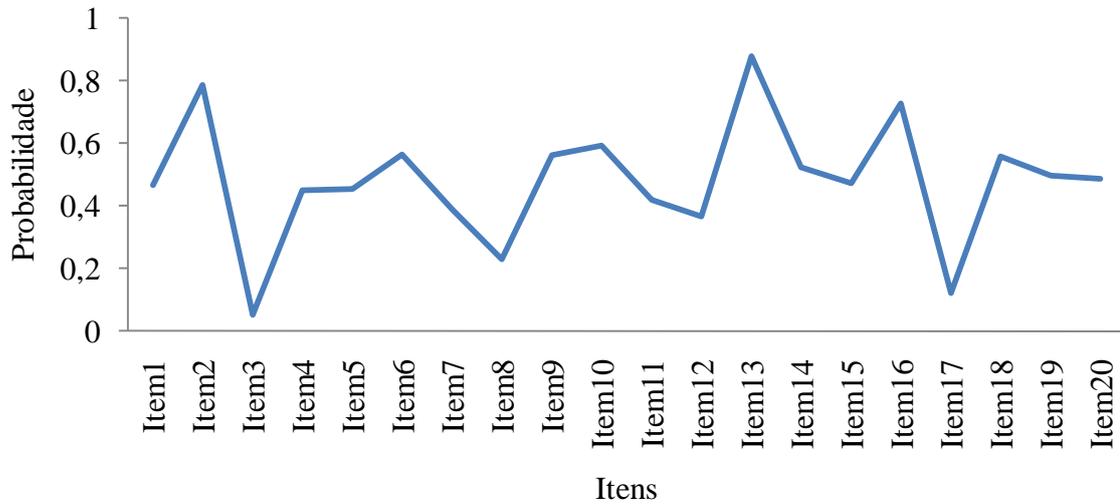
Resultados	Valores
Número esperado de respostas idênticas	10,105
Desvio-padrão do número esperado de respostas idênticas	1,955
Índice ω	0,458
Probabilidade de correspondência (p – valor)	0,32347

Fonte: CopyDetect – dados simulados.

5.1.3.2 Testes Binomial Generalizado

Na Figura 5.2, observa-se claramente que os itens 13, 2 e 16 são os que apresentam as maiores probabilidades de correspondência de resposta, respectivamente. Em contrapartida o item 3 foi o que apresentou menor probabilidade de coincidência de respostas.

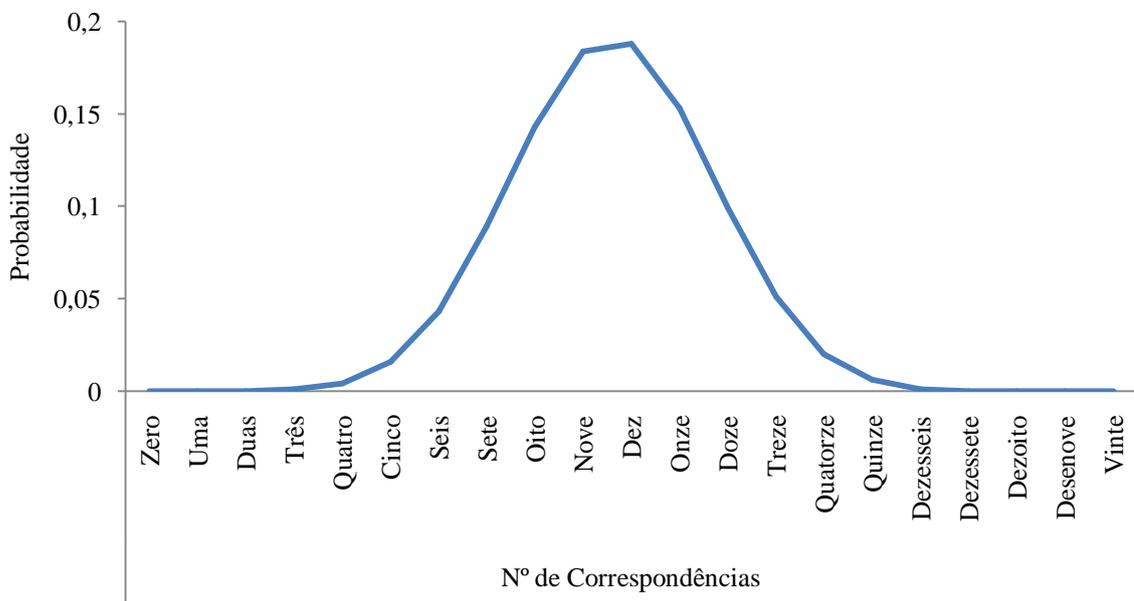
Figura 5.2 - Probabilidades de correspondência de respostas por item.



Fonte: CopyDetect – dados simulados.

Na Figura 5.3, verifica-se que a grande massa de probabilidade está entre seis e treze coincidências de respostas. Por conseguinte, as extremidades apresentam baixíssimas probabilidades no número de correspondências, inclusive com valor zero.

Figura 5.3 – Distribuição de probabilidades exata para determinado número de correspondência de respostas.



Fonte: CopyDetect – dados simulados.

Com base na Tabela 5.7, verifica-se que não há indícios de fraude, ou seja, aceita-se H_0 , ao compararmos a probabilidade de se observar 11 ou mais correspondências de respostas = 0,33148 com um nível de significância $\alpha = 5\%$.

Tabela 5.7 – Valor obtido para o Teste Binomial Generalizado.

Resultados	Valor
Probabilidade de se observar 11 ou mais correspondências (p-values)	0,33148

Fonte: CopyDetect – dados simulados.

5.1.3.3 Índice K

Número de pontuação incorreta para o Examinando 30 (suspeito de cópia) = 12.

Número de pontuação incorreta para o Examinando 70 (fonte suspeita) (w_s) = 3

Número examinandos no subgrupo com número de pontuação incorreta igual a 12 ($m_{c'j'}$) = 25.

Número de respostas incorretas idênticas entre cada examinando neste subgrupo e o examinando fonte ($M_{c'j'}$):

1 1 3 3 3 3 3 2 3 2 2 1 3 2 3 2 3 2 2 3 1 2 3 3 3

Média do número de respostas incorretas idênticas entre cada examinando neste subgrupo e o examinando fonte ($\bar{m}_{c'j'}$)

$$(\bar{m}_{c'j'}) = \frac{\sum M_{c'j'}}{m_{c'j'}} = \frac{59}{25} = 2,36$$

Probabilidade de correspondência de resposta incorreta idênticos com examinando fonte para este subgrupo em relação ao número de respostas incorretas:

$$p_{c'}^* = \frac{(\bar{m}_{c'j'})}{(w_s)} = \frac{2,36}{3} = 0,7867$$

Probabilidade de se observar 3 ou mais respostas incorretas idênticas através da distribuição binomial:

- $m_{c'c}$ número de respostas similares erradas entre c e s ;
- w_s número de respostas erradas de s .

$$p = \sum_{g=3}^3 \binom{3}{3} (0,7867)^3 (1 - 0,7867)^{(3-3)} = 0,4868$$

Através de p , conclui-se que não há existência de fraude ao compararmos com um $\alpha = 5\%$ nível de significância.

5.1.3.4 Variantes do índice K (K_1, K_2, S_1)

Na Tabela 5.8, a análise para constatação de fraude será feita, tendo como referência o grupo em que o indivíduo 30 pertence, ou seja, considerando o quantitativo de pontuação incorreta deste grupo (12 respostas incorretas).

Tabela 5.8 – Resultados obtidos para as variantes do índice (K_1, K_2, S_1).

Nº de pontuação incorreta do grupo	Nº de examinandos	Nº médio de correspondências incorretas com examinando fonte	Previsão do número médio de correspondências incorretas idênticas com examinando fonte		
			Modelo linear (K_1)	Modelo quadrático (K_2)	Modelo log-linear (S_1)
0	0	NA	NA	NA	NA
1	1	1,000	0,972	0,787	1,127
2	6	0,667	1,087	0,963	1,194
3	8	1,125	1,201	1,132	1,266
4	6	1,500	1,316	1,294	1,342
5	17	1,176	1,431	1,449	1,423
6	22	1,591	1,545	1,596	1,508
7	18	1,722	1,660	1,736	1,599
8	18	1,889	1,774	1,869	1,695
9	16	2,000	1,889	1,994	1,797
10	22	2,364	2,003	2,113	1,905
11	24	2,375	2,118	2,223	2,019
12	25	2,360	2,233	2,327	2,140
13	15	2,400	2,347	2,424	2,269
14	20	2,450	2,462	2,513	2,405
15	14	2,571	2,576	2,594	2,550
16	8	2,375	2,691	2,669	2,703
17	6	2,500	2,805	2,736	2,866
18	3	3,000	2,920	2,796	3,038
19	1	3,000	3,035	2,849	3,220
20	0	NA	NA	NA	NA

Fonte: CopyDetect – dados simulados.

Número de pontuação incorreta para o examinando 30 (suspeito de cópia) = 12.

Número de pontuação incorreta para o examinando 70 (fonte suspeita) = 3.

Probabilidade binomial de correspondência em uma resposta incorreta Idêntica com examinando fonte para o subgrupo com 12 pontuações incorretas:

- Para obtenção do índice K_1 , temos $\hat{p}_1^* = \frac{2,233}{3} = 0,744$.
- Para obtenção do índice K_2 , temos $\hat{p}_2^* = \frac{2,327}{3} = 0,776$.
- Os valores de \hat{p}_1^* e \hat{p}_2^* são obtidos de forma similar a p_c^* .

Probabilidade de se observar 3 ou mais correspondências incorretas idênticas através da distribuição Binomial:

- O índice $K_1(p - valor) = \sum_{g=3}^3 \binom{3}{3} (0,744)^3 (1 - 0,744)^{(3-3)} = 0,412$.
- O índice $K_2(p - valor) = \sum_{g=3}^3 \binom{3}{3} (0,776)^3 (1 - 0,776)^{(3-3)} = 0,467$.

Probabilidade de se observar 3 ou mais correspondências incorretas idênticas através da distribuição Poisson, considerando a média ajustada do grupo em que pertence o indivíduo 30, teremos:

$$S_1(p - valor) = \sum_{g=m_c'}^{w_s} \frac{e^{-\lambda} \lambda^g}{g!} = \frac{e^{-2,14} (2,14)^3}{3!} = 0,192.$$

Conclusões para a probabilidade de se observar 3 ou mais correspondências incorretas idênticas adotando $\alpha = 5\%$:

Observou-se que estes índices não detectaram cópia de resposta, ou seja, a tese que o indivíduo 30 teria copiado as respostas do examinando 70 não se confirma. Logo, aceita-se a hipótese nula.

5.1.3.5 Índice S_2

Na tabela 5.9, encontra-se o número de pontuação incorreta onde se encontra o indivíduo 30, apresentando 12 pontuações incorretas com um total de 25 examinandos neste grupo, o qual apresenta média de correspondência de respostas incorretas com o indivíduo fonte de 2,360 e número ponderado médio observado de correspondência de respostas corretas idênticas com o examinando fonte de 0,261.

Tabela 5.9 - Resultados obtidos para o índice S_2 .

Nº de pontuação incorreta do grupo	Nº de examinandos	Nº médio de correspondências incorretas com examinando fonte	Nº ponderado médio observado de correspondências corretas idênticas com examinando fonte
0	0	NA	NA
1	1	1,000	0,004
2	6	0,667	0,011
3	8	1,125	0,015
4	6	1,500	0,049
5	17	1,176	0,050
6	22	1,591	0,073
7	18	1,722	0,089
8	18	1,889	0,111
9	16	2,000	0,151
10	22	2,364	0,216
11	24	2,375	0,251
12	25	2,360	0,261
13	15	2,400	0,305
14	20	2,450	0,322
15	14	2,571	0,270
16	8	2,375	0,333
17	6	2,500	0,251
18	3	3,000	0,060
19	1	3,000	0,000
20	0	NA	NA

Fonte: CopyDetect – dados simulados.

A Tabela 5.10, apresenta a predição de correspondência entre o número de respostas incorreto e corretas ponderadas através do modelo log-linear, o qual apresenta a média estimada para o modelo de Poisson de $\mu = 3,173$ para o grupo de 12 respostas incorretas.

Tabela 5.10 - Predição de correspondência entre o número de respostas incorreto e corretas ponderadas através do modelo log-linear.

Nº de pontuação incorreta do grupo	Valores previstos
0	NA
1	2,078
2	2,159
3	2,244
4	2,332
5	2,423
6	2,519
7	2,617
8	2,720
9	2,827
10	2,938
11	3,053
12	3,173
13	3,297
14	3,427
15	3,561
16	3,701
17	3,846
18	3,997
19	4,154
20	NA

Fonte: CopyDetect – dados simulados.

Para o índice S_2 ($p - valor$) = 0,391, conclui-se que não há evidência de fraude comparando S_2 com $\alpha = 5\%$.

2ª Etapa da Simulação:

Na Tabela 5.11, notou-se que para os pares suspeitos de copiar as respostas da fonte, verificou-se a presença de fraude por parte de alguns indivíduos detectados a partir de alguns indices, conforme explanação a seguir:

Dentre os índices propostos apenas W2 (Ômega) e GBT2 (Teste Binomial Generalizado), detectaram casos de cópias junto aos indivíduos selecionados ao acaso, quando comparados seus resultados com um nível de significância igual a 5%, como segue abaixo:

Através do índice W2, identificou-se como fraudadores os indivíduos do vetor 2, sendo eles 4, 15 e 129, com probabilidades 0,0413, 0,0054 e 0,0065, respectivamente.

Para o índice GTB2, identificou-se como fraudadores os indivíduos do vetor 2, sendo eles 15 e 129, com probabilidades 0,0105 e 0,0261, respectivamente.

Ressalta-se que a fraude foi identificada após a recombinação dos dados, conforme descrito inicialmente para a 2ª etapa.

Tabela 5.11 – Índices para 5 pares de vetores selecionados ao acaso para constatação de fraude considerando nível de significância de $\alpha = 5\%$.

Pares de cópias	Vetores de cópia		Índices													
	V1	V2	W1	W2	GBT1	GBT2	K_1	K_2	K1_1	K1_2	K2_1	K2_2	S1_1	S1_2	S2_1	S2_2
1	66	145	0,7456	0,2617	0,8445	0,3631	0,5923	0,5148	0,8660	0,3906	0,7958	0,3721	0,8584	0,3902	0,7899	0,4116
2	81	4	0,2236	0,0413	0,4252	0,1407	0,5423	0,2764	0,4143	0,1717	0,456	0,2215	0,2649	0,1478	0,3579	0,1974
3	199	15	0,3157	0,0054	0,4204	0,0105	0,5385	0,1555	0,4498	0,0986	0,4965	0,1201	0,2982	0,1173	0,3384	0,1590
4	92	129	0,5068	0,0065	0,6403	0,0261	0,5446	0,1546	0,5523	0,0973	0,6278	0,1178	0,4025	0,1208	0,4236	0,2017
5	55	250	0,9377	0,1087	0,8581	0,0814	0,9300	0,3813	0,9213	0,2892	0,9311	0,3414	0,4590	0,1754	0,6615	0,3252

Fonte: CopyDetect – dados simulados.

5.1.4 Considerações Finais

Dentro da proposta do trabalho podemos averiguar estudo sobre métodos estatísticos para a detecção de fraudes em testes, mostramos novas possibilidades de ferramentas estatísticas baseadas na Teoria Clássica dos Testes ou na Teoria da Resposta ao Item, para tentar subsidiar melhor a lisura de testes avaliativos, dentre vários índices presentes na literatura explicitamos alguns e mostramos através de exemplo um algoritmo implementado para identificar possíveis sujeitos que tenham praticados atos fraudulentos (cola) durante a aplicação do teste.

Averiguamos através da simulação que em um primeiro momento não foi possível detectar fraude, mesmo tendo a suposição que o sujeito denominado como 30 tenha colado da fonte 70.

Porém, observamos que ao adotar as respostas do copiadador para o individuo fonte, criando a possibilidade que o possível copiadador tenha visto a folha de resposta da fonte para tentarmos identificar cinco pares de indivíduos aleatórios suspeitos de fraude que apresentaram as mesmas características de respostas erradas que da fonte. Logo, constatamos que alguns indivíduos haviam colado no teste, pois os índices (Teste Binomial Generalizado e $\hat{\Omega}$) apresentaram valores de probabilidade abaixo de alfa (5%).

Sugerimos a construção de trabalhos futuros com vista a ampliação e aplicação de um destes índices em especial aos adotados pela TRI em banco de dados real, pois através de amostras bem maiores as possibilidades de constatação de fraudes é significativa.

APÊNDICE A

Algoritmo para Geração de Dados

```

install.packages("CopyDetect")
require(CopyDetect)
#Load irtoys package
require(irtoys)
#Define o número de itens e número de simulações
n=20
N=250
#Geração dos parâmetros dos itens para os dados simulados
#Na primeira, segunda e terceira colunas temos os parâmetros de discriminação, parâmetro de
#dificuldade e o parâmetro de acerto ao acaso, respectivamente.
#ipar receberá os parâmetros dos itens.
ipar <- cbind(rlnorm(n, meanlog = 0, sdlog = .5), rnorm(n,0,1),rep(0,n))
# responses receberá as simulações de respostas para itens dicotômicos.
responses <- as.data.frame(sim(ip=ipar, x=rnorm(N)))
# est.ipar receberá as estimativas para os parâmetros dos itens utilizando o modelo logístico
#de 2 parâmetros.
est.ipar <- est(responses, model = "2PL", engine = "ltm")$est
est.ipar
#Neste momento, defini-se que o examinando suspeito (copiador) é o examinando 30 e a
#fonte suspeita é o examinando 70.
CopyDetect1(data=responses,item.par=est.ipar,pair=c(30,70))
#Vamos calcular estes índices para 100 pares aleatórios de examinandos
k=100 # será para 100 indivíduos.
pairs <- as.data.frame(matrix(k,ncol=2))
for(i in 1:k){
  d <- sample(1:N,2,replace=FALSE)
  pairs[i,1]=d[1]
  pairs[i,2]=d[2]
}
pairs$W <- NA
pairs$GBT <- NA
pairs$K <- NA
pairs$K1 <- NA
pairs$K2 <- NA
pairs$S1 <- NA
pairs$S2 <- NA
for(i in 1:k){
  x <- CopyDetect1(data=responses,item.par=est.ipar,pair=c(pairs[i,1],pairs[i,2]))
  pairs[i,]$W=x$W.index$p.value
  pairs[i,]$GBT=x$GBT.index$p.value
  pairs[i,]$K=x$K.index$k.index
  pairs[i,]$K1=x$K.variants$K1.index
  pairs[i,]$K2=x$K.variants$K2.index
}

```

```

pairs[i,]$S1=x$K.variants$S1.index
pairs[i,]$S2=x$K.variants$S2.index
}
#Espera-se um percentual de 5% dentre os 100 indivíduos sejam detectados ao acaso.
length(which(pairs$W<.05))/k
length(which(pairs$GBT<.05))/k
length(which(pairs$K<.05))/k
length(which(pairs$K1<.05))/k
length(which(pairs$K2<.05))/k
length(which(pairs$S1<.05))/k
length(which(pairs$S2<.05))/k
#Agora, calcular esses índices para 5 pares de cópia de resposta.
#Inicialmente, escolhe-se aleatoriamente um examinando suspeito de cópia, em seguida
#escolhe-se um examinando fonte suspeito de ceder à cópia e finalmente seleciona-se
#aleatoriamente 10 itens (50% cópia)
#Desta forma, sustitui-se o vetor de resposta do examinando fonte
#No vetor de resposta do examinando copiado, adotando a idéia que o examinando
#copiado olha para o teste do examinando fonte e copia 10 dos 20 itens do teste.
copy.pairs <- as.data.frame(matrix(nrow=5,ncol=2))
for(i in 1:5){
  d <- sample(1:N,2,replace=FALSE)
  copy.pairs[i,1]=d[1] #Examinando copiado hipotético
  copy.pairs[i,2]=d[2] #Examinando fonte hipotético
}
new.responses <- responses
for(i in 1:5){ #Simula a cópia de resposta para cada par de cópia de resposta
  copy.items <- sample(1:n,10,replace=FALSE)
  new.responses[copy.pairs[i,1],copy.items]=new.responses[copy.pairs[i,2],copy.items]
}
# Calcula os índices para os pares considerando os vetores de resposta originais
copy.pairs$W1 <- NA
copy.pairs$GBT1 <- NA
copy.pairs$K_1 <- NA
copy.pairs$K1_1 <- NA
copy.pairs$K2_1 <- NA
copy.pairs$S1_1 <- NA
copy.pairs$S2_1 <- NA
for(i in 1:5){
  x <- CopyDetect1(data=responses,item.par=est.ipar,pair=c(copy.pairs[i,1],copy.pairs[i,2]))
  copy.pairs[i,]$W1=x$W.index$p.value
  copy.pairs[i,]$GBT1=x$GBT.index$p.value
  copy.pairs[i,]$K_1=x$K.index$k.index
  copy.pairs[i,]$K1_1=x$K.variants$K1.index
  copy.pairs[i,]$K2_1=x$K.variants$K2.index
  copy.pairs[i,]$S1_1=x$K.variants$S1.index
  copy.pairs[i,]$S2_1=x$K.variants$S2.index
}
#Calcula os índices para os mesmos pares considerando as novas cópia simulada para os
#vetores de #resposta
est.ipar2 <- est(new.responses, model = "2PL", engine = "ltm")$est

```

```

copy.pairs$W2 <- NA
copy.pairs$GBT2 <- NA
copy.pairs$K_2 <- NA
copy.pairs$K1_2 <- NA
copy.pairs$K2_2 <- NA
copy.pairs$S1_2 <- NA
copy.pairs$S2_2 <- NA
for(i in 1:5){
  x<-CopyDetect1(data=new.responses,item.par=est.ipar2,
pair=c(copy.pairs[i,1],copy.pairs[i,2]))
  copy.pairs[i,$W2]=x$W.index$p.value
  copy.pairs[i,$GBT2]=x$GBT.index$p.value
  copy.pairs[i,$K_2]=x$K.index$k.index
  copy.pairs[i,$K1_2]=x$K.variants$K1.index
  copy.pairs[i,$K2_2]=x$K.variants$K2.index
  copy.pairs[i,$S1_2]=x$K.variants$S1.index
  copy.pairs[i,$S2_2]=x$K.variants$S2.index
}
print(copy.pairs) #Aqui finaliza-se o algoritmo.

```

REFERÊNCIAS BIBLIOGRÁFICAS

ACFE. Association of Certified Fraud Examiners. **Report to the Nations - On Occupational Fraud and Abuse - 2010 Global Fraud Study.** Disponível em: <http://www.acfe.com/uploadedfiles/acfe_website/content/documents/rtnn-2010.pdf> Acesso: **01-03-2016.**

ACFE-Association of Certified Fraud Examiners. **Report to the Nations - On Occupational Fraud and Abuse - 2012 Global Fraud Study.** Disponível em: <https://www.acfe.com/uploadedFiles/ACFE_Website/Content/rtnn/2012-report-to-nations.pdf> Acesso: **01-03-2016.**

AGRESTI, A. An introduction to categorical data analysis. NY Wiley.1996.

AICPA. American Institute of Certified Public Accountants American. Disponível em: <<http://www.aicpa.org>>. Acesso: **01-03-2016.**

ALLEN, Jeff. **Relationships of Examinee Pair Characteristics and Item Response Similarity,** 2012. Disponível em: <https://cete.ku.edu/sites/cete.drupal.ku.edu/files/docs/2012_test_fraud_presentations/Allen_Relationships%20of%20Examinee%20Pair%20Characteristics.pdf> Acesso em: **08-03-2016.**

ANDRADE, Dalton Francisco de et.al. Teoria da Resposta ao Item: Conceitos e Aplicações. **14ª SINAPE, 2000.**

ANGOFF, W.H. **The development of statistical indices for detecting cheaters.** *Journal of the American Statistical Association*, 69(345): 1974. 44-49.

ARANHA, Maria Lúcia de Arruda. **História da Educação.** 2ª ed. Ver. E atual. São Paulo: Moderna, 1996.

ARAUJO, Eutalia Aparecida Candido de et al. **Teoria da Resposta ao Item.** Artigo publicado pela Rev Esc Enferm USP 2009; 43(Esp):1000-8. Disponível em: <<http://www.scielo.br/pdf/reeusp/v43nspe/a03v43ns.pdf>>. Acesso: **16-11-2015.**

BAY, L. G. Detection of cheating on multiple-choice examinations. Paper presented at the annual meeting of the American Educational Research Association, San Francisco. 1995.

BELLEZZA, F.S. e BELLEZZA, S.F.. **Detection of cheating on multiple-choice tests by using error-similarity analysis.** *Teaching of Psychology*, 16(3): 1989.151- 155.

BELOV, D. I., PASHLEY, P. J., LEWIS, C., & ARMSTRONG, R. D.. **Detecting aberrant responses with Kull- back–Leibler distance.** In K. Shigemasu, A. Okada, T. Imaizumi, & T. Hoshino (Eds.), *New Trends in Psychometrics* (pp. 7-14). Tokyo: Universal Academy Press.2007.

BELOV, Dmitry I e ARMSTRONG, Ronald D.. **Detection of Answer Copying via Kullback–Leibler Divergence and K-Index.** Artigo publicado por Law School Admission Council Research Report - LSAC, 2009. **Disponível em:** < [http://www.lsac.org/docs/default-source/research-\(lsac-resources\)/rr-09-01.pdf](http://www.lsac.org/docs/default-source/research-(lsac-resources)/rr-09-01.pdf)>**Acesso em:** 13-04-2016.

BELOV, Dmitry I. **Detection of Answer Copying Based on the Structure of a High-Stakes Test:** Aplicado a Mensuração Psicológica. Artigo publicado por SAGE Publishing, 2011. **Disponível em:** <https://www.researchgate.net/publication/254078208_Detection_of_Answer_Copying_Based_on_the_Structure_of_a_High-Stakes_Test>**Acesso em:** 12-04-2016.

Biblioteca Digital da Câmara dos Deputados – Centro de Documentação e Informação, **Constituição da República Federativa Brasil.** 35ª Edição, 2012. **Disponível em:**<http://bd.camara.gov.br/bd/bitstream/handle/bdcamara/15261/constituicao_federal_35ed.pdf?sequence=9>. **Acesso em: 10-05-2016.**

BLANQUÉ, Pascal. *Crisis and Fraud.* *Journal of Financial Regulation and Compliance*, v. 11 n. 1, p. 1, 2002.

BLISS, T. J. **Statistical Methods to Detect Cheating on Tests:** A Review of the Literature. Artigo publicado na National Conference of Bar Examiners, 2012. **Disponível em:**< <http://www.ncbex.org/pdfviewer/?file=%2Fdmsdocument%2F57> >**Acesso em:** 05-04-2016.

BRASIL, Código Penal pela Lei nº 12.550, de 15 de Dezembro de 2011. **Disponível em:** <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/Lei/L12550.htm>. **Acesso em: 13-05-2016.**

BRASIL, Constituição da Republica dos Estados Unidos do Brasil 34| Constituição da Republica dos Estados Unidos do Brasil (de 16 de Julho de 1934). **Disponível em:**

<<http://presrepublica.jusbrasil.com.br/legislacao/92083/constituicao-da-republica-dos-estados-unidos-do-brasil-34#art-168>>. **Acesso em: 10-05-2016.**

BUNN, D. N., CAUDILL, S. B. and GROPPER, D. M.. **Crime in the Classroom: An Economic Analysis of Undergraduate Student Cheating Behavior.** Journal of Economic Education, 23, pp. 197–207, 1992. **Disponível em:**<<http://www.sciencedirect.com/science/article/pii/S1053535796900116>>. **Acesso em: 02-05-2016.**

CAMACHO, X. G. O. Propiedades de algunos estadísticos empleados para la detección de copia de respuestas en La medición educativa. Tese de Doutorado apresentada a Faculdade de Educação da Universidade de Madri, Espanha, 2011. **Disponível em:**<<http://eprints.ucm.es/13293/1/T33095.pdf>>**Acesso em: 30-05-2016.**

CHACHAMOVICH, E.. **Teoria da Resposta ao Item: aplicação ao modelo de Rasch em desenvolvimento e validação de instrumentos em saúde mental.** Unpublished doctoral dissertation, Programa de Pós-Graduação em Ciências Médicas: Psiquiatria, Faculdade de Medicina, Universidade Federal do Rio Grande do Sul/UFRGS, Porto Alegre, Brasil.2007.

CIZEK, G. J. **Cheating on tests: How to do it, detect it, and prevent it.** Mahwah (p. 268). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.1999.

COUTINHO, Lindinalva de Almada, **Ferramenta para Estimar a Proficiência de Indivíduos Baseada na Teoria de Resposta ao Item.** Dissertação de Mestrado em Engenharia da Computação e Sistemas, na área de concentração: Informática na Educação, apresentada à Coordenadoria do Programa de Pós-Graduação em Engenharia da Computação e Sistemas da Universidade Estadual do Maranhão, São Luis, Brasil, 2013. **Disponível em:**http://www.mestrado.engcomp.uema.br/wp-content/uploads/2014/03/Disserta%C3%A7ao_final_Lidinalva.pdf>. **Acesso: 01-12-2015.**

CRISTOFARO, Elizabeth Agnes Urban. **Uma Abordagem Bayesiana para Análise de Fraude de Subscrição em Telecomunicações.** São Carlos: Dissertação de Mestrado apresentada ao Departamento de Estatística da Universidade Federal de São Carlos – Des/UFSCar, São Carlos, Brasil, 2006. **Disponível em:** <http://www.bdtd.ufscar.br/htdocs/tedeSimplificado//tde_busca/arquivo.php?codArquivo=4330>**Acesso: 04-03-2016.**

EMBRETSON, S. E., & REISE, S. P.. Item Response Theory for psychologists. Mahwah, NJ: Lawrence Erlbaum.2000.

Estudantes de medicina são presos suspeitos de fraudar vestibulares. **Disponível em:**<<http://g1.globo.com/goias/noticia/2015/11/estudantes-de-medicina-sao-presos-suspeitos-de-fraudar-vestibulares.html>>. **Acesso:10-05-2016.**

FASK, Alan, ENGLANDER, Fred, e WANG, Zhaobo. **On the Integrity of Online Testing for Introductory Statistics Courses: A Latent Variable Approach. Practical Assessment, Research & Evaluation.** A peer-reviewed electronic journal. .2015. ISSN: 1531-7714 Vol. 20 N° 10, 2015. **Disponível em:** <<http://pareonline.net/getvn.asp?v=20&n=10>>**Acesso em:** 17-03-2016.

CARVALHO FILHO, José dos Santos. **Manual de direito administrativo.** 18. ed. Rio de Janeiro. Lumens, 2007.

FONTANIVE, N. S.; ELLIOT, L. G.; KLEIN, R. Os desafios da apresentação dos resultados da avaliação de sistemas escolares a diferentes públicos. **REICE - Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación,** v. 5, n. 2e, 2007.

FRARY, R. B., TIDEMAN, T N., & WATTS, T M. Indices of cheating on multiple-choice tests. *Journal of Educational Statistics,* 6, 152 -165.1977.

Fraudes nos canais eletrônicos. **Disponível em:** <http://www.hsbc.com.br/1/2/br/informacoes/seguranca?WT.ac=HBBR_PVLINK03#/fraudes-nos-canais-eletronicos>**Acesso: 29-02-2016.**

FREITAS, Eduardo Costa de. **Avaliação do sistema de combate às Fraudes Empresariais no Brasil.** São Paulo: Programa de Pós-Graduação em Ciências Contábeis da Universidade Presbiteriana Mackenzie, São Paulo, Brasil, 2012. **Disponível em:** <http://up.mackenzie.br/fileadmin/user_upload/_imported/fileadmin/PUBLIC/UP_MACKENZIE/servicos_educacionais/stricto_sensu/Ciencias_Contabeis/Dissertacoes/Eduardo_Costa_de_Freitas.pdf>**Acesso: 01-03-2016.**

Grupo quebrou sigilo do Enem. **Disponível em:**http://patriciafrancisco.jusbrasil.com.br/noticias/112266909/grupo-quebrou-o-sigilo-do-enem?ref=topic_feed. **Acesso em: 12-05-2016.**

GULLIKSEN, H.. *Theory of mental tests.* New York: Wiley.1950.

HAMBLETON, R. K., & SWAMINTHAN, H.. Item Response Theory: Principles and applications. Boston, MA: Kluwer Nijhoff.1985.

HANSON, B. A.. Statistical Indexes of Response Similarity Derived from the Compound Binomial Distribution. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.1994.

HOLLAND, P. W. **Assessing unusual agreement between the incorrect answers of two examinees**: Using the K-index. (p. 41). Princeton.1996.

Infográfico: **A História dos Concursos Públicos no Brasil**. Disponível em: <<https://www.iobconcursos.com/noticias-sobre-concursos/a-historia-dos-concursos-publicos-no-brasil/1168>>. Acesso em: **10-05-2016**.

MOREIRA JUNIOR, Fernando de Jesus. A Utilização da Teoria da Resposta ao Item como Ferramenta para a Construção de Escalas na Área de Gestão. **RGO Revista Gestão Organizacional** | VOL. 6 - Edição Especial – 2013. pp. 243-259. Disponível em: <[http://www.sigmees.com/files/1519-7178-1-PB%20\(1\)ARTIGOFERNANDO2013.pdf](http://www.sigmees.com/files/1519-7178-1-PB%20(1)ARTIGOFERNANDO2013.pdf)>. Acesso: **29-11-2015**.

KARABATSOS, G. **Comparing the aberrant response detection performance of thirty-six person-fit statistics**. *Applied Measurement in Education*, 16(4), 2003. 277-298. doi:10.1207/S15324818AME1604_2

KHALID, L. S., MEHMOOD, Z., e REHMAN, N. A.. **Statistical Methods for Answer Copying – A Brief Overview**. British Journal of Arts and Social Sciences, 2011. ISSN: 2046-9578 Vol. 1 N° 1, 2011. Disponível em: <<http://www.bjournal.co.uk/BJASS.aspx>> Acesso em: 10-03-2016.

KOVACH, S.. **Detecção de Fraudes em Transações Financeiras via Internet em Tempo Real**. Tese de Doutorado apresentada a Escola Politécnica da Universidade de São Paulo, São Paulo - Brasil, 2011. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/3/3141/tde-09082011-155153/pt-br.php>> Acesso em: 26-04-2016.

LEWIS, C. & THAYER, D. T. . The power of the k-index (or PMIR) to detect copying (ETS Research Report No. 98-49). Princeton, NJ: Educational Testing Service.1998.

LORD, F. M., & NOVICK, M. R. *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley, 1968.

LUCKESI, Cipriano Carlos. **Filosofia da Educação**. São Paulo: Cortez, 1991.

_____. *In*: FIDALGO, Sueli Salles. **A avaliação de ensino aprendizagem: ferramenta para a formação de agentes críticos**. Pontifícia Universidade Católica de São Paulo, 2002, Dissertação de Mestrado.

MEIJER, R. R.. **Person-fit research: An introduction**. *Applied Measurement in Education*, 9, 3–8. 1996.

MELLO, Celso Antônio Bandeira de. **Curso de direito administrativo**. 21. ed. São Paulo. Malheiros, 2006.

_____. **Regime constitucional dos servidores da administração direta e indireta**. São Paulo: RT, 1990.

MINGOT, Sueli Aparecida. *Análise de dados através de estatística multivariada. Uma abordagem aplicada*. V. 01, editora UFMG, 2005.

Monitor das Fraudes. **Disponível em:** <<http://www.fraudes.org/showpage1.asp?pg=2>>
Acesso: 25-02-2016.

PASQUALI, Luiz e Primi, Ricardo. *Fundamentos da Teoria da Resposta ao Item –TRI. Avaliação Psicológica*, 2003, (2) 2, pp 99-110. **Disponível em:** http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712003000200002
Acesso: 16-11-2015.

PIMENTA, Maria Alzira de Almeida e PIMENTA, Sônia de Almeida. **Fraude em Avaliações de aprendizagens: Estudo Comparativo entre o Nordeste e o Sudeste do Brasil, 2011**. **Disponível em:** <<http://www.saece.org.ar/docs/congreso4/trab37.pdf>> **Acesso em: 02-03-2016.**

PINHEIRO, Alessandra Noeli Craveiro, **Modelos de Resposta ao Item com função de ligação t-Assimétrica**. São Carlos: Dissertação de Mestrado apresentado ao Departamento de Estatística da Universidade Federal de São Carlos – Des/UFSCar, São Carlos, Brasil, 2007.
Disponível em:

<http://www.btdt.uhttp://www.btdt.ufscar.br/htdocs/tedeSimplificado//tde_busca/arquivo.php?codArquivo=1549>. **Acesso: 24-11-15.**

PRIMI, Ricardo, *Psicometria: fundamentos matemáticos da Teoria Clássica dos Testes. Avaliação Psicológica*, 2012, *11(2)*, pp. 297-307. Artigo publicado pela Red de Revistas Científicas de América Latina y el Caribe, España y Portugal. **Disponível em:** <<http://www.redalyc.org/html/3350/335027501015/>>. **Acesso em: 01-06-2016**

QUADROS, Fernanda Maciel de, **Comparação da Teoria Clássica dos Testes com a Teoria da Resposta ao Item aplicado ao subteste Semelhanças da WASI**. Rio Grande do Sul: Trabalho de Conclusão de Curso apresentado ao Departamento de Estatística da Universidade Federal do Rio Grande do Sul/UFRGS, Porto Alegre, Brasil, 2010. **Disponível em:** <<https://www.lume.ufrgs.br/bitstream/handle/10183/24863/000749619.pdf?sequence=1>> **Acesso: 21-11-2015.**

RAMOS, François Silva. **Fraude Acadêmica: Uma Análise Ético-Legislativa**. Uberaba-MG: Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Educação da Universidade de Uberaba, Uberaba, Brasil, 2006. **Disponível em:** <<http://www.uniube.br/biblioteca/novo/base/teses/BU000279533.pdf>> **Acesso: 04-03-2016.**

REGO, Sergio. Índice H, autoria e integridade na produção científica. *Rev. bras. educ. med.* [online]. 2010, vol.34, n.2.

ROCKNESS, Howard; ROCKNESS, Joane. Legislated Ethics: From Enron to Sarbanes Oxley, the Impact on Corporate America. **Journal of Business Ethics** v. 57, p. 31-54, 2005.

SACRISTAN, José Gimeno. A educação que temos e a educação que queremos. In: SACRISTAN, José Gimeno. **A Educação no século XXI: os desafios do futuro**. Porto Alegre: Editora Artmed, 2007.

SAMEJIMA, F. A. Estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph*, 17.1969.

SARTES, Andreoli et.al. Avanços na Psicometria: Da Teoria Clássica dos Testes à Teoria de Resposta ao Item. **Psicologia: Reflexão e Crítica**, Vol. 26, núm. 2, 2013, pp. 241-250; **Universidade Federal do Rio Grande do Sul Porto Alegre, Brasil. Disponível em:** <www.scielo.br/prc>. **Acesso: 21-11-2015.**

SAUPE, J.L.. **An Empirical Model for the Corroboration of Suspected Cheating on Multiple-Choice Tests.** *Educational and Psychological Measurement*, 20, 1960. 475-489.

SILVA, Etienne Pereira da. Teoria da Resposta ao Item: Análise de Atitude dos Graduandos em Relação às Disciplinas de Estatística. Monografia apresentada ao Departamento de Estatística da Universidade de Brasília-UNB, Brasil, 2014, pp.21. **Disponível em:** <http://bdm.unb.br/bitstream/10483/10076/1/2014_%20EtiennePereiradaSilva.pdf>. **Acesso: 03-12-2015.**

SILVA, Jéssica Maria Afonso. Modelos de Resposta ao Item considerando efeitos de ansiedade: Como a ansiedade pode afetar o aluno frente ao teste. **Monografia apresentada ao Curso de Estatística da Universidade Federal de Juíz de Fora-UFJF**, 2013, Brasil, pp.10. **Disponível em:** <<http://www.ufjf.br/cursoestatistica/files/2014/04/Modelo-de-Teoria-de-Resposta-ao-Item-considerando-efeitos-de-ansiedade.-%E2%80%93Como-a-ansiedade-pode-afetar-o-aluno-frente-ao-teste.pdf>>. **Acesso: 24-11-2015.**

SILVA, Rommel Viégas. **A Terceirização na Administração Pública ante o princípio do Concurso.** Trabalho de Conclusão de Curso na forma de Artigo científico, apresentado à FESP Faculdades, João Pessoa-PB, Brasil, 2012. **Disponível em:** <[http://www.fespfaculdades.com.br/painel/uploads/arquivos/TCC%20%20ROMMEL%20\(1\).pdf](http://www.fespfaculdades.com.br/painel/uploads/arquivos/TCC%20%20ROMMEL%20(1).pdf)>. **Acesso: 10-05-2016.**

SOEIRO, Leda e AVELINE, Suelly. **Avaliação Educacional.** Porto Alegre: Sulina, 1982.

SOTARIDONA, L. S. &Meijer, R. R. Two New Statistics to Detect Answer Copying. Artigo publicado pelo Journal of Educational Measurement Spring. Vol. 40. Nº 1, pp 53-69, 2003. **Disponível em:** <https://www.deepdyve.com/lp/wiley/two-new-statistics-to-detect-answer-copying-QzOc0r8oQe?shortRental=true> **Acesso em:** 28-05-2016.

SOTARIDONA, Leonardo Sitchirita. **Statistical Methods for the Detection of Answer Copying on Achievement Tests.** Tese de Doutorado apresentada ao Departamento de Metodologia de Pesquisa, Medição e Análise de dados da Universidade de Twente – Holanda, 2003. **Disponível em:** < http://doc.utwente.nl/41573/1/thesis_Sotaridona.pdf > **Acesso em:** 31-03-2016.

SPEARMAN, C. “General Intelligence”, objectively determined and measured. *Am J Psychol.* 1904;15(2):201-93.

THIESSEN, Bradley Adam. **Relationship Between Test Security Policies and Test Score Manipulations**. Tese de Doutorado em filosofia em bases Psicológicas e Quantitativa (Avaliação Educacional e Estatística) no Colégio de Pós-Graduação da Universidade de Iowa – EUA, 2008. **Disponível em:** <<http://www.bradthiessen.com/html5/docs/thiessen.pdf>> **Acesso em:** 20-03-2016.

THURSTONE, L. L. A Law of comparative judgment. *Psychol Rev.* 1927;34(2):273-86.

THURSTONE, L. L. Attitudes can be measured. *Am J Sociol.* 1928;26(2):249-69.

VAN DER LINDEN, W.J. & SOTARIDONA, L.. **A statistical test for detecting answer copying on multiple-choice tests**. *Journal of Educational Measurement*, 41(4): 361-377.2004.

VICENTIM, Marisa Aparecida Boso. **O Professor PDE e os Desafios da Escola Pública Paranaense: Produção Didático-Pedagógica**. Volume II. **Disponível em:** <http://www.diaadiaeducacao.pr.gov.br/portals/cadernospde/pdebusca/producoes_pde/2012/2012_uel_ped_pdp_marisa_aparecida_boso_vicentim.pdf> **Acesso:** 06-03-2016.

Wikipédia, a enciclopédia livre. **Disponível em:** <https://pt.wikipedia.org/wiki/Esquema_em_pir%C3%A2mide> **Acesso:** 26-02-2016.

WOLLACK, J. A. & COHEN, A. S.. Detection of answer copying with Unknown item and trait parameters. *Applied Psychological Measurement*, 22, 1998, 144-152.

WOLLACK, J. A. A nominal response model approach to detect answer copying. *Applied Psychological Measurement*, 21, 1997.307-320.

WOLLACK, J.A. e COHEN, A. S. Detection of answer copying with Unknown item and trait parameters. *Applied Psychological Measurement*, 22, 1998. 144-152.

ZERPA, Carlos e BARNEVELD, Christina Van.. An Exploratory Case Study of the Use of Video Digitizing Technology to Detect Answer-Copying on a Paper-and-Pencil Multiple-Choice Test. Artigo publicado Canadian Engineering Education Association, (CEEA13), 2013. **Disponível em:** <<http://ojs.library.queensu.ca/index.php/PCEEA/article/viewFile/4809/4756>> **Acesso em:** 28-05-2016.

ZOPLUOGLU, Cengiz. CopyDetect: An R Package for Computing Statistical Indices to Detect Answer Copying on Multiple-Choice Examinations, aplicado a medidas psicológicas. Artigo publicado por SAGE Journals, 2013. **Disponível em:** <<http://apm.sagepub.com/content/37/1/93>>. **Acesso em: 31-05-2016.**