UNIVERSIDADE ESTADUAL DE CAMPINAS

VANDERLI MARINO MELEM

# TEORIA DE RESPOSTA AO ITEM: UMA APLICAÇÃO DO MODELO DE CRÉDITO PARCIAL DE MASTERS

Campinas
1998

UNICAMP
BIBLIOTECA CENTRAL

TEORIA DE RESPOSTA AO ITEM:

UMA APLICAÇÃO DO MODELO DE CRÉDITO PARCIAL DE MASTERS

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por VANDERLI MARINO MELEM e aprovada pela comissão julgadora.

Campinas, 24 de agosto de 1998.

Profª. Drª. Eugênia M. Reginato Charnet
Orientadora

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial à obtenção do Título de MESTRE em ESTATÍSTICA.

**Dissertação de Mestrado defendida e aprovada em 24 de agosto de 1998**

**pela Banca Examinadora composta pelos Profs. Drs.**

**Prof (a). Dr (a). EUGÊNIA MARIA REGINATO CHARNET**

**Prof (a). Dr (a). CLARICE AZEVEDO DE LUNA FREIRE**

**Prof (a). Dr (a). JAMES PATRICK MAHER**

**Universidade Estadual de Campinas**
**Instituto de Matemática, Estatística e Computação Científica**
**Departamento de Estatística**

# TEORIA DE RESPOSTA AO ITEM: UMA APLICAÇÃO DO MODELO DE CRÉDITO PARCIAL DE MASTERS

**Autora: VANDERLI MARINO MELEM**
**Orientadora: Profa. Dra. EUGÊNIA MARIA REGINATO CHARNET**

**Campinas**
**1998**

# Índice

**RESUMO** ..... i

**CAPÍTULO 1**
**1-** Introdução ..... 1

**CAPÍTULO 2**
TEORIA DE RESPOSTA AO ITEM PARA ITENS DICOTÔMICOS ..... 9
**2.1-** Unidimensionalidade e Independência Condicional ..... 13
**2.2-** Parâmetros da TRI ..... 19
**2.2.1-** Modelo Logístico-Linear ..... 25

**CAPÍTULO 3**
TEORIA DE RESPOSTA AO ITEM PARA ITENS POLITÔMICOS ..... 31
**3.1-** Itens com Respostas Categóricas ..... 34
**3.2-** Modelo de Crédito Parcial (MCP)-Masters ..... 42
**3.3-** Modelo de Crédito Parcial Generalizado (MCPG)–Muraki. 53

**CAPÍTULO 4**
ESTIMAÇÃO DOS PARÂMETROS ..... 54
**4.1-** Estimação dos Parâmetros para o Caso de Itens Dicotômicos ..... 56
**4.1.1-** Estimação de Máxima Verossimilhança Conjunta ..... 61
**4.1.2-** Suposições ..... 64
**4.1.3-** Método de Newton-Raphson ..... 65
**4.2-** Métrica da Habilidade ($\theta$) ..... 72
**4.3-** EMV Conjunta dos parâmetros - Itens Categóricos ..... 75
**4.3.1-** EMV Conjunta dos parâmetros dos indivíduos ..... 82
**4.4-** EMV Marginal dos parâmetros sob o Modelo de Crédito Parcial ..... 85
**4.4.1-** Solução de Bock e Aitkin ..... 89
**4.4.1.1-** Algoritmo EM ..... 92

**CAPÍTULO 5**
FUNÇÃO DE INFORMAÇÃO ..... 99
**5.1-** Função de Informação do Item – Dicotômico ..... 101
**5.1.1-** Função de Informação do Teste ..... 104
**5.2-** Função de Informação do Item para Respostas Categorizadas ..... 109

**CAPITULO 6**
APLICAÇÃO: TEORIA DE RESPOSTA AO ITEM ..... 114
**6.1-** Análise Descritiva ..... 120
**6.2-** Estimação dos Parâmetros dos Itens sob o Modelo de Crédito Parcial de Masters ..... 123

**6.2.1** - Interpretação Gráfica dos Itens - MCPM ..... 129
**6.3**- Estimação dos Parâmetros dos Itens sob o Modelo de Crédito Parcial Generalizado de Muraki ..... 137
**6.3**- Estimação dos Parâmetros dos Itens sob o Modelo de Resposta Categórica de Samejima ..... 140

**APENDICE A** - Análise Fatorial Multivariada ..... 150

**APENDICE B** - Prova de Inglês do Vestibular/97 - UNICAMP 164

**APENDICE C** - Programas para uso do MULTILOG ..... 172

**REFERÊNCIAS BIBLIOGRÁFICAS** ..... 175

# RESUMO

Um dos problemas enfrentados por profissionais que corrigem um exame (ou teste) é a dificuldade em atribuir um escore a um indivíduo. É conhecido que dois indivíduos que obtém o mesmo escore em um teste podem não ter a mesma habilidade em relação a variável latente sendo medida pelo teste.

A Teoria de Resposta ao Item fornece modelos para estudar, entre outras coisas, as diferentes categorias de resposta envolvidas em um item e, a habilidade do indivíduo sendo medida pelo teste. Alguns dos modelos propostos são os dicotômicos (Rasch, 1980), politômicos (Samejima, 1969, 1972) e contínuos (Samejima, 1969). Neles se incluem parâmetros do item e do indivíduo que servem para medir fundamentalmente a dificuldade do item e o poder de discriminação do item e, a habilidade dos indivíduos em relação a variável latente sendo medida pelo teste.

Como a quantidade de parâmetros a serem estimados depende do número de itens e de indivíduos envolvidos no teste, é necessário utilizar métodos iterativos para sua estimação.

Uma aplicação dessa teoria foi realizada utilizando o modelo politômico de crédito parcial de Masters (Masters, 1982) na prova de inglês do Vestibular/97 da UNICAMP.

# AGRADECIMENTOS

Especialmente agradeço ao meu marido Claudio, pelo incentivo, carinho e compreensão nesses últimos anos em que estive envolvida com o mestrado.

A Deus agradeço pela oportunidade de conhecer pessoas maravilhosas como meus amigos, Alexandro, Carmen Sylvia, Fernando, Petty, Rose e Rui, desejo de coração que estas amizades sejam eternas.

Agradeço a minha orientadora e amiga Eugênia, pela orientação, dedicação e amizade, essenciais para a realização deste trabalho.

Agradeço a professora Matilde do Instituto de Estudos da Linguagem (IEL), pela colaboração no capítulo de aplicação e, a professora e amiga Clarice da Universidade Federal do Paraná pelo incentivo e carinho.

E por fim, dedico este trabalho e tudo de bom que aprendi a minha mãe Alice e a minha irmã Erci, desejo que estejam com Deus e felizes por mim.

Vanderli

# CAPÍTULO 1

## 1- INTRODUÇÃO

Um dos problemas enfrentados por profissionais que corrigem um exame ou por um professor que corrige uma prova (ou teste) é a dificuldade em atribuir um escore a um indivíduo da forma mais justa possível. Conhecemos que dois indivíduos que obtém o mesmo escore em um teste podem não ter o mesmo valor da variável, ou característica latente que está sendo medida pelo teste. O problema está em detectar esta diferença através da chamada Teoria Clássica de Testes (TCT). Esta variável, ou característica latente é uma variável não observável, como: inteligência, habilidade em matemática, leitura, etc. Alternativamente, uma teoria conhecida como Teoria de Resposta ao Item (TRI) vem para nos ajudar nesse sentido, uma vez que, através dela obtemos um parâmetro de discriminação e um parâmetro de dificuldade para cada item, os quais fornecem muita informação sobre o "comportamento" dos mesmos, possibilitando a identificação e uma possível solução de problemas. A TRI permite expressar as propriedades dos itens em termos de parâmetros invariantes com respeito aos indivíduos. Diferente da TCT onde os parâmetros dos itens variam de acordo com o grupo de indivíduos que faz o teste.

Estas duas teorias, utilizadas para análise de itens, são encontradas na literatura.

A Teoria de Resposta ao Item (TRI) é também conhecida como Teoria de Traço Latente ou Teoria de Curva Característica do Item. Os trabalhos de Richardson (1936), Lawley (1944), Tucker (1951), Birnbaum (1957,1958 e 1968) e Rasch (1960) são geralmente considerados como os precursores no desenvolvimento da TRI. No entanto, o real desenvolvimento desta teoria foi devido aos trabalhos de Lord (1952, 1953) e de Rasch (1980).

A Teoria de Resposta ao Item (TRI) foi também estudada por outros autores, dentre eles: Lord e Novick (1968), Bock e Lieberman (1970), Samejima (1972), Lord (1980), Andersen (1977), Cressie e Holland (1983), Molenaar (1983), Thissen e Steimberg (1984), Stocking e Pearlman (1989). Mais recentemente, vem sendo estudada por Baker (1992), Muraki (1992), Fischer (1995) e outros.

Autores como Lord e Novick (1968), Stocking e Pearlman (1989) enfocaram a análise de testes quando os itens são dicotômicos, ou seja, testes cujas respostas aos itens são classificadas apenas como certas ou erradas. Os testes de múltipla escolha são exemplos de testes com itens dicotômicos, onde se considera uma alternativa correta e as demais incorretas. Outros autores, em especial

Samejima (1972), se estendem ao caso de itens politômicos, ou seja, com mais de duas categorias de respostas por item, por exemplo - itens onde as respostas são classificadas como:

→ Fraco Regular Bom

→ Discordo Fortemente Discordo Concordo Concordo Fortemente

→ Liberal Moderado Conservador

A Teoria Clássica de Testes (TCT) está baseada principalmente nas correlações entre as notas de um novo teste e as de outro teste considerado padrão, ou nas correlações entre as notas de subtestes de um mesmo teste. Estatísticas importantes, como a fidedignidade e a validade de um teste são ferramentas utilizadas para avaliar um teste nessa teoria.

A **fidedignidade** de um teste avalia o grau com que o escore observado representa o escore verdadeiro. É definida como o quadrado da correlação entre o escore observado (X) e o escore verdadeiro (T), ou seja, $\rho^2(X,T)$ sendo portanto uma função da variável aleatória T, conhecida como variável latente. Como as variáveis latentes não são observáveis torna-se necessário a definição da fidedignidade como função de outras variáveis que sejam observáveis.

Na Teoria Clássica de Testes considera-se o escore observado X, como uma função linear do escore verdadeiro T, ou seja X=T+**E**, onde **E** é o erro aleatório e assume-se não correlacionado com a variável T.

Assim, pela definição do coeficiente de correlação de Pearson: $\rho^2(X,T)=\left(\frac{COV(X,T)}{\sqrt{\sigma^2(X).\sigma^2(T)}}\right)^2=\frac{\sigma^2(T)}{\sigma^2(X)}$, sendo $\sigma^2$(T) e $\sigma^2$(X) as variâncias das variáveis T e X, respectivamente.

Vejamos,

$$COV(X,T) = COV(T+\mathbf{E},T)$$

$$= E[(T+\mathbf{E})T]-E(T+\mathbf{E})E(T)$$

$$= E[T^2+\mathbf{E}T]-[E(T)+E(\mathbf{E})]E(T)$$

$$=E(T^2)+E(\mathbf{E}T)-E^2(T)-E(\mathbf{E})E(T)$$

$$= [E(T^2)-E^2(T)]+E(\mathbf{E}T)-E(\mathbf{E})E(T)$$

$$= \sigma^2(T)+COV(\mathbf{E},T)$$

Como **E** e T não são correlacionados, COV(**E**,T)=0,

$$COV(X,T)=\sigma^2(T)$$

Portanto,

$$\rho^2(X,T)=\frac{\sigma^2(T)}{\sigma^2(X)} \quad .\square$$

A **validade** de um teste, aplicado a um indivíduo selecionado ao acaso, é definida como a correlação entre os escores X e Y, onde X é o escore observado e, Y pode ser outras medidas de desempenho (critério),obtidas independentemente do primeiro teste. A validade é então ρ (X,Y).

A TRI descreve as respostas dos indivíduos a um item do teste através de uma função, denominada função de resposta ao item e, esta função define a probabilidade de acerto ao item por um indivíduo dotado de certa habilidade. Esta função compreende parâmetros referentes ao item.

Enquanto a Teoria Clássica de Testes avalia o teste como um todo, ou descreve o comportamento do grupo de indivíduos em relação ao teste que lhes foi aplicado, a Teoria de Resposta ao Item trabalha com parâmetros referentes aos itens e parâmetros referentes aos indivíduos, com a finalidade de avaliar o indivíduo ou ainda poder construir testes a partir de itens com parâmetros previamente conhecidos.

A vantagem da Teoria Clássica de Testes é a de não ser muito complexa, uma vez que a fidedignidade e a validade de um

teste são correlações entre notas obtidas nos testes. Entretanto, a Teoria de Resposta ao Item conduz a resultados mais ricos em informações. Um inconveniente é que na TRI o número de parâmetros a serem estimados é muito grande e a probabilidade de acerto ao item depende tanto da habilidade do indivíduo como da natureza do item.

A TRI começou a ser viabilizada com o desenvolvimento de modelos matemáticos devido a Birnbaum (1957), conhecidos como curvas logísticas. Mas, sua aplicação ainda era difícil pela complexidade da matemática envolvida na estimação dos parâmetros. A sua utilização realmente teve início nas últimas décadas com o desenvolvimento da informática com sofisticados softwares.

Hoje, a TRI tem sido incorporada em muitos programas de análise de itens, como: **LOGIST** (Wingersky e Lord, 1973), **BICAL** (Wright, Mead e Bell,1979), **BILOG** (Mislevy e Bock, 1984), **BILOG-MG** (Zimowski, Muraki, Mislevy e Bock, 1990), **MULTILOG** (Thissen,1991) e os programas: **LRSMG** (G. H. Fischer e P. Parzer, 1991); **LRSM** (G. H. Fischer e P. Parzer, 1991); **LLTM** (A K. Formann,1981, B. Wild,1989 e G. H. Fischer, 1973 a 1983); **LPCM** (P. Parzer, I. Ponocny e G. H. Fischer, 1994).

Um dos objetivos da TRI é estimar o valor da habilidade do indivíduo que é uma variável, ou característica latente,

para cada indivíduo que respondeu ao teste, independente do instrumento utilizado para medir esta variável. Se o modelo nesse contexto é adequado aos dados (respostas dos indivíduos aos itens do teste) e se existir unidimensionalidade, ou seja, se o teste contém itens que estejam medindo a mesma habilidade nos indivíduos, o objetivo acima se cumpre. Podemos nos auxiliar da Análise Fatorial Multivariada, que é explorada no apêndice A desse trabalho, para verificar a unidimensionalidade dos itens.

Na TRI os tipos de modelos podem variar:

→ com respeito à dimensão do espaço latente (modelos que permitem mais que uma variável ou característica latente, são complexos e menos desenvolvidos que os modelos unidimensionais);

→ com respeito aos níveis de respostas (número de categorias de resposta para cada item);

→ com respeito aos parâmetros (número de parâmetros de itens definidos);

→ com respeito as suposições estatísticas (forma da Curva Característica do Item, CCI).

Nos Capítulos 2 e 3 são apresentados alguns modelos da TRI para os casos de itens dicotômicos e politômicos, respectivamente. No Capítulo 4 são desenvolvidos procedimentos de Estimação por Máxima Verossimilhança como a Estimação de Máxima Verossimilhança Marginal(EMVM), algoritmo EM, etc. No Capítulo 5 estudamos a função de informação do item e a função de informação do teste.

Finalmente, no Capítulo 6, aplicamos os modelos estudados no Capítulo 3 para os dados da prova de inglês (segunda fase) do vestibular da UNICAMP, realizado em janeiro de 1997.

# CAPÍTULO 2

## TEORIA DE RESPOSTA AO ITEM PARA ITENS DICOTÔMICOS

Um conceito central da TRI é a Curva Característica do Item (CCI). Esta curva descreve a relação entre a probabilidade de um indivíduo acertar um a item, e os valores da variável, ou característica latente, quanto maior a habilidade do indivíduo maior a probabilidade dele acertar o item. Geralmente esta variável, ou característica latente, se refere à inteligência, habilidade em leitura, etc. Por isso passaremos a representá-la com o nome de "habilidade" e a denotá-la por $\theta$.

A falta de tal relação sugere que o item não será capaz de discriminar os indivíduos com maior nível de habilidade dos outros. Um teste composto por tais itens será então ineficiente em mensurar as diferenças em habilidade entre os indivíduos.

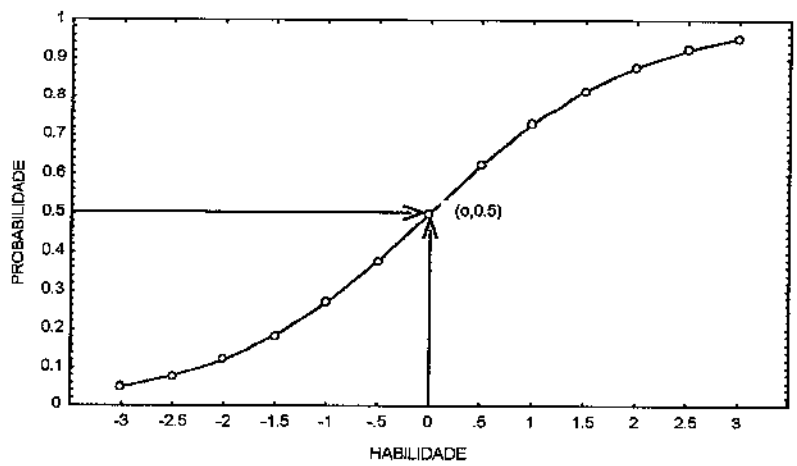



**FIGURA 2.1**- Curva Característica do Item.

Na Figura 2.1 acima podemos verificar que a CCI é uma função não decrescente em θ, ou seja, quanto maior a habilidade do indivíduo maior a probabilidade dele responder o item corretamente. A habilidade θ pode assumir valores de $-\infty$ a $+\infty$.

Então, $P_i(\theta)$ denota a probabilidade de que um indivíduo de habilidade θ responda corretamente o i-ésimo item do teste e, θ é a variável, ou característica latente mensurada pelo item.

Observamos que $P_i(\theta)$ é a probabilidade condicional de resposta correta ao i-ésimo item, dado θ e, é uma função não decrescente em θ. Desta forma, a CCI se assemelha a uma função logística, $f(W)=1/(1+e^{-W})$, cuja representação gráfica é idêntica à da Figura 2.1 com W=θ.

Vamos conceituar uma subpopulação de indivíduos para cada ponto na escala da habilidade. Essa subpopulação será

considerada homogênea com respeito a variável latente θ. Uma ilustração encontra-se na Figura 2.2 a seguir.

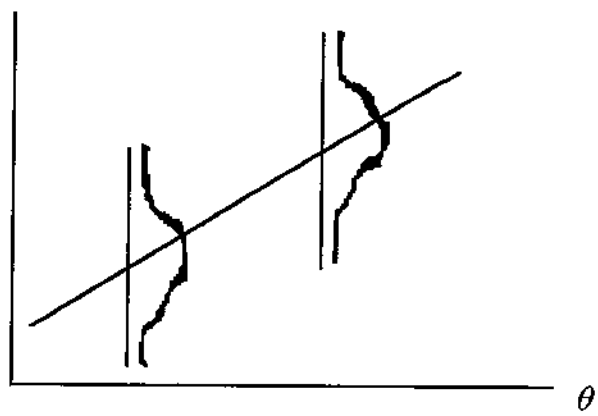



**FIGURA 2.2**- Duas subpopulação homogêneas.

A probabilidade de um indivíduo responder corretamente a um item é interpretada como a probabilidade que um indivíduo, escolhido aleatoriamente de uma subpopulação homogênea, responda corretamente a um item. Por exemplo: considere uma CCI que indica que para $\theta=\kappa$, (onde $\kappa$ é uma constante) a probabilidade de responder o item corretamente é 0,87. Isso pode ser interpretado como: a probabilidade de um indivíduo, escolhido aleatoriamente de uma subpopulação com $\theta=\kappa$, responder corretamente ao item é de 0,87.

Um teste capaz de mensurar uma habilidade de um grupo de indivíduos deve ser composto de itens que meçam esta única habilidade nos indivíduos. A esta propriedade do teste daremos o nome de UNIDIMENSIONALIDADE. Além disso, sempre que todos os

itens estiverem sendo tratados simultaneamente, a suposição de INDEPENDÊNCIA CONDICIONAL será necessária.

Nas próximas seções nos dedicaremos a estudar conceitos fundamentais da TRI.

## 2.1- UNIDIMENSIONALIDADE E INDEPENDÊNCIA CONDICIONAL

**Definição 2.1:** Um teste é dito unidimensional, se todos os itens que o compõem medem uma única variável latente (habilidade).

Especificamente, a exigência de que um teste seja unidimensional é a de que a dependência estatística entre os itens possa ser explicada por uma única variável, ou característica latente. Ou seja, se os itens são dependentes estatisticamente com respeito a uma habilidade que queremos medir, então todos os itens medirão a mesma habilidade e o teste é assim chamado unidimensional. Tem-se observado que a Análise Fatorial Multivariada é uma das técnicas mais apropriadas para (testar) comprovar a unidimensionalidade, mas não há um critério claro à partir do qual se possa afirmar a unidimensionalidade.

Na Análise Fatorial Multivariada descartam-se os itens que formam fatores periféricos, fazendo-se o mesmo nas análises posteriores até obter (se existir) uma análise em que um fator explica a maior parte da variação dos itens. Este processo de polimento de itens para obter unidimensionalidade tem sentido quando já na primeira análise surge uma unidimensionalidade aproximada, caso contrário, se converte em um processo artificial para descartar itens,

que conduzem necessariamente a um teste unidimensional mas pouco a ver com os objetivos iniciais. (Ver apêndice A - Análise Fatorial Multivariada).

Um outro conceito importante nessa teoria é a de independência condicional, também chamada de independência local, foi introduzido no trabalho de Lord e Novick (1968) como a seguir.

**Definição 2.2:** Dizemos que, em um teste existe independência condicional se os itens que o compõem são estatisticamente independentes para alguma subpopulação homogênea, com respeito a uma variável latente (habilidade).

Dessa definição temos que a suposição de independência condicional requer que a probabilidade de um indivíduo obter um particular padrão de corretas e incorretas respostas seja igual ao produto das probabilidades correspondentes a cada resposta individual. Esta interpretação mais estatística do que é independência condicional deve-se a Hambleton e Swaminathan (1985).

Embora comumente seja usado o termo independência local, sempre usaremos nesse trabalho o termo independência condicional que é considerado mais descritivo e apropriado, particularmente porque o termo independência local é freqüentemente usado para descrever independência que é local em tempo e espaço.

Da Definição 2.2, independência condicional significa que dentro de alguma subpopulação homogênea de indivíduos, nos quais as variáveis latentes (habilidades) $\theta_1,\theta_2,...,\theta_N$ assumem os mesmos valores, as distribuições condicionais dos escores dos itens são todas mutuamente independentes. Isto não sugere que os escores dos itens são não relacionados uns com os outros para o grupo total de indivíduos, não existindo unidimensionalidade. Significa que os escores dos itens são relacionados uns com os outros somente através das variáveis latentes (habilidades) $\theta_1,\theta_2,...,\theta_N$, onde N é o número de indivíduos.

Se a independência condicional não se verifica, a probabilidade de algum indivíduo responder corretamente a um grupo de itens é, digamos, maior que o produto das probabilidades individuais para cada resposta correta. Podemos esperar então que esses indivíduos obtenham escores maiores que outros de mesmo nível de habilidade. Isto nos indicará que o teste mede mais de uma variável latente (habilidade), ou seja, não existe unidimensionalidade.

Se um determinado teste for unidimensional, então a suposição de independência condicional será necessariamente satisfeita (Hambleton, 1979).

Considerando um teste com L itens e, se estes itens medem uma única habilidade nos indivíduos, pela independência condicional, fixados os valores de $\theta_1,\theta_2,\ldots,\theta_N$, a distribuição conjunta dos escores dos itens ($x_i$) é igual ao produto das distribuições marginais $f_i(x_i|\theta)$, ou seja,

$$f(x_1, x_2,\ldots,x_L|\theta) = \prod_{i=1}^{L} f_i(x_i|\theta). \qquad (2.1)$$

Considerando primeiro os itens 2,3,...,L, a distribuição conjunta para estes itens será

$$f(x_2, x_3,\ldots,x_L|\theta)=\prod_{i=2}^{L} f_i(x_i|\theta) \qquad (2.2)$$

Dividindo (2.1) por (2.2), temos:

$$\frac{f(x_1,x_3,\ldots,x_L|\theta)}{f(x_2,x_2,\ldots,x_L|\theta)} = f_1(x_1|\theta)$$

$$h_1(x_i|\theta,x_2,x_3,\ldots,x_L) = f_1(x_1|\theta)$$

Isto mostra que sob a hipótese de independência condicional a distribuição condicional $h_1$ de $x_1$, para valores fixos de ($\theta$, $x_2$, $x_3$,...,$x_L$) não depende de ($x_2$, $x_3$,...,$x_L$), ou seja, um indivíduo acertar o i-ésimo item independe dele acertar ou não o i'-ésimo item, para i≠i'=1,2,3...,L. Repetindo este procedimento, provamos que a

distribuição condicional $h_i$ de $x_i$, para valores fixos de $(\theta, x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_L)$ não depende de $(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_L)$. □

Para itens dicotômicos a suposição de independência condicional pode ser escrita como:

$$\text{Prob}\,(X_1=x_1,\ X_2=x_2,\ \ldots,\ X_L=x_L|\theta) = \prod_{i=1}^{L} \text{Prob}(X_i = x_i|\theta), \qquad (2.3)$$

onde $x_i$ assume valor zero ou um.

A função de probabilidade de um escore de item dicotômico para θ fixo pode ser escrita como:

$$f_i(x_i|\theta) \cong P_i^{x_i} Q_i^{1-x_i}, \qquad (2.4)$$

onde $Q_i = 1 - P_i$ sendo: $P_i(\theta)=P_i$, $Q_i(\theta)=Q_i$ e $f_i(x_i|\theta) \cong \begin{cases} P_i & se \quad x_i = 1 \\ Q_i & se \quad x_i = 0 \end{cases}$

As respostas aos itens dadas pelo indivíduo de habilidade θ podem ser denotadas pelo vetor de variáveis aleatórias $\mathbf{X} \cong (X_1, X_2, \ldots, X_L)$. Assim por (2.3) e (2.4) a distribuição condicional de **X** dado θ é

$$P(\mathbf{X}|\theta) \cong \prod_{i=1}^{L} P_i^{x_i} Q_i^{1-x_i}. \qquad (2.5)$$

Se para alguma população de indivíduos a habilidade $\theta$ tem uma função de densidade de probabilidade denotada por $g(\theta)$ então $P(\mathbf{X}=\mathbf{x})$, a distribuição (não condicional) de $\mathbf{X}$ para o grupo total de indivíduos, é dada por:

$$P(\mathbf{X}=\mathbf{x})=\int\prod_{i=1}^{L}P_i(\theta)^{x_i}\left(1-P_i(\theta)\right)^{1-x_i}dF(\theta), \qquad (2.6)$$

onde $F(\theta)$ é uma função de distribuição acumulada; $P_i(\theta)$ é uma função monótona não decrescente em $\theta$ com $0\leq P_i(\theta)\leq 1$ para todo $\theta$, e a integração se dá sobre o conjunto n-dimensional correspondente aos elementos de $\theta$.

A equação (2.6) mostra que as respostas aos itens $\mathbf{X}=(X_1,X_2,\ldots,X_L)$ são condicionalmente independentes dado uma variável latente $\theta$.

## 2.2- PARÂMETROS DA TRI

A TRI nos possibilita analisar individualmente os itens de um teste nos fornecendo parâmetros referentes aos itens e parâmetros referentes aos indivíduos. Bem interpretados, esses parâmetros fornecem uma grande quantidade de informação, necessária para a análise dos itens, identificando itens por sua dificuldade e por seu poder de discriminação entre os indivíduos de maior ou menor nível de habilidade. Existirá um parâmetro, denotado por $\theta$, para cada indivíduo, que o classificará com respeito à variável, ou característica latente mensurada pelo teste. Como já sabemos, aqui esta variável é chamada de habilidade do indivíduo. Em geral, existirão três parâmetros referentes a cada item que serão denotados por $\alpha$, $\beta$ e $\psi$, e estão localizados na CCI exemplificada na Figura 2.3 a seguir:

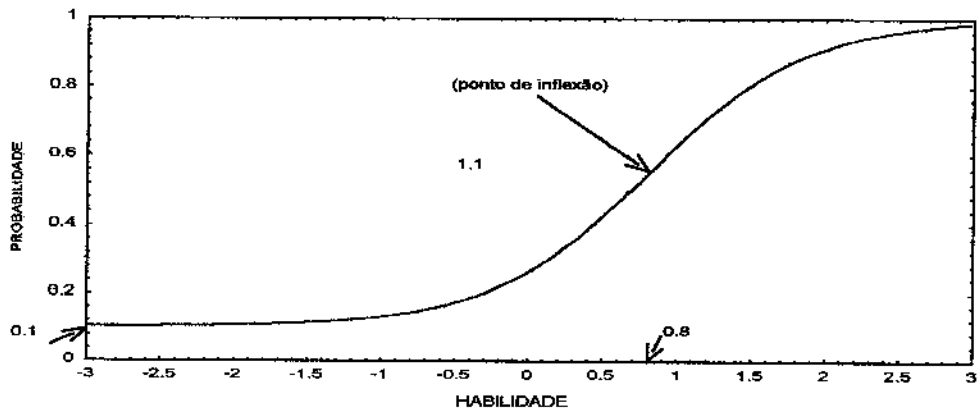



**FIGURA 2.3-** Relação entre a probabilidade de resposta correta e a habilidade dos indivíduos, fixados os valores dos parâmetros do item em: $\alpha$=1.1, $\beta$=0.8 e $\psi$=0.1.

Os parâmetros de discriminação e de dificuldade, denotados por $\alpha$ e $\beta$, respectivamente, se relacionam com a Curva Característica do Item como segue: A CCI tem um único ponto de inflexão que ocorre em $\theta=\beta$ e, o coeficiente angular da reta tangente à curva no ponto $\theta=\beta$, é proporcional ao valor de $\alpha$. O parâmetro $\beta$ é um parâmetro de locação que determina a posição na escala da habilidade $\theta$ onde ocorre o ponto de inflexão.

O parâmetro $\alpha$ é proporcional à inclinação da curva no ponto de inflexão $\theta=\beta$ e representa o poder de discriminação do item. Quanto maior seu valor absoluto, maior o poder de discriminação do item. Se ocorrer $\alpha$ negativo significará que o item correspondente é pontuado em direção oposta aos outros itens.

O parâmetro $\psi$, chamado de parâmetro de adivinhação (ou acerto casual), é a probabilidade mínima de um indivíduo responder corretamente o item, ou seja, um indivíduo que não tenha nenhum conhecimento responderá o i-ésimo item corretamente com probabilidade $\psi > 0$.

Seja a função $P_i(\theta)$ da CCI a função de distribuição de probabilidade de uma distribuição normal padrão,

$$P_i(\theta) = \int_{-\infty}^{\omega} f(z)dz,$$

onde $\omega$ é um número real que satisfaz, $\omega = \alpha(\theta - \beta)$. Dessa forma, $P_i(\theta)$ é a probabilidade de que indivíduos com a mesma habilidade $\theta$ respondam o item i corretamente.

Por exemplo, suponhamos um certo item i com parâmetro de discriminação $\alpha = 0.5$, parâmetro de dificuldade $\beta = 1.0$, e um indivíduo com habilidade $\theta = 2$. Temos então que $\omega = 0.5$ e portanto, a probabilidade de indivíduos com habilidade $\theta = 2$ responderem corretamente ao i-ésimo item é de 69%.

Os parâmetros $\alpha$, $\beta$ e $\psi$ variam com os itens e por isso serão indexados como $\alpha_i$, $\beta_i$ e $\psi_i$. O parâmetro $\psi_i$, interpretado como parâmetro de adivinhação, quando não for utilizado no modelo

significará que estaremos assumindo que, um indivíduo com nenhum conhecimento tem probabilidade zero de acertar o item, ou seja, não existe a possibilidade de adivinhar a resposta correta.

Assim, considerando $\psi_i=0$, podemos escrever

$$P_i(\theta)=\int_{-\infty}^{\alpha_i(\theta-\beta_i)} f(z)dz,$$

para i=1,2,...,L.

Exemplificaremos a seguir –Figura 2.4- a forma das curvas características de três itens, com o mesmo parâmetro de discriminação $\alpha_i=1.1$ e diferentes valores do parâmetro de dificuldade, sejam eles: $\beta_1=0.5$, $\beta_2=1.0$ e $\beta_3=1.5$.

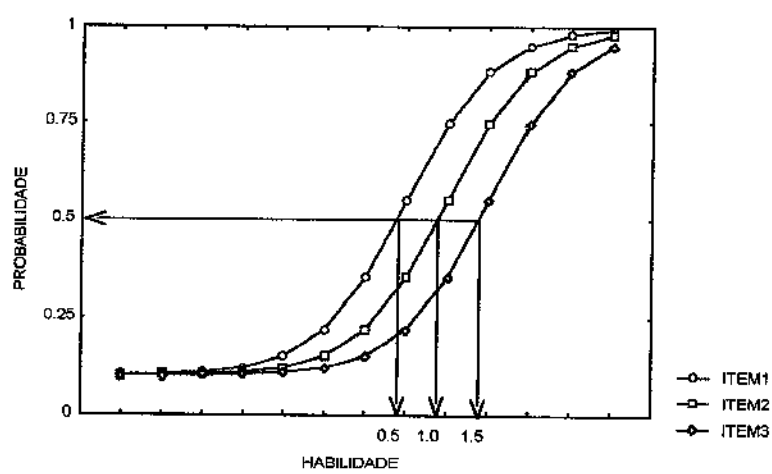



**FIGURA 2.4-** Curvas Características de Itens para os diferentes valores de $\beta_i$.

Observa-se que $P_i(\theta)=0.5$ corresponde a, $\theta=\beta_1=0.5$ no item 1, $\theta=\beta_2=1$ no item 2 e $\theta=\beta_3=1.5$ no item 3. Assim, $\beta_i$ é igual ao escore da habilidade para o qual metade dos indivíduos respondem o i-ésimo item corretamente. Podemos observar na Figura 2.4 que, para um indivíduo ter 50% de chance de responder o 3° item corretamente, ele precisa de um escore de habilidade ($\theta$) maior do que para ter a mesma chance de responder o 1° item corretamente. Logo o primeiro item é o mais fácil pois requer uma habilidade menor para o indivíduo respondê-lo corretamente.

Suponhamos agora três itens com o mesmo parâmetro de dificuldade $\beta=1.5$ e diferentes valores do parâmetro de discriminação, sejam eles: $\alpha_1=0.1$, $\alpha_2=1.0$ e $\alpha_3=100$.

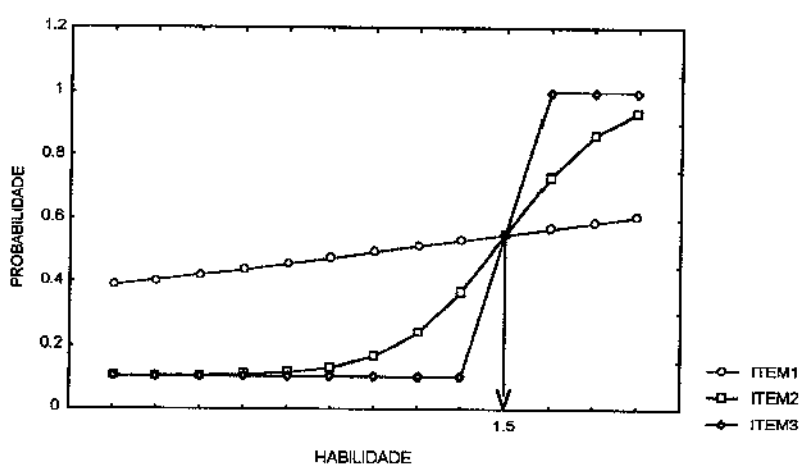



**FIGURA 2.5-** Curvas Características de Itens para os diferentes valores de $\alpha_i$.

Nessa Figura 2.5 verificamos que, o primeiro item não discrimina bem porque a probabilidade de responder ao item corretamente é aproximadamente a mesma para todos os níveis de habilidade. No entanto o terceiro item, com $\alpha_3$=100, discrimina muito bem entre os indivíduos com escores menores que 1,45 daqueles com escores maiores que 1,55. De fato, para indivíduos com $\theta \leq 1,45$ a probabilidade de resposta correta para o terceiro item é menor que 0,01, enquanto que para indivíduos com $\theta \geq 1,55$ a probabilidade de resposta correta para o terceiro item é maior que 0,99. O item 2 é intermediário em termos de discriminação.

## 2.2.1- MODELO LOGÍSTICO-LINEAR

Vimos que a CCI assume a forma de um S, sendo que a probabilidade de acerto a um item, $P_i(\theta)$, é próxima de zero para certos valores baixos da habilidade $\theta$ e, próxima de um para altos valores da habilidade $\theta$, e cresce suavemente na parte intermediária.

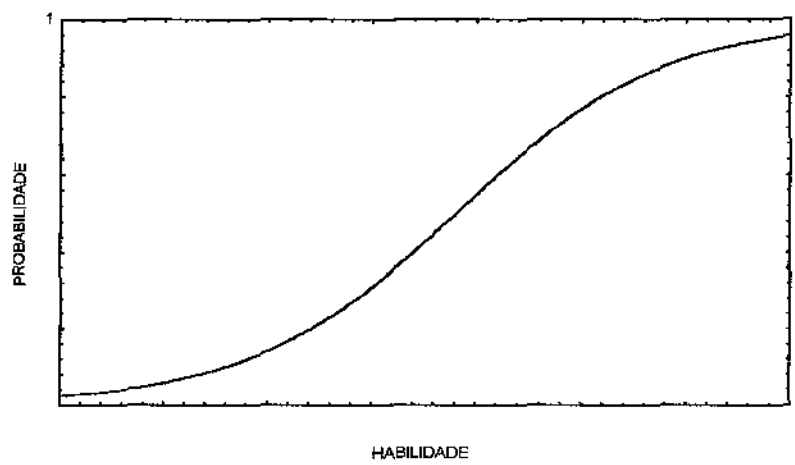



**FIGURA 2.6**- Função de Distribuição Acumulada.

Embora um grande número de famílias de funções de dois parâmetros possam ser usadas, somente as distribuições acumuladas das funções de distribuição normal ou logística serão usadas como modelo matemático para a CCI. A distribuição logística é a mais utilizada pela sua simplicidade matemática. O uso da ogiva logística como um modelo da CCI é devido a Birnbaum (1957).

Considerando um item i e o parâmetro de habilidade do indivíduo $(\theta_j)$, uma fácil justificativa do uso das distribuições açumuladas das funções de distribuição normal e logística é dada a seguir.

A função de distribuição acumulada (f.d.a.) da normal é dada por,

$$P_i(\theta_j) \cong \phi(Z) = \int_{-Z_{ji}=-\frac{(\theta_j-\mu_i)}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}} dZ,$$

onde $\mu_i$ é a média, $\sigma_i$ o desvio padrão, $\theta_j$ a habilidade e $Z_{ji}$ o desvio normal.

Agora essa f.d.a. é nossa equação da CCI. Como $\beta_i$ é definido como o ponto na escala da habilidade em que a probabilidade de resposta correta é 0.5 (veja Figura 2.4), assim $\beta_i = \mu_i$. As duas distribuições tem um ponto de inflexão em $\beta_i = \mu_i$ e o seu valor numérico tem o mesmo significado em ambos os modelos.

O parâmetro $\alpha_i$ é usado como um índice do declive da CCI e, seu valor numérico produz diferentes declives sob os dois modelos, pois as variâncias das funções de distribuições normal e logística são diferentes.

O desvio padrão mede a extensão da distribuição normal e: → se $\sigma_i$ é grande, a ogiva normal não terá muito declive em torno de $\mu_i = \beta_i$; → se $\sigma_i$ é pequeno, a ogiva normal terá mais declive em torno de $\mu_i = \beta_i$. Dessa forma, o parâmetro de discriminação é inversamente proporcional ao desvio padrão, $\alpha_i = \frac{1}{\sigma_i}$. Logo,

$$Z_{ji} = \frac{\left(\theta_j - \mu_i\right)}{\sigma_i} = \alpha_i\left(\theta_j - \beta_i\right). \qquad (2.7)$$

O modelo logístico-linear considera que ($x_1$, $x_2,\ldots,x_L$) são observações dicotômicas independentes, caracterizadas por duas possibilidades de resposta: $P(x_i=1)=P_i(\theta)$, probabilidade de resposta correta; $P(x_i=0)=1-P_i(\theta)=Q_i(\theta)$, probabilidade de resposta incorreta. A dependência de $P_i(\theta)$ sobre a variável latente segue a forma logística $P_i(\theta) = 1/\left(1+e^{-Z_{ji}}\right)$, onde $Z_{ji} = \log\left\{P_i(\theta)/\left[1-P_i(\theta)\right]\right\} = \log\left[P_i(\theta)/Q_i(\theta)\right]$, é chamado transformação logística da probabilidade $P_i(\theta)$ ou simplesmente "log odds".

Os modelos logísticos que serão apresentados diferem dependendo do número de parâmetros considerados.

**Modelo de 2 parâmetros**: é conveniente começar com o modelo logístico de 2 parâmetros (originalmente obtido por Birnbaum (1958), por sua semelhança com o modelo normal. Nesse

caso $Z_{ji}=D\alpha_i(\theta_j-\beta_i)$, onde D é uma constante que pode ser arbitrariamente estabelecida. É comum usar D=1.7 porque assim $P_i(\theta)$ para as ogivas normal e logística não diferirá em mais de 0.01, para algum valor de $\theta$, Lord e Novick (1968). O modelo neste caso é

$$P_i(\theta) = \frac{e^{D\alpha_i(\theta_j-\beta_i)}}{1+e^{D\alpha_i(\theta_j-\beta_i)}}. \qquad (2.8)$$

Podem ser também determinados outros três modelos de 1,2, e 3 parâmetros usando os modelos de ogiva normal, os quais assumem que a CCI é expressa por uma função da curva normal acumulada. Porém, segundo Lord (1980), não há razões substancialmente sólidas para se eleger a priori um tipo ou outro de modelo (logístico/normal) e, a nível prático, os resultados são muito similares. Pela simplicidade matemática, os modelos logísticos são preferíveis.

Nota-se na literatura que até Lord (1952), pioneiro em modelos de ogiva normal, se converteu aos modelos logísticos, Lord (1968 e 1980). Parece que isto se deve às influências de Birnbaum, como sugerem Wright e Stone (1979).

**Modelo de 1 Parâmetro:** é um caso especial do modelo logístico com 2 parâmetros onde todos os itens tem o mesmo

parâmetro de discriminação ($\alpha_i = \alpha$, i=1,2,...,L), em geral igual a 1 (um). A equação para a CCI uniparamétrica é

$$P_i(\theta) = \frac{e^{D\alpha(\theta_j - \beta_i)}}{1 + e^{D\alpha(\theta_j - \beta_i)}}. \qquad (2.9)$$

Nesse caso a proporção de indivíduos que respondem corretamente ao item i, é função da habilidade do indivíduo (θ) e da dificuldade do item $(\beta)$. Esse modelo é o chamado **Modelo de RASCH**, por ter sido formulado originalmente por RASCH (1960), recebendo posteriormente notável atenção, especialmente na Universidade de Chicago por Wrigth. (Wrigth, 1977 a e b; Wrigth e Stone, 1979).

Considerando uma extensão do modelo de Rasch para itens politômicos nos deparamos com o Modelo de Crédito Parcial (MCP), Masters(1982), que será estudado com mais detalhes no Capítulo 3.

**Modelo de 3 Parâmetros:** em um teste de múltipla escolha existe a possibilidade de um indivíduo com habilidade próxima de zero ter a chance de adivinhar a resposta correta e acertar o item.

Já vimos que $\psi_i$ é o parâmetro de adivinhação e, quando incluído no modelo obtém-se o modelo logístico com 3 parâmetros, que é dado por:

$$P_i(\theta) = \psi_i + \frac{(1-\psi_i)\, e^{Da_i(\theta_j-\beta_i)}}{1+e^{Da_i(\theta_j-\beta_i)}}, \qquad (2.10)$$

onde $\psi_i \in [0,1)$.

# CAPÍTULO 3

# TEORIA DE RESPOSTA AO ITEM PARA ITENS POLITÔMICOS

Podemos encontrar, na prática, várias formas de respostas aos itens de um teste. Alem disso, estes itens podem ser classificados, de acordo com sua variável resposta, como **Discreta** ou **Contínua**. Se a variável resposta de um item é considerada como discreta encontramos três formas importantes de resposta discreta, a saber: forma de resposta dicotômica (já discutida no Capítulo 2), forma de resposta nominal ((a) a seguir) e forma de resposta categórica (seção 3.1).

As formas de resposta nominal e categórica podem conter mais de duas categorias e, por isso são denominadas respostas multicategóricas. Os modelos para respostas multicategóricas também são encontrados na literatura nos trabalhos de Samejima (1969), Bock (1972), Masters (1982), Thissen e Steinberg (1984) e outros.

O teste de múltipla escolha, por exemplo, pode ser considerado um caso de itens com respostas multicategóricas, quando cada uma das alternativas recebe certa ponderação segundo sua pertinência, ao invés de serem consideradas apenas como respostas

dicotômicas (certo/errado). Chamaremos de itens politômicos aqueles itens com respostas multicategóricas.

a) Itens com Resposta Nominal: nesse caso as respostas para um item são classificadas em m categorias mutuamente exclusivas, exaustivas e não ordenadas.

Por exemplo, no caso (a) o indivíduo pode escolher livremente uma das m categorias de resposta em um item i, sendo estas respostas nominais e não ordenadas, não sendo acumulado um escore total, uma vez que medidas nominais não podem ser somadas. Para esta situação Bock (1972) sugere uma versão modificada do Modelo de Resposta Categórica de Samejima (1969), denominado Modelo Nominal de Bock (MNB). Essa teoria não será desenvolvida nesse trabalho onde a ênfase será dada aos Modelos de Itens com Resposta Categórica.

Thissen e Steinberg (1984) sugerem um modelo para itens de múltipla escolha como uma extensão do modelo proposto por Bock (1972) e subseqüentemente estendido por Samejima (1979) que adicionou uma categoria de resposta latente, rotulada como "zero" ou "don't know"(DK).

Itens contínuos são aqueles em que o indivíduo estima certo nível em uma escala. Samejima (1972) apresenta um

modelo que trabalha com situações onde as variáveis do item são contínuas. Ela descreve duas situações de nível de resposta contínua: Nível de Resposta Contínua Como uma Situação Limite do Nível de Resposta Graduada e; Situações de Resposta Aberta e Fechada. Na literatura encontramos poucos trabalhos sobre itens com resposta contínua, relativamente a itens com resposta discreta, sendo a grande maioria desses trabalhos atribuídos a Samejima.

É possível o uso de um formato dicotômico dentro dos itens multicategóricos e contínuos, estabelecendo-se para isso um ponto de corte. Com esse procedimento muitas vezes a perda de informação pode reverter em ganho em operacionalidade. Por outro lado, no Capítulo 5 veremos que a quantidade de informação do item aumenta quando usamos mais de duas categorias de resposta, visto que cada categoria contribui individualmente para a quantidade de informação do item.

## 3.1- ITENS COM RESPOSTAS CATEGÓRICAS

Existem algumas formas de respostas categóricas em que são associados modelos específicos da TRI para as Curvas Características das Categorias de Respostas dos Itens (CCCRI), assim chamadas por produzirem uma curva característica para cada categoria de resposta do item.

Masters (1982) classifica quatro tipos de categorias ordenadas de resposta: "Repeated Trials"(Medidas Repetidas),"Counts"(Contagem), "Rating Scales"(Classificação de Escalas) e "Partial Credit"(Crédito Parcial). Este último tipo, no qual o indivíduo recebe crédito parcial pelo seu resultado em cada item, será o objetivo do nosso estudo na seção 3.2.

É comum a utilização de itens onde as alternativas de respostas são elaboradas numa certa ordem, como por exemplo: "discordo fortemente", "discordo", "indiferente", "concordo" e "concordo fortemente". Em outro formato, o indivíduo responde livremente às questões e depois estas respostas são ordenadas em $m_i$ categorias de resposta, sendo $k=1,2\ldots, m_i$, onde $m_i$ é o número de categorias de resposta do i-ésimo item. A resposta de um indivíduo a

um item deve pertencer a somente uma das $m_i$ categorias e, $x_{jik}=1$ se a resposta é dada para a categoria k do item i pelo indivíduo de habilidade $\theta_j$, e $x_{jik}=0$ caso contrário. É importante verificar que, para cada nível de habilidade $\theta_j$, temos $\sum_{k=1}^{m_i} P_{ik}(\theta_j)=1$.

$P_{ik}(\theta_j)$ é a probabilidade de um indivíduo com habilidade $\theta_j$ responder na a k-ésima categoria de um item i, dada pela área entre os limites $\lambda_s$ e $\lambda_{s-1}$. Sendo $\lambda$ os limites entre as categorias e $s=1,2,\ldots,m_i-1$.

Considerando que as categorias de um item são ordenadas, podemos representá-las em relação aos limites $\lambda_s$, como segue:

*CATEGORIAS*

1 2 3 4
↑ ↑ ↑
$\lambda_1$ $\lambda_2$ $\lambda_3$

logo, temos $m_i=4$ categorias e portanto 3 limites entre as categorias. É importante lembrar que nessa situação as categorias de respostas são sempre ordenadas.

O gráfico de $P_{ik}(\theta_j)$ para $k=1,2,\ldots,m_i$, em função de $\theta$, produzirá uma CCCRI para cada categoria de resposta. Para

exemplificar graficamente estas curvas características, vamos considerar os valores dos parâmetros de um item com cinco categorias, dado no exemplo número oito do manual do programa MULTILOG (Thissen, 1991), para o modelo de resposta categórica de Samejima.

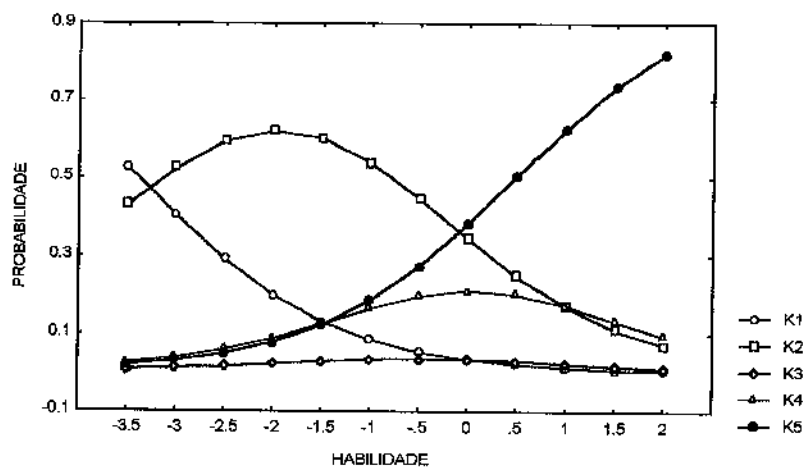



**FIGURA 3.1-** Cinco CCCRI, onde os limites entre as categorias são estimados por $\lambda_1$=-3.4, $\lambda_2$=-0.51, $\lambda_3$=-0.37, $\lambda_4$=0.48 e $\alpha$=1.

Verificamos na Figura 3.1 que os limites $\lambda_1$ e $\lambda_4$ são os pontos onde $P_{i1}(\theta_j)=0.5$ e $P_{i5}(\theta_j)=0.5$, respectivamente. Quando temos três ou mais categorias de resposta, somente o primeiro e o último limite são facilmente localizados em gráficos como esse. Segundo Masters(1982,p.155), para mais de três categorias não existe uma relação simples entre as CCCRI e os limites entre as categorias.

A falta de uma forma consistente para as CCCRI de todas as categorias de respostas de um item apresenta problemas para

modelar estas curvas e estimar os parâmetros do item. Para contornar esse problema usaremos um mecanismo simples com os limites $\lambda_s$ e essas curvas serão chamadas de Curvas Características Limites(CCL).

Nesse caso, definimos a probabilidade de uma "resposta correta" como a probabilidade de responder em uma categoria maior do que um dado limite $\lambda_s$, onde $s=1,2,\ldots,m_i-1$, aqui não consideramos $s=0$ porque a categoria zero não existe. A probabilidade de responder em uma categoria menor que $\lambda_s$ é considerada como a probabilidade de uma resposta incorreta.

Podemos perceber que, dessa forma, estamos trabalhando com probabilidades acumuladas e denotaremos como $P^*_{ik}(\theta_j)$, a probabilidade de responder em uma categoria maior que o limite $\lambda_s$, sendo $s=k-1$, ou seja, de responder em na categoria k ou maior. Continuando com o exemplo da Figura 3.1, para $m_i=5$ categorias teremos 4 limites entre as categorias. Assim,

$$\rightarrow P^*_{i1}(\theta_j)=P_{i1}(\theta_j)+\ldots+P_{i5}(\theta_j)=1,$$

então escolher uma categoria maior que o limite $\lambda_0$ (aqui consideramos s=o para exemplificar) tem probabilidade acumulada igual a 1;

$$\rightarrow P^*_{i2}(\theta_j)=P_{i2}(\theta_j)+\ldots+P_{i5}(\theta_j),$$

a probabilidade de escolher uma categoria maior que o limite $\lambda_1$ é a probabilidade de escolher a categoria 2, 3, 4 ou 5;

$$\rightarrow P^*_{i3}(\theta_j)=P_{i3}(\theta_j)+\ldots+P_{i5}(\theta_j),$$

a probabilidade de escolher uma categoria maior que o limite $\lambda_2$, ou seja, é a probabilidade de escolher a categoria 3, 4 ou 5;

$$\rightarrow P^*_{i4}(\theta_j)=P_{i4}(\theta_j)+P_{i5}(\theta_j),$$

a probabilidade de escolher uma categoria maior que o limite $\lambda_3$, ou seja, é a probabilidade de escolher a categoria 4 ou 5;

Finalmente,

$$\rightarrow P^*_{i5}(\theta_j)=P_{i5}(\theta_j),$$

a probabilidade de escolher uma categoria maior que o limite $\lambda_4$ é a probabilidade de escolher a categoria 5.

Generalizando, $P^*_{ik}(\theta_j) = \sum_{t=k}^{m} P_{it}(\theta_j)$.

Dessa forma a probabilidade de uma resposta pertencer a uma dada categoria k de um item i dada pelo indivíduo j, $P_{ik}(\theta_j)$, é definida pela diferença entre sucessivas probabilidades

acumuladas $P^*_{ik}(\theta_j)$. Para simplificar a notação excluiremos os índices j e i.

Logo,

$$P_k(\theta) = P^*_k(\theta) - P^*_{k+1}(\theta) \qquad (3.1)$$

A probabilidade de um indivíduo j responder ao item i na categoria k ou maior, é dada por

$$P^*_{ik}(\theta_j) = \frac{\exp\left[\alpha_i\left(\theta_j - \lambda_{is}\right)\right]}{1+\exp\left[\alpha_i\left(\theta_j - \lambda_{is}\right)\right]}, \qquad (3.2)$$

onde $\lambda_{is}$ é o limite entre as categorias k-1 e k associadas ao item i e s=k-1.

Assim, de (3.1) e (3.2), obtemos o Modelo de Resposta Categórica:

$$P_{ik}(\theta_j) = P^*_{ik}\left(\theta_j\right) - P^*_{i,k+1}\left(\theta_j\right) = \frac{1}{1+\exp\left[-\alpha_i\left(\theta_j - \lambda_{is}\right)\right]} - \frac{1}{1+\exp\left[-\alpha_i\left(\theta_j - \lambda_{i,s+1}\right)\right]} \qquad (3.3)$$

O modelo (3.3) acima é chamado de **Modelo de Resposta Categórica de Samejima (MRCS)** (1969,1972) e, quando aplicado possibilita que sejam estimadas: uma habilidade $\theta_j$ para cada indivíduo, um parâmetro de discriminação $\alpha_i$ e um conjunto de parâmetros limites $\lambda_{i1}, \lambda_{i2}, \ldots, \lambda_{i,m_i-1}$ para cada item.

No Modelo de Resposta Categórica de Samejima (MRCS), os limites entre as categorias, $\lambda_s$, são os parâmetros de locação das CCL.

Esses parâmetros são ordenados da forma $\lambda_{m_i-1}>\ldots>\lambda_s>\ldots>\lambda_1$, onde os $-\infty<\lambda_S<\infty$, devem ser monotonamente decrescentes mas não é necessário que sejam igualmente espaçados.

Observamos que quando estimamos os parâmetros do item, obtemos $m_i$-1 parâmetros de locação ($\lambda_{is}$, onde s=1,2,...,$m_i$-1) associados aos limites, pois temos ($m_i$-1) limites entre as categorias. Na realidade desejamos estimar $m_i$ parâmetros de locação, um para cada categoria de resposta. Este problema, que ocorre devido a restrição $\sum_{k=1}^{m_i} P_{ik}(\theta_j)=1$, pode ser contornado mostrando que o ponto modal de uma curva característica de categorias de resposta ao item é dado por $\lambda_k = (\lambda_{S+1}+\lambda_S)/2$ exceto para k=1 e k=m. Assim,

$$\lambda_1 = \lambda_1$$

$$\lambda_2 = (\lambda_2 + \lambda_1)/2$$

$$\lambda_3 = (\lambda_3 + \lambda_2)/2$$

$$\vdots$$

$$\lambda_k = (\lambda_{S+1} + \lambda_S)/2$$

$$\vdots$$

$$\lambda_m = \lambda_{m-1}$$

Os parâmetros $\lambda_k$ podem ser usados para descrever a locação de uma categoria de resposta ao item sob a escala de habilidade $\theta$ de maneira análoga à utilizada nos casos de itens com respostas dicotômicas. As CCL têm um valor comum para o parâmetro de discriminação $\alpha_i$, assim admite-se um mesmo parâmetro de discriminação para todas as categorias de resposta dentro de cada item.

## 3.2- MODELO DE CRÉDITO PARCIAL (MCP) - MASTERS

O MCP de Masters (1982) é uma extensão do modelo dicotômico de Rasch (1960) tendo mais de duas categorias de resposta por item, assim como, o MRC de Samejima (visto na seção 3.1) é uma extensão do modelo logístico de dois parâmetros de Birnbaum (1958). O MCP será estudado aqui, porque sua forma de resposta é utilizada nas questões do vestibular da UNICAMP, onde são atribuídos os escores parciais de 0,1,2,3,4 e 5 em cada questão.

O Modelo de Crédito Parcial (MCP) é um dos membros da chamada família de modelos de Rasch (Masters e Wright, 1984). Os cinco modelos que pertencem a esta família se diferenciam pelo formato das respostas aos itens do teste e, possuem a mesma forma algébrica sendo de fato casos especiais do modelo logístico simples de Rasch, $\exp.(\theta_j - \beta_i)/[(1+ \exp.( \theta_j - \beta_i)]$.

A saber, os modelos são: "The Dichotomous" (Rasch, 1960), "Poisson Counts" (Rasch, 1960), "Binomial Trials" (Rasch, 1972), "Rating Scale" (Andrich, 1978b) e "Partial Credit" (Masters, 1982).

No MCP o item é caracterizado por uma série de passos ordenados, onde o indivíduo recebe crédito parcial para cada passo completado com sucesso. Por exemplo, consideremos um item tendo 6 possíveis escores, x=0,1,2,3,4,5 e conseqüentemente $m_i=6$ categorias de resposta, onde o escore zero denota que o indivíduo respondeu totalmente errado (ou não respondeu), o escore 5 (cinco) denota que o indivíduo respondeu adequadamente, e os escores intermediários representam créditos parciais sobre o item. Essa situação pode ser representada como segue:

**QUADRO 3.1**- Representação das probabilidades condicionais dos passos.

| | | | | | | |
|---|---|---|---|---|---|---|
| | | $\varphi_{i1}(\theta_j)$ ⇓ | $\varphi_{i2}(\theta_j)$ ⇓ | $\varphi_{i3}(\theta_j)$ ⇓ | $\varphi_{i4}(\theta_j)$ ⇓ | $\varphi_{i5}(\theta_j)$ ⇓ | |
| CATEGORIAS | | 1 | 2 | 3 | 4 | 5 | 6 |
| ESCORES | | 0 | 1 | 2 | 3 | 4 | 5 |
| PASSOS | 0 | 1 | 2 | 3 | 4 | 5 | |
| PROBABILIDADES DOS ESCORES | $P_{i0}(\theta_j)$ | $P_{i1}(\theta_j)$ | $P_{i2}(\theta_j)$ | $P_{i3}(\theta_j)$ | $P_{i4}(\theta_j)$ | $P_{i5}(\theta_j)$ | |

Note, por exemplo, que $\varphi_{i3}(\theta_j)$ é a probabilidade condicional do indivíduo de habilidade $\theta_j$ passar do escore 2 para o

escore 3 no item i dado que ele já obteve escore 2, e, a dificuldade dessa passagem, que será denominada de dificuldade individual do passo x=3, é dada por $\beta_{i3}$. $P_{ix}(\theta_j)$ é a probabilidade de um indivíduo de habilidade $\theta_j$ obter escore x no item i. Masters (1982, pg158), diz que não importa saber sobre a probabilidade de alcançar o nível 2 em primeiro lugar, que depende da habilidade $\theta_j$ e das dificuldades individuais dos dois primeiros passos ($\beta_{i1}$ e $\beta_{i2}$) no item i. Aqui partimos da condição de que o indivíduo já obteve escore 2 e queremos informação sobre a probabilidade deste indivíduo passar do escore 2 para o escore 3, ou seja, de completar o terceiro passo.

Para cada passo completado com sucesso o indivíduo recebe escore um. Podemos interpretar x como: uma contagem do número de passos que o indivíduo de habilidade $\theta_j$ completa no item i.

O indivíduo de habilidade $\theta_j$ será classificado em uma das $m_i$ categorias do item i. Assim, vai obter um dos $m_i$ possíveis escores no item i. Desta forma,

$$P_{i0}(\theta_j)+P_{i1}(\theta_j)+P_{i2}(\theta_j)+P_{i3}(\theta_j)+P_{i4}(\theta_j)+P_{i5}(\theta_j)=1 \quad (3.4)$$

No MCP, os passos que o indivíduo de habilidade $\theta_j$ completa no item i são considerados independentes uns dos outros, tendo cada passo as seguintes probabilidades condicionais:

$$\varphi_{i1}(\theta_j) = \frac{P_{i1}(\theta_j)}{P_{i0}(\theta_j) + P_{i1}(\theta_j)} = \frac{e^{(\theta_j - \beta_{i1})}}{1 + e^{(\theta_j - \beta_{i1})}}. \qquad (3.5)$$

$\varphi_{i1}(\theta_j)$ é a probabilidade condicional do indivíduo de habilidade $\theta_j$ completar o primeiro passo, dado que ele completou um passo (ou nenhum) e $\beta_{i1}$ é a dificuldade individual de completar o primeiro passo;

$$\varphi_{i2}(\theta_j) = \frac{P_{i2}(\theta_j)}{P_{i1}(\theta_j) + P_{i2}(\theta_j)} = \frac{e^{(\theta_j - \beta_{i2})}}{1 + e^{(\theta_j - \beta_{i2})}}. \qquad (3.6)$$

$\varphi_{i2}(\theta_j)$ é a probabilidade condicional do indivíduo de habilidade $\theta_j$ completar o segundo passo, dado que ele completou um ou dois passos e $\beta_{i2}$ é a dificuldade individual de completar o segundo passo;

E assim sucessivamente até,

$$\varphi_{i5}(\theta_j) = \frac{P_{i5}(\theta_j)}{P_{i4}(\theta_j) + P_{i5}(\theta_j)} = \frac{e^{(\theta_j - \beta_{i5})}}{1 + e^{(\theta_j - \beta_{i5})}}. \qquad (3.7)$$

$\varphi_{i5}(\theta_j)$ é a probabilidade condicional do indivíduo de habilidade $\theta_j$ completar o quinto passo, dado que ele completou quatro ou cinco passos e $\beta_{i5}$ é a dificuldade individual de completar o quinto passo.

Podemos perceber que $\beta_{i0}$, a dificuldade de escore zero no item i (ou não completar nenhum passo) é zero.

Das expressões (3.4) a (3.7) obtemos a expressão para a probabilidade $P_{ix}(\theta_j)$ de um indivíduo de habilidade $\theta_j$ obter escore x no item i. Este é o **Modelo de Crédito Parcial de Masters (MCPM)**.

$$P_{ix}(\theta_j)=\frac{e^{\sum_{n=0}^{x}(\theta_j-\beta_{in})}}{\sum_{k=0}^{m_i} e^{\sum_{n=0}^{k}(\theta_j-\beta_{in})}},\ x=0,1,2,3,\ldots,m_i-1 \qquad (3.8)$$

Vejamos, as expressões (3.5), (3.6) e (3.7) acima podem ser simplificadas como segue:

$$P_{i1}(\theta_j)=P_{i0}(\theta_j)e^{(\theta_j-\beta_{i1})} \qquad (3.9)$$

$$P_{i2}(\theta_j)=P_{i1}(\theta_j)e^{(\theta_j-\beta_{i2})} \qquad (3.10)$$

$$P_{i3}(\theta_j) = P_{i2}(\theta_j) e^{(\theta_j - \beta_{i3})} \qquad (3.11)$$

$$P_{i4}(\theta_j) = P_{i3}(\theta_j) e^{(\theta_j - \beta_{i4})} \qquad (3.12)$$

$$P_{i5}(\theta_j) = P_{i4}(\theta_j) e^{(\theta_j - \beta_{i5})}. \qquad (3.13)$$

Substituindo (3.9) em (3.10),

$$P_{i2}(\theta_j) = P_{i0}(\theta_j) e^{(\theta_j - \beta_{i1})} e^{(\theta_j - \beta_{i2})}. \qquad (3.14)$$

E, (3.14) substituir em (3.11),

$$P_{i3}(\theta_j) = P_{i0}(\theta_j) e^{(\theta_j - \beta_{i1})} e^{(\theta_j - \beta_{i2})} e^{(\theta_j - \beta_{i3})},$$

e assim sucessivamente até (3.13) (ou até $m_i$-1), obteremos

$$P_{ix}(\theta_j) = P_{i0}(\theta_j) e^{\sum_{n=1}^{x} (\theta_j - \beta_{in})}, \quad x = 1, 2, 3, \ldots, m_i - 1 \qquad (3.15)$$

Somando (3.15) em k (de 1 a $m_i$), teremos

$$\sum_{k=1}^{m_i} P_{ik}(\theta_j) = P_{i0}(\theta_j) \sum_{k=1}^{m_i} e^{\sum_{n=1}^{k} (\theta_j - \beta_{in})} \qquad (3.16)$$

De (3.4), temos que,

$\sum_{k=0}^{6} P_{ik}(\theta_j)=1 \Rightarrow \sum_{k=1}^{6} P_{ik}(\theta_j)=1-P_{i0}(\theta_j)$, generalizando para $m_i$ categorias,

$$\sum_{k=1}^{m_i} P_{ik}(\theta_j)=1-P_{i0}(\theta_j) \qquad (3.17)$$

Substituindo (3.17) em (3.16), teremos

$$1-P_{i0}(\theta_j)=P_{i0}(\theta_j)\sum_{k=1}^{m_i} e^{\sum_{n=1}^{k}(\theta_j-\beta_{in})} \Rightarrow$$

$$P_{i0}(\theta_j)=\frac{1}{1+\sum_{k=1}^{m_i} e^{\sum_{n=1}^{k}(\theta_j-\beta_{in})}} \qquad (3.18)$$

Substituindo (3.18) em (3.15), teremos

$$P_{ix}(\theta_j)=\frac{e^{\sum_{n=1}^{x}(\theta_j-\beta_{in})}}{1+\sum_{k=1}^{m_i} e^{\sum_{n=1}^{k}(\theta_j-\beta_{in})}},\ x=1,2,3,\ldots,m_i-1 \qquad (3.19)$$

Associando (3.18) e (3.19) e considerando $\sum_{n=0}^{0}(\theta_j-\beta_{in})\cong 0$ e $\sum_{n=0}^{k}(\theta_j-\beta_{in})\cong\sum_{n=1}^{k}(\theta_j-\beta_{in})$ podemos escrever uma única expressão para a probabilidade do indivíduo de habilidade $\theta_j$ obter

escore x no item i,

$$P_{ix}(\theta_j)=\frac{e^{\sum_{n=0}^{x}(\theta_j-\beta_{in})}}{\sum_{k=0}^{m_i} e^{\sum_{n=0}^{k}(\theta_j-\beta_{in})}},\ x=0,1,2,3,\ldots,m_i-1 \quad \square \qquad (3.20)$$

O parâmetro de discriminação $\alpha_i$ é considerado igual a um, para qualquer item e, a probabilidade em (3.20) é função da habilidade do indivíduo e das dificuldades individuais dos passos no item i.

No MCP o numerador contém somente termos para o passo completado pelo indivíduo de habilidade $\theta_j$ e, o denominador é a soma de todos os possíveis termos do numerador. Por causa disso, Thissen e Steinberg (1984) classificaram a abordagem do MCP de Masters (1982) como um modelo "dividido pela soma total".

Para melhor compreensão da diferença entre os parâmetros de dificuldade do MCP de Masters e do MRC de Samejima vamos considerar dois itens na Figura 3.2, um para cada modelo. Em (a), temos seis Curvas Características das Categorias de Resposta dadas pelo MCP, obtidas por um dos itens da prova de inglês do vestibular/97 da UNICAMP, onde as estimativas dos parâmetros são: $\beta_1$=-0.78, $\beta_2$=-0.37, $\beta_3$=1.73, $\beta_4$=-0.37 e $\beta_5$=1.09. Em (b), temos três Curvas Características das Categorias de Resposta dadas pelo MRC de

Samejima, sendo: Liberal, Moderado e Conservador as três categorias ordenadas, onde as estimativas dos parâmetros são: $\lambda_1$=-0.93 e $\lambda_2$=0.97.

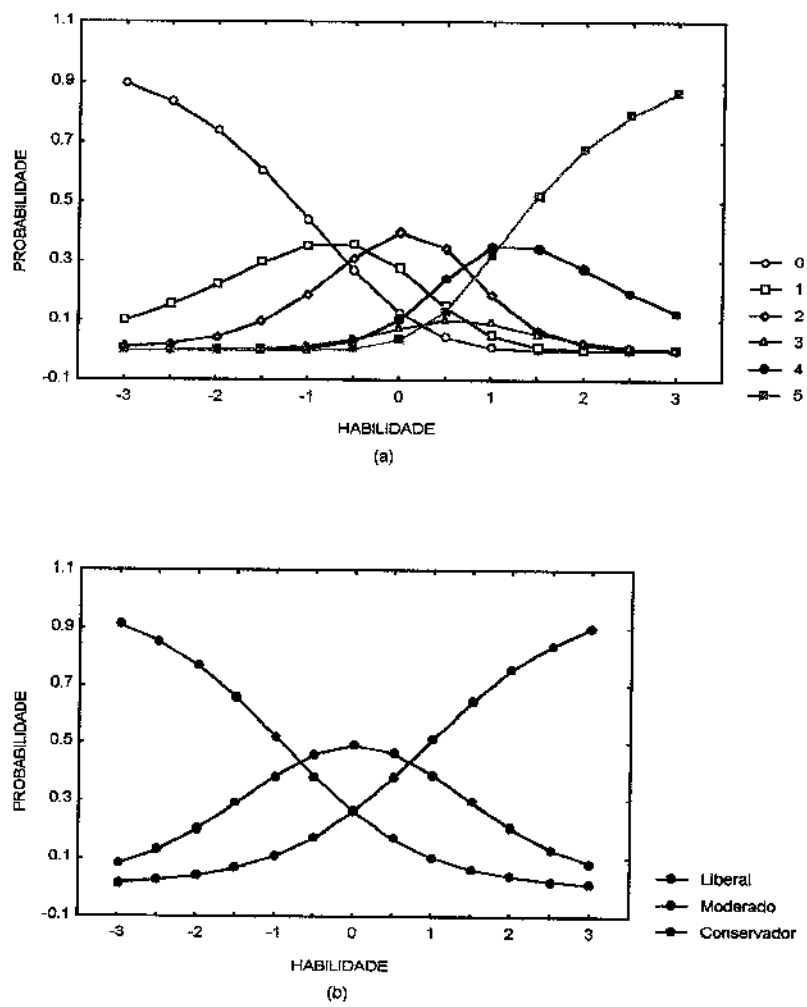



**FIGURA 3.2-** (a) Curvas Características para seis categorias de resposta dadas pelo MCP. (b) Curvas Características para três categorias de resposta dadas pelo MRCS.

Na Figura 3.2 (a), os parâmetros das categorias de resposta do item obtidos através do MCP são localizados através das

interseções das curvas de probabilidades de respostas 0 e 1, 1 e 2, 2 e 3, 3 e 4 e 4 e 5, ou seja: $\beta_{i1}$ na escala de habilidade corresponde ao ponto onde $P_{i0}(\theta_j)=P_{i1}(\theta_j)$ e $\beta_{i2}$ corresponde ao ponto onde $P_{i1}(\theta_j)=P_{i2}(\theta_j)$ e assim sucessivamente. Enquanto que sob o modelo de resposta categórica de Samejima o parâmetro $\lambda_{is}$ da curva característica limite (CCL) corresponde à escala de habilidade onde $P^*_{ik}(\theta_j)=\sum_{s=k}^{m_i} P_{is}(\theta_j)=0,5$ (Figura 3.2(b)) e, $\lambda_{is}$ está relacionado com probabilidades acumuladas por isso são sempre ordenados $\lambda_{i1} \leq \lambda_{i2} \leq \ldots \leq \lambda_{i,m_i-1}$.

No MCP as dificuldades individuais dos passos $(\beta_{ix})$ não precisam necessariamente serem ordenadas, pois, como as dificuldades para completar cada passo são consideradas independentes pode ocorrer que a dificuldade individual do último passo seja menor que as dificuldades individuais anteriores (verifique a Figura 3.2 (a)).

Quando os dados se adequarem ao MCP os valores $\beta_{i1},\beta_{i2},\beta_{i3},\beta_{i4}$ e $\beta_{i5}$ não dependem dos valores de $\beta_{i'1},\beta_{i'2},\beta_{i'3},\beta_{i'4}$ e $\beta_{i'5}$ de outro item i' qualquer no teste. Dentro de um item, no entanto, os parâmetros de dificuldade dos passos não podem ser interpretados separadamente (Molenaar,1983).

Quando as categorias de resposta são definidas da mesma maneira para todos os itens, como em um escala "rating" a equação (3.20) pode ser simplificada pela proposição de que as dificuldades individuais dos passos em cada item são invariantes de item para item e os itens diferem somente em sua locação sob a escala de habilidade. Esta simplificação leva ao "Rating Scale Model" (Andrich, 1978b).

Nesse caso a dificuldade do passo x para o item *i* é reformulada como $\beta_{ix} = \beta_i + \tau_x$, onde $\beta_i$ é a locação do item sobre a escala de habilidade e $\tau_x$ é a locação do x-ésimo passo relativo à locação do item.

Assim $\beta_{ix}$ na expressão geral é substituído por $\beta_i + \tau_x$ e, sob o "rating scale model" existem $m_i$-1 passos e N parâmetros de habilidade para serem estimados enquanto que sob o modelo de crédito parcial, ($m_i$-1)L passos e N parâmetros de habilidade são estimados.

## 3.3- MODELO DE CRÉDITO PARCIAL GENERALIZADO (MCPG) -MURAKI

O Modelo de Crédito Parcial Generalizado de Muraki (1992) é assim chamado por considerar também o parâmetro de discriminação do item que no modelo de Masters (1982) é considerado constante a todos os itens. Logo, a relação entre a probabilidade de ocorrência de uma resposta e a habilidade é medida através do modelo de Muraki onde o parâmetro de discriminação ($\alpha_i$, i=1,2,...,L) no Modelo de Crédito Parcial de Masters (1982).

$$P_{ix}(\theta_j)=\frac{exp\left[\sum_{n=0}^{x}1.7\alpha_i(\theta_j-\beta_{in})\right]}{\sum_{k=0}^{m_i}exp\left[\sum_{n=0}^{k}1.7\alpha_i(\theta_j-\beta_{in})\right]},\ x=0,1,2,\ldots,m_i-1 \quad (3.21)$$

onde $\beta_{i0}=0$, $\sum_{n=0}^{k}1.7\alpha_i(\theta_j-\beta_{in})\cong\sum_{n=1}^{k}1.7\alpha_i(\theta_j-\beta_{in})$.

$\rightarrow P_{ix}(\theta_j)$ representa a probabilidade condicional de obter escore x no item i;

$\rightarrow \alpha_i$ é o parâmetro de discriminação do item i;

$\rightarrow \beta_{ix}$ é o parâmetro de dificuldade individual associado à transição do passo x, e

$\rightarrow m_i$ é o número de categorias do item i.

# CAPÍTULO 4

## ESTIMAÇÃO DOS PARÂMETROS

A preocupação central da Teoria de Resposta ao Item é classificar as habilidades dos indivíduos sem que importe o instrumento utilizado bem como estimar os parâmetros dos itens independentemente da amostra empregada. Se o modelo escolhido se ajusta bem aos dados os dois objetivos acima se cumprem, o que pode ser verificado através da calibração dos itens. Quanto maior o tamanho da amostra de indivíduos e quanto melhor cubram o "rank" valores de $\theta$, mais precisas serão as estimativas dos parâmetros.

Na literatura observamos que para a estimação dos parâmetros dos itens, somente no modelo dicotômico de Rasch é natural o uso da Estimação de Máxima Verossimilhança Condicional (EMV Condicional). Nesse caso conhecemos a função de distribuição das habilidades dos indivíduos fazendo com que as estimativas dos parâmetros de dificuldade dos itens sejam independentes dos parâmetros dos indivíduos. Nesse procedimento de EMV Condicional, desenvolvido por Andersen (1972), são estimados somente os parâmetros de dificuldade dos itens.

O problema está em obter a função de distribuição das habilidades dos indivíduos que seja consistente, por isso, a Estimação de Máxima Verossimilhança Marginal (EMVM) tem seu importante papel na estimação de parâmetros nos modelos da TRI. Na seção 4.4 é desenvolvido o procedimento de EMVM para estimar os parâmetros sob o Modelo de Crédito Parcial, visto que o programa MULTILOG (Thissen, 1991) utilizado na aplicação do Capítulo 6, utiliza este método para estimar os parâmetros dos itens.

Thissen (1982) descreve dois algoritmos usando Estimação de Máxima Verossimilhança Marginal para a estimação dos parâmetros do modelo logístico de um parâmetro (Modelo de Rasch).

A Estimação de Máxima Verossimilhança Conjunta (EMVC) estima juntamente parâmetros dos itens e parâmetros dos indivíduos iterativamente em dois estágios. Este procedimento será apresentado na seção 4.1 para itens dicotômicos, usando o modelo logístico de dois parâmetros de Birnbaun (1968), e na seção 4.3 para itens categóricos, usando o modelo de resposta categórica de Samejima (1972).

## 4.1- ESTIMAÇÃO DOS PARÂMETROS PARA O CASO DE ITENS DICOTÔMICOS

Depois de selecionar um dos modelos já descritos nos Capítulos anteriores, que se ajuste aos dados, o passo seguinte é estimar os parâmetros de cada item e o parâmetro de cada indivíduo. Assim, para o caso de respostas dicotômicas serão estimados ML+N parâmetros, onde M=1,2,3, é o número de parâmetros utilizados no modelo e L e N representam, respectivamente, o número de itens utilizados e o número de indivíduos avaliados.

O que temos são as respostas dos indivíduos aos itens e com estas respostas empíricas iremos estimar os parâmetros referentes aos itens e os parâmetros referentes aos indivíduos. Esses parâmetros devem ser estimados de forma a maximizar as probabilidades de ocorrência dos dados, ao aplicar os itens aos indivíduos.

Como são muitos os parâmetros a serem estimados se torna indispensável o uso de programas computacionais. Atualmente dispomos de três programas importantes para o caso dicotômico: BICAL (Wrigth e Mead, 1979) para modelos logísticos de um parâmetro; LOGIST (Wingersky, Barton e Lord, 1982) e BILOG (Mislery e Bock, 1984) para modelos logísticos de um, dois e três

parâmetros. Outros programas, para o caso de itens politômicos, que serão citados posteriormente, poderão ser adaptados para o caso de itens dicotômicos, sendo esses, casos especiais de itens politômicos.

Exposições detalhadas sobre a estimação dos parâmetros podem ser encontradas em Baker (1987 e 1992), Birnbaum (1968), Lord (1980) ou Swanithan (1983), sendo especialmente recomendado o trabalho de Hambleton e Swanithan (1985,1991,1996) por sua clareza e utilização de exemplos numéricos que ajudam a compreensão.

Nesse contexto usaremos somente o modelo logístico, pois sua vantagem sobre o modelo de distribuição normal padrão é a de não envolver integrais além da distribuição acumulada da densidade logística possuir forma fechada e poder ser facilmente obtida.

Consideremos que L itens dicotômicos sejam aplicados a N indivíduos. Seja $X_{ji}$ a variável aleatória indicadora de acerto ao i-ésimo item pelo j-ésimo indivíduo, dotado de habilidade $\theta_j$, j = 1, 2,...,N e i = 1,2,..., L.

Para cada indivíduo existirá um vetor de respostas aos itens, de comprimento L, denotado por $(x_{j1}, x_{j2},...,x_{jL}|\theta_j)$.

Vimos que a CCI descreve a probabilidade de acerto ou erro a um item para um dado valor de $\theta$, ou seja:

$$P(X_{ji}=1|\theta_j)= P_i(\theta_j) \text{ e } P(X_{ji}=0|\theta_j) = Q_i(\theta_j)=1-P_i(\theta_j)$$

Assim, a resposta a um item para um determinado nível de $\theta$ é uma variável aleatória com distribuição de Bernoulli com parâmetro $P_i(\theta_j)$. Então,

$$P_i\left(X_{ji}|\theta_j\right) = \left[P_i\left(\theta_j\right)\right]^{x_{ji}}\left[Q_i\left(\theta_j\right)\right]^{1-x_{ji}}, \qquad (4.1)$$

e, sob a suposição de independência condicional, para todos os indivíduos com a mesma habilidade a função de probabilidade conjunta de $X_{j1},X_{j2},\ldots,X_{jL}$, no ponto $x_{j1}, x_{j2},\ldots, x_{jL}$ é dada por

$$P(x_{j1},x_{j2},\ldots,x_{jL}|\theta_j,\alpha,\beta,\psi)= \prod_{i=1}^{L}\left[P_i\left(\theta_j\right)\right]^{x_{ji}}\left[Q_i\left(\theta_j\right)\right]^{1-x_{ji}} \qquad (4.2)$$

Os valores de $P_i\left(\theta_j\right)$ e $Q_i\left(\theta_j\right)$ são dados pela CCI, portanto variam com o modelo de TRI utilizado.

Para os N indivíduos a função de verossimilhança (4.2) será escrita como:

$$P(\theta,\alpha,\beta,\psi|x_1,x_2,\ldots,x_N)= \prod_{j=1}^{N}\prod_{i=1}^{L}\left[P_i\left(\theta_j\right)\right]^{x_{ji}}\left[Q_i\left(\theta_j\right)\right]^{1-x_{ji}}, \qquad (4.3)$$

Maximizar (4.3) é equivalente a maximizar o logaritmo da função de verossimilhança,

$$\ell = \sum_{j=1}^{N} \sum_{i=1}^{L} \left[ x_{ji} \ln P_i(\theta_j) + (1 - x_{ji}) \ln Q_i(\theta_j) \right] \qquad (4.4)$$

O vetor de parâmetros, $P_y$, é estimado pelo método da máxima verossimilhança (MV), ou seja, a estimativa de $P_y$ é obtida maximizando-se a função de verossimilhança dada em (4.4).

O vetor de estimativas de máxima verossimilhança (EMV) dos parâmetros será o vetor de soluções da equação $\frac{\partial \ell}{\partial P_y} = 0$, onde $P_y$ é o vetor de parâmetros a estimar, $P_y = (\theta, \alpha, \beta, \psi)$ e y é o número de parâmetros.

Supondo conhecidos os parâmetros dos itens os valores de θ são obtidos pelo método de Estimação de Máxima Verossimilhança Condicional (EMVC), assim chamada por estar condicionada ao conhecimento prévio dos parâmetros dos itens. Esse método de estimação é considerado pioneiro na estimação dos parâmetros nos modelos dicotômicos da TRI e sua aplicação é limitada aos modelos que pertencem a chamada "família de modelos Rasch" (Masters e Wright, 1984).

O fato é que geralmente não conhecemos os parâmetros dos itens e assim é necessário estimar em conjunto os parâmetros dos itens e os parâmetros dos indivíduos; tal procedimento é denominado Estimação de Máxima Verossimilhança Conjunta (EMVConjunta), e foi proposto por Birnbaum (1968).

### 4.1.1- ESTIMAÇÃO DE MÁXIMA VEROSSIMILHANÇA CONJUNTA

Lord (1980) argumenta que, quando estimativas dos parâmetros de itens e parâmetros de indivíduos são obtidas simultaneamente, tais estimativas não convergem para os parâmetros verdadeiros quando o número de indivíduos é muito grande.

Na EMVConjunta um procedimento de estimação em dois estágios foi desenvolvido: num primeiro estágio são estimados os parâmetros dos itens supondo conhecidos os parâmetros dos indivíduos e, num segundo estágio são estimados os parâmetros dos indivíduos supondo conhecidos os parâmetros dos itens (os estimados no primeiro estágio). Para dar início ao processo, os escores (escores brutos) padronizados do teste são geralmente usados como valores iniciais "conhecidos" das habilidades dos indivíduos (Wood, Wingersky, e Lord, 1976). Esse processo cíclico, que opera iterativamente com esses dois estágios em cada ciclo, se repete até que seja atingido um critério de convergência pré determinado.

Alguns programas como BIMAIN e BILOG que calculam as estatísticas da Teoria Clássica de Testes (TCT), como índice de facilidade do item e índice de discriminação[1], as utilizam como valores iniciais dos parâmetros dos itens para a estimação iterativa no primeiro estágio.

(1) Sob a suposição de que $\theta$ se distribui normalmente com média igual a zero e desvio padrão igual a um na população de indivíduos, os parâmetros da distribuição normal da TRI se relacionam com as estatísticas da TCT como segue: **Índice de facilidade do item** (proporção de indivíduos que respondem o item i corretamente) : $p_i = \Phi\left(-\alpha_i\beta_i / \sqrt{1+\alpha_i^2}\right)$, (4.5)

onde $p_i$ é o valor da distribuição normal padronizada no ponto $-\alpha_i\beta_i / \sqrt{1+\alpha_i^2}$.

**Índice de discriminação** (correlação biserial entre os escores do item i e o escore da característica latente) : $\rho_i = \alpha_i \Big/ \sqrt{1+\alpha_i^2}$ , (4.6)

que na TCT é estimado como a correlação item-total. Essas aproximações se mantém quando se aplicam os modelos logísticos, mas são expressos na métrica normal, Algina e Crocker (1986,p.350e351).

O programa LOGIST, desenvolvido por Wingersky e Lord (1973), tem a finalidade de obter estimativas com base no procedimento proposto por Birnbaum (1968) para o modelo logístico de três parâmetros.

Algumas suposições deverão ser satisfeitas para que seja possível a realização das derivadas da função $\ell$ dada em (4.4) para cada indivíduo e para cada item.

## 4.1.2- SUPOSIÇÕES

1. Independência entre os parâmetros dos indivíduos (habilidades): $COV(\theta_j, \theta_{j'})=0$, para todo $j \neq j'$.

2. Independência entre os parâmetros de distintos itens:

$COV(\alpha_i, \alpha_{i'})=COV(\alpha_i, \beta_{i'})=.....=COV(\psi_i, \psi_{i'})=0$,

para todo $i \neq i'$.

3. Independência entre os parâmetros dos indivíduos e os parâmetros dos itens: $COV(\theta_j, \alpha_i)=0$, $COV(\theta_j, \beta_i)=0$ e $COV(\theta_j, \psi_i)=0$, para todo $i=1, 2,..., L$ e $j=1, 2,..., N$.

## 4.1.3- MÉTODO DE NEWTON - RAPHSON

Se as duas primeiras derivadas de (4.4) existem e a função é monótona, o método de Newton-Raphson pode ser utilizado para encontrar suas raízes. O método consiste em estimações sucessivas dos parâmetros até que se obtenha a convergência, isto é, até que o valor obtido na iteração t não difira significativamente da obtida na iteração t-1.

Usando a equação do método de Newton-Raphson em forma matricial, temos

$A_{t+1}=A_t- B_t^{-1}.F_t$ , onde: (4.7)

A = vetor coluna dos parâmetros de habilidades e parâmetros do item de dimensão (ML+N)×(1);

B = matriz de derivadas parciais de segunda ordem de dimensão (ML+N)×(ML+N). B é chamada matriz HESSIANA ou matriz de informação de Newton-Raphson;

F = vetor coluna de derivadas de primeira ordem da função de probabilidade $\ell$, de comprimento (ML+N)×(1);

t = índice que indica a t-ésima iteração.

Assim, a matriz de probabilidade (NxL) de respostas aos itens é dada pela função de verossimilhança (4.4) e, considerando o modelo logístico de dois parâmetros dado em (2.8), ou seja M=2, a equação (4.7) é expressa abaixo, onde os vetores A e F tem a dimensão (2L+N)×(1) (L é igual ao n° de itens e, N é igual ao n° de individuos) e, a matriz B tem dimensão (2L+N)×(2L+N).

$$\begin{bmatrix} \hat{\alpha}_1 \\ \hat{\vartheta}_1 \\ \hat{\alpha}_2 \\ \hat{\vartheta}_2 \\ \vdots \\ \hat{\alpha}_L \\ \hat{\vartheta}_L \\ \hat{\theta}_1 \\ \hat{\theta}_2 \\ \vdots \\ \hat{\theta}_{N-1} \\ \hat{\theta}_N \end{bmatrix}_{t+1} = \begin{bmatrix} \hat{\alpha}_1 \\ \hat{\vartheta}_1 \\ \hat{\alpha}_2 \\ \hat{\vartheta}_2 \\ \vdots \\ \hat{\alpha}_L \\ \hat{\vartheta}_L \\ \hat{\theta}_1 \\ \hat{\theta}_2 \\ \vdots \\ \hat{\theta}_{N-1} \\ \hat{\theta}_N \end{bmatrix}_t - \begin{bmatrix} \frac{\partial^2 \ell}{\partial^2 \alpha_1} & \frac{\partial^2 \ell}{\partial \alpha_1 \partial \vartheta_1} & \frac{\partial^2 \ell}{\partial \alpha_1 \partial \alpha_2} & \frac{\partial^2 \ell}{\partial \alpha_1 \partial \vartheta_2} & \cdots & \frac{\partial^2 \ell}{\partial \alpha_1 \partial \vartheta_L} & \frac{\partial^2 \ell}{\partial \alpha_1 \partial \theta_1} & \frac{\partial^2 \ell}{\partial \alpha_1 \partial \theta_2} & \cdots \cdots & \frac{\partial^2 \ell}{\partial \alpha_1 \partial \theta_N} \\ \frac{\partial^2 \ell}{\partial \vartheta_1 \partial \alpha_1} & \frac{\partial^2 \ell}{\partial^2 \vartheta_1} & \frac{\partial^2 \ell}{\partial \vartheta_1 \partial \alpha_2} & \frac{\partial^2 \ell}{\partial \vartheta_1 \partial \vartheta_2} & \cdots & \frac{\partial^2 \ell}{\partial \vartheta_1 \partial \vartheta_L} & \frac{\partial^2 \ell}{\partial \vartheta_1 \partial \theta_1} & \frac{\partial^2 \ell}{\partial \vartheta_1 \partial \theta_2} & \cdots \cdots & \frac{\partial^2 \ell}{\partial \vartheta_1 \partial \theta_N} \\ \frac{\partial^2 \ell}{\partial \alpha_2 \partial \alpha_1} & \frac{\partial^2 \ell}{\partial \alpha_2 \partial \vartheta_1} & \frac{\partial^2 \ell}{\partial^2 \alpha_2} & \frac{\partial^2 \ell}{\partial \alpha_2 \partial \vartheta_2} & \cdots & \frac{\partial^2 \ell}{\partial \alpha_2 \partial \vartheta_L} & \frac{\partial^2 \ell}{\partial \alpha_2 \partial \theta_1} & \frac{\partial^2 \ell}{\partial \alpha_2 \partial \theta_2} & \cdots \cdots & \frac{\partial^2 \ell}{\partial \alpha_2 \partial \theta_N} \\ \frac{\partial^2 \ell}{\partial \vartheta_2 \partial \alpha_1} & \frac{\partial^2 \ell}{\partial \vartheta_2 \partial \vartheta_1} & \frac{\partial^2 \ell}{\partial \vartheta_1 \partial \alpha_2} & \frac{\partial^2 \ell}{\partial^2 \vartheta_1} & \cdots & \frac{\partial^2 \ell}{\partial \vartheta_2 \partial \vartheta_L} & \frac{\partial^2 \ell}{\partial \vartheta_2 \partial \theta_1} & \frac{\partial^2 \ell}{\partial \vartheta_2 \partial \theta_2} & \cdots \cdots & \frac{\partial^2 \ell}{\partial \vartheta_2 \partial \theta_N} \\ \vdots & & & & \ddots & & & & \ddots & \\ \frac{\partial^2 \ell}{\partial \alpha_L \partial \alpha_1} & \frac{\partial^2 \ell}{\partial \alpha_L \partial \vartheta_1} & \frac{\partial^2 \ell}{\partial \alpha_L \partial \alpha_2} & \frac{\partial^2 \ell}{\partial \alpha_L \partial \vartheta_2} & \cdots & \frac{\partial^2 \ell}{\partial \alpha_L \partial \vartheta_L} & \frac{\partial^2 \ell}{\partial \alpha_L \partial \theta_1} & \frac{\partial^2 \ell}{\partial \alpha_L \partial \theta_2} & \cdots \cdots & \frac{\partial^2 \ell}{\partial \alpha_L \partial \theta_N} \\ \frac{\partial^2 \ell}{\partial \vartheta_L \partial \alpha_1} & \frac{\partial^2 \ell}{\partial \vartheta_L \partial \vartheta_1} & \frac{\partial^2 \ell}{\partial \vartheta_L \partial \alpha_2} & \frac{\partial^2 \ell}{\partial \vartheta_L \partial \vartheta_2} & \cdots & \frac{\partial^2 \ell}{\partial^2 \vartheta_L} & \frac{\partial^2 \ell}{\partial \vartheta_L \partial \theta_1} & \frac{\partial^2 \ell}{\partial \vartheta_L \partial \theta_2} & \cdots \cdots & \frac{\partial^2 \ell}{\partial \vartheta_L \partial \theta_N} \\ \frac{\partial^2 \ell}{\partial \theta_1 \partial \alpha_1} & \frac{\partial^2 \ell}{\partial \theta_1 \partial \vartheta_1} & \frac{\partial^2 \ell}{\partial \theta_1 \partial \alpha_2} & \frac{\partial^2 \ell}{\partial \theta_1 \partial \vartheta_2} & \cdots & \frac{\partial^2 \ell}{\partial \theta_1 \partial \vartheta_L} & \frac{\partial^2 \ell}{\partial^2 \theta_1} & \frac{\partial^2 \ell}{\partial \theta_1 \partial \theta_2} & \cdots \cdots & \frac{\partial^2 \ell}{\partial \theta_1 \partial \theta_N} \\ \frac{\partial^2 \ell}{\partial \theta_2 \partial \alpha_1} & \frac{\partial^2 \ell}{\partial \theta_2 \partial \vartheta_1} & \frac{\partial^2 \ell}{\partial \theta_2 \partial \alpha_2} & \frac{\partial^2 \ell}{\partial \theta_2 \partial \vartheta_2} & \cdots & \frac{\partial^2 \ell}{\partial \theta_2 \partial \vartheta_L} & \frac{\partial^2 \ell}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 \ell}{\partial^2 \theta_2} & \cdots \cdots & \frac{\partial^2 \ell}{\partial \theta_2 \partial \theta_N} \\ \vdots & & & & \ddots & & & & \ddots & \\ \frac{\partial^2 \ell}{\partial \theta_{N-1} \partial \alpha_1} & \frac{\partial^2 \ell}{\partial \theta_{N-1} \partial \vartheta_1} & \frac{\partial^2 \ell}{\partial \theta_{N-1} \partial \alpha_2} & \frac{\partial^2 \ell}{\partial \theta_{N-1} \partial \vartheta_2} & \cdots & \frac{\partial^2 \ell}{\partial \theta_{N-1} \partial \vartheta_L} & \frac{\partial^2 \ell}{\partial \theta_{N-1} \partial \theta_1} & \frac{\partial^2 \ell}{\partial \theta_{N-1} \partial \theta_2} & \cdots \cdots & \frac{\partial^2 \ell}{\partial \theta_{N-1} \partial \theta_N} \\ \frac{\partial^2 \ell}{\partial \theta_N \partial \alpha_1} & \frac{\partial^2 \ell}{\partial \theta_N \partial \vartheta_1} & \frac{\partial^2 \ell}{\partial \theta_N \partial \alpha_2} & \frac{\partial^2 \ell}{\partial \theta_N \partial \vartheta_2} & \cdots & \frac{\partial^2 \ell}{\partial \theta_N \partial \vartheta_L} & \frac{\partial^2 \ell}{\partial \theta_N \partial \theta_1} & \frac{\partial^2 \ell}{\partial \theta_N \partial \theta_2} & \cdots \cdots & \frac{\partial^2 \ell}{\partial^2 \theta_N} \end{bmatrix}_t^{-1} \begin{bmatrix} \frac{\partial \ell}{\partial \alpha_1} \\ \frac{\partial \ell}{\partial \vartheta_1} \\ \frac{\partial \ell}{\partial \alpha_2} \\ \frac{\partial \ell}{\partial \vartheta_2} \\ \vdots \\ \frac{\partial \ell}{\partial \alpha_L} \\ \frac{\partial \ell}{\partial \vartheta_L} \\ \frac{\partial \ell}{\partial \theta_1} \\ \frac{\partial \ell}{\partial \theta_2} \\ \vdots \\ \frac{\partial \ell}{\partial \theta_{N-1}} \\ \frac{\partial \ell}{\partial \theta_N} \end{bmatrix}_t$$

Como $P_i(\theta)$ tem uma relação não linear com $\theta$ foi aplicada uma transformação adequada obtendo uma relação linear entre a nova variável, denotada por Z, e $\theta$. Logo, de (2.7), $Z_{ji} = \alpha_i(\theta_j - \beta_i) = -\alpha_i\beta_i + \alpha_i\theta_j$. Denotaremos por $\vartheta_i$ e $\alpha_i$ o intercepto e o coeficiente angular da reta, respectivamente, onde $\vartheta_i = -\alpha_i\beta_i$.

BIBLIOTECA CENTRAL

Devido a independência entre os parâmetros de distintos indivíduos, entre os parâmetros de distintos itens e entre os parâmetros de indivíduos e itens, dado na seção 4.1.2 como suposições, a forma matricial da equação de Newton-Raphson se reduz a forma matricial abaixo.

$$\begin{bmatrix} \hat{\alpha}_1 \\ \hat{\vartheta}_1 \\ \hat{\alpha}_2 \\ \hat{\vartheta}_2 \\ \vdots \\ \hat{\alpha}_L \\ \hat{\vartheta}_L \\ \hat{\theta}_1 \\ \\ \vdots \\ \hat{\theta}_N \end{bmatrix}_{(t-1)} = \begin{bmatrix} \hat{\alpha}_1 \\ \hat{\vartheta}_1 \\ \hat{\alpha}_2 \\ \hat{\vartheta}_2 \\ \vdots \\ \hat{\alpha}_L \\ \hat{\vartheta}_L \\ \hat{\theta}_1 \\ \\ \vdots \\ \hat{\theta}_N \end{bmatrix}_{t} - \begin{bmatrix} \frac{\partial^2 \ell}{\partial^2 \alpha_1} & \frac{\partial^2 \ell}{\partial \alpha_1 \partial \vartheta_1} & & & & & & & & \\ \frac{\partial^2 \ell}{\partial \vartheta_1 \partial \alpha_1} & \frac{\partial^2 \ell}{\partial^2 \vartheta_1} & & & & & & & & \\ & & \frac{\partial^2 \ell}{\partial^2 \alpha_2} & \frac{\partial^2 \ell}{\partial \alpha_2 \partial \vartheta_2} & & & & & & \\ & & \frac{\partial^2 \ell}{\partial \vartheta_2 \partial \alpha_2} & \frac{\partial^2 \ell}{\partial^2 \vartheta_2} & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & \frac{\partial^2 \ell}{\partial^2 \alpha_L} & \frac{\partial^2 \ell}{\partial \alpha_L \partial \vartheta_L} & & & \\ & & & & & \frac{\partial^2 \ell}{\partial \vartheta_L \partial \alpha_L} & \frac{\partial^2 \ell}{\partial^2 \vartheta_L} & & & \\ & & & & & & & \frac{\partial^2 \ell}{\partial^2 \theta_1} & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \frac{\partial^2 \ell}{\partial^2 \theta_N} \end{bmatrix}_{t}^{-1} \begin{bmatrix} \frac{\partial \ell}{\partial \alpha_1} \\ \frac{\partial \ell}{\partial \vartheta_1} \\ \frac{\partial \ell}{\partial \alpha_2} \\ \frac{\partial \ell}{\partial \vartheta_2} \\ \vdots \\ \frac{\partial \ell}{\partial \alpha_L} \\ \frac{\partial \ell}{\partial \vartheta_L} \\ \frac{\partial \ell}{\partial \theta_1} \\ \\ \vdots \\ \frac{\partial \ell}{\partial \theta_N} \end{bmatrix}_{t}$$

Podemos observar que os vetores A e F e a matriz B continuam com as mesmas dimensões, apesar de muitos termos em B terem sido anulados com as suposições da seção 4.1.2.

Considerando separadamente cada item, as variáveis aleatórias correspondentes as respostas dos indivíduos a um único item é apenas função dos parâmetros $\vartheta_i$ e $\alpha_i$. E, a função de verossimilhança em (4.4) será dada por,

$$\ell\left(\vartheta_i,\alpha_i / x_{1i},x_{2i},\ldots,x_{Ni}\right)=\sum_{j=1}^{N}\left[x_{ji}\ln P_i\left(\theta_j\right)+\left(1-x_{ji}\right)\ln Q_i\left(\theta_j\right)\right] \qquad (4.8)$$

Observando que existe uma submatriz 2x2 para cada item na matriz B da equação de Newton-Raphson e um termo correspondente a cada um dos N indivíduos na diagonal da matriz B, temos, para o i-ésimo item:

$$\begin{bmatrix}\hat{\alpha}_i\\ \hat{\vartheta}_i\end{bmatrix}_{t+1}=\begin{bmatrix}\hat{\alpha}_i\\ \hat{\vartheta}_i\end{bmatrix}_t-\begin{bmatrix}\frac{\partial^2\ell}{\partial^2\alpha_i} & \frac{\partial^2\ell}{\partial\alpha_i\partial\vartheta_i}\\ \frac{\partial^2\ell}{\partial\vartheta_i\partial\alpha_i} & \frac{\partial^2\ell}{\partial^2\vartheta_i}\end{bmatrix}_t^{-1}\cdot\begin{bmatrix}\frac{\partial\ell}{\partial\alpha_i}\\ \frac{\partial\ell}{\partial\vartheta_i}\end{bmatrix}_t \qquad (4.9)$$

O método de Newton-Raphson para obtenção das estimativas de MV de $\vartheta_i$ e $\alpha_i$, consiste na iteração de (4.9) com $\begin{bmatrix}\hat{\alpha}_i\\ \hat{\vartheta}_i\end{bmatrix}_t$ substituído pelo valor de $\begin{bmatrix}\hat{\alpha}_i\\ \hat{\vartheta}_i\end{bmatrix}_{t+1}$ obtido na iteração anterior. Para a primeira iteração é escolhido um valor inicial para $\alpha_i^{(0)}$ e $\vartheta_i^{(0)}$. Muitos programas usam os valores obtidos da TCT como o índice facilidade e o índice de discriminação do item, definidos no rodapé da página 62.

Se por outro lado, substituirmos as derivadas de segunda ordem em (4.9) pelas suas esperanças em cada iteração, o procedimento é conhecido como método do escore de Fisher, isto é

$$\begin{bmatrix} \hat{\alpha}_i \\ \hat{\vartheta}_i \end{bmatrix}_{t+1} = \begin{bmatrix} \hat{\alpha}_i \\ \hat{\vartheta}_i \end{bmatrix}_t - \begin{bmatrix} E\left[\frac{\partial^2 \ell}{\partial^2 \alpha_i}\right] & E\left[\frac{\partial^2 \ell}{\partial \alpha_i \partial \vartheta_i}\right] \\ E\left[\frac{\partial^2 \ell}{\partial \vartheta_i \partial \alpha_i}\right] & E\left[\frac{\partial^2 \ell}{\partial^2 \vartheta_i}\right] \end{bmatrix}_t^{-1} \cdot \begin{bmatrix} \frac{\partial \ell}{\partial \alpha_i} \\ \frac{\partial \ell}{\partial \vartheta_i} \end{bmatrix}_t \qquad (4.10)$$

A inversa da matriz de esperanças acima é a matriz de variâncias e covariâncias das estimativas dos parâmetros do item i, para i=1,2,...,L.

Resolvendo (4.10) para todos os itens, teremos concluído o **primeiro estagio** e, as estimativas dos parâmetros do item estimadas por (4.10) são usadas como parâmetros verdadeiros dos itens no **segundo estágio,** onde encontraremos as estimativas dos parâmetros dos indivíduos (habilidades) usando a equação:

$$\left[\hat{\theta}_j\right]_{t+1} = \left[\hat{\theta}_j\right]_t - \left[E\left(\frac{\partial^2 \ell}{\partial^2 \theta_j}\right)\right]_t^{-1} \cdot \left[\frac{\partial \ell}{\partial \theta_j}\right]_t, \quad j=1,2,\ldots,N. \qquad (4.11)$$

Novamente a derivada de segunda ordem foi substituída pela sua esperança sob o método do escore de Fisher.

Finalmente o vetor A de estimativas de máxima verossimilhança de comprimento (2L+N) é obtido depois de encontradas as estimativas dos parâmetros de todos os itens e de todos os indivíduos, ou seja, resolvendo (4.10) para cada item e (4.11) para cada indivíduo.

No caso dicotômico, onde a soma dos escores dos itens para o j-ésimo indivíduo é uma estatística suficiente para o parâmetro de habilidade $\theta_j$, todos os indivíduos com o mesmo escore no teste terão a mesma estimativa da habilidade $\hat{\theta}_j$, mesmo que seus verdadeiros parâmetros de habilidade sejam diferentes. Dessa forma, podemos agrupar os N indivíduos de acordo com seus escores (ou soma de pontos na prova, ou teste), em g=1,2,...,G grupos. No caso dicotômico serão possíveis g=1,2,...,L-1 grupos formados por indivíduos de mesmo escore no teste.

Quando sabemos que um determinado número de ciclos é necessário para a convergência? O critério para a convergência é arbitrário; Wood, Wingersky e Lord (1976) usaram a própria função de verossimilhança para determinar a convergência. Isto requer as estimativas de todos os parâmetros para substituir na equação dada em (4.4) para cada ciclo iterativo, interrompendo quando $|\ell_t - \ell_{t-1}| \leq \xi$.

Logo, a diferença em probabilidade entre dois sucessivos ciclos tem que ser menor ou igual que um específico valor ($\xi$) para terminar o processo.

Se o "critério de parada" não for estabelecido, as estimativas dos parâmetros do item e do parâmetro de habilidade serão "acumulados" caminhando para o infinito. Ou seja, quando o número de ciclos aumenta a convergência nunca será atingida.

## 4.2- MÉTRICA DA HABILIDADE ($\theta$)

Sendo a locação e a escala da métrica das habilidades não conhecidas, existem somente (N-2) parâmetros $\hat{\theta}_j$ independentes, ainda que N valores serão obtidos. O programa LOGIST calcula a média e a variância dos valores $\hat{\theta}_j$ dos indivíduos e faz uma padronização do escore obtendo a transformação $\hat{\theta}_j^* = \frac{(\hat{\theta}_j - \bar{\hat{\theta}})}{S_{\hat{\theta}}}$ para cada ciclo do processo iterativo de Newton-Raphson.

Para que a probabilidade $P_i(\theta_j)$ não se altere é necessário ajustar os parâmetros do item para a mesma escala usando as seguintes equações:

$$\hat{b}_i^* = \frac{(\hat{b}_i - \bar{\hat{\theta}})}{S_{\hat{\theta}}} \quad \text{e} \quad \hat{a}_i^* = S_{\hat{\theta}}\,\hat{a}_i .$$

A métrica das habilidades estimadas $\hat{\theta}_j$ é única somente fazendo uma transformação linear (Lord e Novick (1968), Wood, Wingersky e Lord (1976)).

Segundo Fernández (1990), podemos fazer algumas transformações de θ para os modelos logísticos de um, dois e três parâmetros desde que estas transformações sejam feitas também nos parâmetros dos itens, da seguinte forma:

→ se for adicionado ou multiplicado (ou ambos) uma constante em θ o mesmo deve ser feito no parâmetro de dificuldade do item, $\beta$;

→ o parâmetro de discriminação do item, $\alpha$, deve ser dividido pela constante multiplicada em θ e;

→ o parâmetro de adivinhação do item, $\psi$, se mantém.

Se fizermos uma transformação desta maneira, estaremos realizando uma transformação admissível e $P_i(\theta_j)$ se manterá inalterado. Por exemplo:

$$P_i(\theta_j^*) = \psi_i^* + \frac{(1-\psi_i^*)\ \exp\left[D\alpha_i^*(\theta_j^* - \beta_i^*)\right]}{1+\exp\left[D\alpha_i^*(\theta_j^* - \beta_i^*)\right]}$$

$$= \psi_i + \frac{(1-\psi_i)\exp\left\{DS_\theta\alpha_i\left[(\theta_j - \bar{\theta})/S_\theta - (\beta_i - \bar{\theta})/S_\theta\right]\right\}}{1+\exp\left\{DS_\theta\alpha_i\left[(\theta_j - \bar{\theta})/S_\theta - (\beta_i - \bar{\theta})/S_\theta\right]\right\}}$$

$$= \psi_i + \frac{(1-\psi_i)\exp\left\{D\alpha_i\left[(\theta_j-\bar{\theta})-(\beta_i-\bar{\theta})\right]\right\}}{1+\exp\left\{D\alpha_i\left[(\theta_j-\bar{\theta})-(\beta_i-\bar{\theta})\right]\right\}}$$

$$= \psi_i + \frac{(1-\psi_i)\exp\left\{D\alpha_i\left[\theta_j-\beta_i\right]\right\}}{1+\exp\left\{D\alpha_i\left[\theta_j-\beta_i\right]\right\}} = P_i(\theta_j).$$

## 4.3- EMV CONJUNTA DOS PARÂMETROS - ITENS CATEGÓRICOS

Considerando que indivíduos podem ter habilidades próximas umas das outras, podemos agrupar essas habilidades em G grupos, g=1,2,...,G, onde $\theta_g$ será o ponto médio do intervalo de habilidade do grupo g. O número de indivíduos no g-ésimo grupo de habilidade é dado por $N_g$ e, dentro de cada grupo o número de indivíduos que respondem em cada uma das categorias de resposta do i-ésimo item é dado por $r_{g1}, r_{g2},\ldots,r_{gk},\ldots,r_{gm}$. Considerando apenas um determinado item do teste que tenha m categorias, k=1,2,...,m. A matriz de observações será dada por $R_{(Gxm)}$, onde as linhas representam os grupos de indivíduos que tem habilidades próximas e as colunas representam as categorias do item.

$$R_{(Gxm)} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1k} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2k} & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \vdots & & \vdots & & \vdots \\ r_{g1} & r_{g2} & \cdots & r_{gk} & \cdots & r_{gm} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \vdots & & \vdots & & \vdots \\ r_{G1} & r_{G2} & \cdots & r_{Gk} & \cdots & r_{Gm} \end{bmatrix} \begin{matrix} \rightarrow N_1 \\ \rightarrow N_2 \\ \vdots \\ \vdots \\ \rightarrow N_g \\ \vdots \\ \vdots \\ \rightarrow N_G \end{matrix}$$

Portanto, $\sum_{g=1}^{G} N_g = N$, $\sum_{k=1}^{m} r_{gk} = N_g$ e $p_{gk} = r_{gk} / N_g$, sendo

$p_{gk}$ a proporção observada de escolha da k-ésima categoria de resposta por indivíduos do g-ésimo grupo e considerando que em cada grupo os $N_g$ indivíduos devem selecionar uma das categorias de resposta do i-ésimo item do teste,

$$\sum_{k=1}^{m} p_{gk} = 1 \Rightarrow p_{gm} = 1 - \sum_{k=1}^{m-1} p_{gk}.$$

Assim, a probabilidade de um indivíduo com habilidade $\theta_g$ cair na k-ésima categoria de resposta será denotada por $P_k(\theta_g)$ e,

$$\sum_{k=1}^{m} P_k(\theta_g) = 1 \Rightarrow P_m(\theta_g) = 1 - \sum_{k=1}^{m-1} P_k(\theta_g)$$

Considerando o Modelo de Resposta Categórica de Samejima (1972) dado em (3.3) e uma transformação da forma $Z_{gi} = \alpha_i(\theta_g - \lambda_{iS}) = -\alpha_i \lambda_{iS} + \alpha_i \theta_g = \zeta_{iS} + \alpha_i \theta_g$, onde $\alpha_i$ e $\zeta_{iS}$ são, o coeficiente angular da reta e o intercepto, respectivamente. Para simplificar a notação passaremos a usar $P_k(\theta_g) = P_{gk}$ e $P^*_k(\theta_g) = P^*_{gk}$.

Portanto de (3.2) a probabilidade acumulada do indivíduo de habilidade $\theta_g$ responder na categoria k ou maior será dada por:

$P^*_{gk} = 1/(1+e^{-(\zeta_S+\alpha\theta_g)})$, sendo s=k-1.

Para um dado item, temos a seguinte função de verossimilhança, $P(R/\zeta_s,\alpha,\theta) = \prod_{g=1}^{G} \frac{N_g!}{r_{g1}!r_{g2}!\ldots r_{gm}!} \prod_{k=1}^{m} P_{gk}^{r_{gk}}$ (4.12)

onde R é a matriz (Gxm) de $r_{gk}$ e (4.12) é o produto de multinomiais com parâmetros $N_g$, $P_{g1},\ldots,P_{g2},\ldots,P_{gm}$, g=1,2,...,G, onde $\sum_{k=1}^{m} r_{gk} = N_g$ e $\sum_{k=1}^{m} P_{gk} = 1$. Tomando logaritmo natural,

$$\ell = \text{constante} + \sum_{g=1}^{G}\sum_{k=1}^{m} r_{gk} \ln P_{gk} \qquad (4.13)$$

Substituindo $P_{gk} = P^*_{gk} - P^*_{g(k+1)}$ de (3.1) em (4.13),

$$\ell = \text{constante} + \sum_{g=1}^{G}\sum_{k=1}^{m} r_{gk} \ln\left(P^*_{gk} - P^*_{g(k+1)}\right).$$

Derivando a função $\ell$ com relação aos parâmetros $\alpha$ e $\zeta_S$, s=1,2,...,m-1 e usando: $r_{gk}=r_k$, $P^*_{gk} = P^*_k$ e $P^*_{g(k+1)} = P^*_{k+1}$, para simplificar a notação excluímos o índice g, teremos:

$$\frac{\partial \ell}{\partial \zeta_S} = \frac{\partial}{\partial \zeta_S} \sum_{g=1}^{G}\sum_{k=1}^{m} r_k \ln\left(P^*_k - P^*_{k+1}\right)$$

$$= \sum_{g=1}^{G} \sum_{k=1}^{m} \frac{r_k}{\left(P_k^* - P_{k+1}^*\right)} \quad \frac{\partial\left(P_k^* - P_{k+1}^*\right)}{\partial \zeta_S}$$

Desmembrando o somatório acima identificaremos melhor as parcelas que são constantes no somatório ao derivarmos a função $\ell$ em relação a $\zeta_S$, s=1,2,...,$m_i$-1.

$$\sum_{k=1}^{m} r_k \ln\left(P_k^* - P_{k+1}^*\right) = r_1 \ln\left(1 - P_2^*\right) + r_2 \ln\left(P_2^* - P_3^*\right) + \ldots$$

$$+ r_{k-1} \ln\left(P_{k-1}^* - P_k^*\right) \quad + \quad r_k \ln\left(P_k^* - P_{k+1}^*\right) + r_{k+1} \ln\left(P_{k+1}^* - P_{k+2}^*\right) +$$

$$\ldots + r_m \ln\left(P_m^* - 0\right).$$

Assim a derivada da função $\ell$ com respeito a um parâmetro $\zeta_S$ envolverá dois termos adjacentes no somatório acima, pois, de (3.2) o limite $\zeta_S$ está entre as categorias k-1 e k.

$$\frac{\partial \ell}{\partial \zeta_s} = \sum_{g=1}^{G} \left[ \frac{r_{k-1}}{\left(P_{k-1}^* - P_k^*\right)} \frac{\partial\left(P_{k-1}^* - P_k^*\right)}{\partial \zeta_s} + \frac{r_k}{\left(P_k^* - P_{k+1}^*\right)} \frac{\partial\left(P_k^* - P_{k+1}^*\right)}{\partial \zeta_s} \right]$$

$$\frac{\partial \ell}{\partial \zeta_s} = \sum_{g=1}^{G} \left[ \frac{r_{k-1}}{\left(P_{k-1}^* - P_k^*\right)} \left( - \frac{\partial P_k^*}{\partial \zeta_s} \right) + \frac{r_k}{\left(P_k^* - P_{k+1}^*\right)} \left( \frac{\partial P_k^*}{\partial \zeta_s} \right) \right]$$

$$\frac{\partial \ell}{\partial \zeta_s} = \sum_{g=1}^{G} \left[ \frac{\partial P_k^*}{\partial \zeta_s} \left( \frac{r_k}{\left(P_k^* - P_{k+1}^*\right)} - \frac{r_{k-1}}{\left(P_{k-1}^* - P_k^*\right)} \right) \right]$$

$$\frac{\partial \ell}{\partial \zeta_s} = \sum_{g=1}^{G}\left[ - P_k^* Q_k^* \left( \frac{r_k}{P_k} - \frac{r_{k-1}}{P_{k-1}} \right) \right]$$

$$\frac{\partial \ell}{\partial \zeta_s} = \sum_{g=1}^{G}\left[ P_k^* Q_k^* \left( \frac{r_{k-1}}{P_{k-1}} - \frac{r_k}{P_k} \right) \right],$$

seja $w_k = P_k^* Q_k^*$, $r_k = N_g p_{gk}$ e $r_{k-1} = N_g p_{g(k-1)}$.

$$\frac{\partial \ell}{\partial \zeta_s} = \sum_{g=1}^{G} w_k N_g \left[ \frac{p_{g(k-1)}}{P_{k-1}} - \frac{p_{gk}}{P_k} \right] \qquad (4.14)$$

Logo, derivando a função $\ell$ com respeito ao parâmetro $\alpha$ teremos,

$$\frac{\partial \ell}{\partial \alpha} = \sum_{g=1}^{G} \sum_{k=1}^{m} \frac{r_k}{\left( P_k^* - P_{k+1}^* \right)} \frac{\partial \left( P_k^* - P_{k+1}^* \right)}{\partial \alpha},$$

sendo $\dfrac{\partial \left( P_k^* - P_{k+1}^* \right)}{\partial \alpha} = \theta_g P_k^* Q_k^* - \theta_g P_{k+1}^* Q_{k+1}^*$

$$\frac{\partial \ell}{\partial \alpha} = \sum_{g=1}^{G} \theta_g \sum_{k=1}^{m} \frac{r_k \left( P_k^* Q_k^* - P_{k+1}^* Q_{k+1}^* \right)}{\left( P_k^* - P_{k+1}^* \right)},$$

sendo $P_k^* Q_k^* = w_k$, $P_{k+1}^* Q_{k+1}^* = w_{k+1}$ e $r_k = N_g p_{gk}$

$$\frac{\partial \ell}{\partial \alpha} = \sum_{g=1}^{G} \theta_g N_g \sum_{k=1}^{m} \frac{p_{gk}}{P_k} \left( w_k - w_{k+1} \right) \qquad (4.15)$$

Usando a equação do método de Newton-Raphson dada em (4.7), teremos:

$$\begin{bmatrix} \hat{\zeta}_1 \\ \hat{\zeta}_2 \\ \vdots \\ \vdots \\ \hat{\zeta}_{m_i-1} \\ \hat{\alpha} \end{bmatrix}_{t+1} = \begin{bmatrix} \hat{\zeta}_1 \\ \hat{\zeta}_2 \\ \vdots \\ \vdots \\ \hat{\zeta}_{m_i-1} \\ \hat{\alpha} \end{bmatrix}_t - \begin{bmatrix} \frac{\partial^2 \ell}{\partial^2 \zeta_1} & \frac{\partial^2 \ell}{\partial \zeta_1 \partial \zeta_2} & \cdots & & & \\ \frac{\partial^2 \ell}{\partial \zeta_2 \partial \zeta_1} & \frac{\partial^2 \ell}{\partial^2 \zeta_2} & \frac{\partial^2 \ell}{\partial \zeta_2 \partial \zeta_3} & & & \\ \cdots & \frac{\partial^2 \ell}{\partial \zeta_3 \partial \zeta_2} & \frac{\partial^2 \ell}{\partial^2 \zeta_3} & \frac{\partial^2 \ell}{\partial \zeta_3 \partial \zeta_4} \cdots & & \\ & & \vdots & \ddots & & \\ & & & & \frac{\partial^2 \ell}{\partial^2 \zeta_{m_i-1}} & \frac{\partial^2 \ell}{\partial \zeta_{m_i-1} \partial \alpha} \\ & & & & \cdots & \frac{\partial^2 \ell}{\partial^2 \alpha} \end{bmatrix}_t^{-1} \begin{bmatrix} \frac{\partial \ell}{\partial \zeta_1} \\ \frac{\partial \ell}{\partial \zeta_2} \\ \vdots \\ \frac{\partial \ell}{\partial \zeta_{m_i-1}} \\ \frac{\partial \ell}{\partial \alpha} \end{bmatrix}_t$$

As derivadas de segunda ordem são dadas por:

$$\frac{\partial^2 \ell}{\partial^2 \zeta_S} = \sum_{g=1}^{G} \left[ \left( \frac{N_g P_{g(k-1)}}{P_{k-1}} - \frac{N_g P_{gk}}{P_k} \right) \left( Q_k^* - P_k^* \right) w_k \right]$$

$$+ \left[ \left( - \frac{N_g P_{g(k-1)}}{P_{k-1}^2} - \frac{N_g P_{gk}}{P_k^2} \right) w_k^2 \right]$$

$$\frac{\partial^2 \ell}{\partial^2 \alpha} = \sum_{g=1}^{G} \sum_{k=1}^{m} N_g \frac{P_{gk}}{P_k} \frac{\partial^2 P_k}{\partial^2 \alpha} - \sum_{g=1}^{G} \sum_{k=1}^{m} N_g \frac{P_{gk}}{P_k^2} \left( - w_{k+1} + w_k \right)^2 \theta_g^2$$

Na matriz de derivadas de segunda ordem com respeito aos parâmetros só existirá a diagonal principal, onde os termos são dados por (4.14) para s=1,2,...,$m_i$-1 e por (4.15), e as diagonais secundárias acima e abaixo da diagonal principal, onde os termos são dados pelas derivadas parciais de segunda ordem. Isso ocorre porque o parâmetro $\zeta_s$ está entre duas categorias adjacentes.

As proporções observadas ($p_{gk}$) de indivíduos respondendo na k-ésima categoria serão substituídas por suas esperanças, desta forma estaremos utilizando o método do escore de Fisher.

Então a equação do escore de Fisher para ser resolvida iterativamente para os parâmetros das Curvas Características Limites (CCL) para um item de resposta categórica será,

$$\begin{bmatrix} \hat{\zeta}_1 \\ \hat{\zeta}_2 \\ \vdots \\ \vdots \\ \hat{\zeta}_{m_i-1} \\ \hat{\alpha} \end{bmatrix}_{t+1} = \begin{bmatrix} \hat{\zeta}_1 \\ \hat{\zeta}_2 \\ \vdots \\ \vdots \\ \hat{\zeta}_{m_i-1} \\ \hat{\alpha} \end{bmatrix}_{t} - \begin{bmatrix} E\left(\frac{\partial^2 \ell}{\partial^2 \zeta_1}\right) & E\left(\frac{\partial^2 \ell}{\partial \zeta_1 \partial \zeta_2}\right) & \cdots & & \\ E\left(\frac{\partial^2 \ell}{\partial \zeta_2 \partial \zeta_1}\right) & E\left(\frac{\partial^2 \ell}{\partial^2 \zeta_2}\right) & E\left(\frac{\partial^2 \ell}{\partial \zeta_2 \partial \zeta_3}\right) & \cdots & \\ \vdots & E\left(\frac{\partial^2 \ell}{\partial \zeta_3 \partial \zeta_2}\right) & E\left(\frac{\partial^2 \ell}{\partial^2 \zeta_3}\right) & E\left(\frac{\partial^2 \ell}{\partial \zeta_3 \partial \zeta_4}\right) & \cdots \\ & \vdots & & & \ddots \\ & & & E\left(\frac{\partial^2 \ell}{\partial^2 \zeta_{m_i-1}}\right) & E\left(\frac{\partial^2 \ell}{\partial \zeta_{m_i-1} \partial \alpha}\right) \\ & & & \cdots & E\left(\frac{\partial^2 \ell}{\partial^2 \alpha}\right) \end{bmatrix}_t^{-1} \begin{bmatrix} \frac{\partial \ell}{\partial \zeta_1} \\ \frac{\partial \ell}{\partial \zeta_2} \\ \vdots \\ \vdots \\ \frac{\partial \ell}{\partial \zeta_{m_i-1}} \\ \frac{\partial \ell}{\partial \alpha} \end{bmatrix}_t$$

Resolvendo a equação acima para cada item, i=1,2,...,L, teremos as estimativas dos parâmetros de todas as categorias de todos os itens.

Para voltar ao parâmetro $\lambda_k$ para interpretação, fazemos $\lambda_S = -\frac{\zeta_S}{\alpha}$ então $\lambda_k = (\lambda_{S+1} + \lambda_S)/2$ e s=(k-1).

## 4.3.1- EMV CONJUNTA DOS PARÂMETROS DOS INDIVÍDUOS

Como na seção 4.1, para itens dicotômicos, aqui também vamos considerar os parâmetros dos itens, estimados na seção 4.3, como verdadeiros. Assim podemos usar a MVConjunta para estimar as habilidades dos indivíduos.

A função de verossimilhança para um dado indivíduo com habilidade $\theta_j$ é a verossimilhança de um particular vetor de resposta de itens $X_j=(X_{1k},X_{2k},\ldots,X_{Lk})$, onde $x_{ik}=1$ se o indivíduo escolhe a k-ésima categoria de resposta para o i-ésimo item e, $x_{ik}=0$ caso contrário.

$$P(X_j|\zeta,\alpha,\theta_j)=\prod_{i=1}^{L}\prod_{k=1}^{m_i}P_{ik}^{x_{ik}}. \qquad (4.16)$$

A equação (4.16) é o produto de Binomiais com parâmetros $(m_i, P_{ik})$, i=1,2,...,L. Tomando o logaritmo natural em (4.16) e derivando com respeito ao parâmetro $\theta_j$, temos

$$\ell = \ln P\left(X_j|\zeta,\alpha,\theta_j\right)=\sum_{i=1}^{L}\sum_{k=1}^{m_i}x_{ik}\ln P_{ik}\ , \text{ e}$$

$$\frac{\partial\ell}{\partial\theta_j}=\sum_{i=1}^{L}\sum_{k=1}^{m_i}\frac{x_{ik}}{\left(P^*_{i,k}-P^*_{i,k+1}\right)}\frac{\partial\left(P^*_{i,k}-P^*_{i,k+1}\right)}{\partial\theta_j}$$

$$\text{sendo } \frac{\partial\left(P^{*}_{i,k} - P^{*}_{i,k+1}\right)}{\partial\theta_j} = \alpha_i\left(w_{i,k} - w_{i,k+1}\right)$$

$$\frac{\partial\ell}{\partial\theta_j} = \sum_{i=1}^{L}\sum_{k=1}^{m_i} x_{ik}\alpha_i\left[\frac{w_{i,k} - w_{i,k+1}}{P_{ik}}\right], \qquad (4.17)$$

e a segunda derivada com respeito ao parâmetro $\theta_j$, fica

$$\frac{\partial\ell^2}{\partial\theta_j} = \sum_{i=1}^{L}\alpha_i^2\sum_{k=1}^{m_i} x_{ik}\left[\frac{-w_{ik}\left(Q^{*}_{ik} - P^{*}_{ik}\right)}{P_{ik}} + w_{k-1}\frac{\left(Q^{*}_{i,k-1} - P^{*}_{i,k-1}\right)}{P_{ik}} - \frac{\left(w_{i,k-1} - Q_{ik}\right)^2}{P^2_{ik}}\right] \qquad (4.18)$$

Usando (4.17) e (4.18) existirá uma equação do escore de Fisher para cada indivíduo, da forma

$$\hat{\theta}_{j(t+1)} = \hat{\theta}_{j(t)} - \left[\frac{\partial\ell/\partial\theta_j}{\partial^2\ell/\partial^2\theta_j}\right]_{(t)}, \quad j=1,2,\ldots,N, \qquad (4.19)$$

onde t denota a t-ésima iteração.

O procedimento de máxima verossimilhança conjunta será empregado da mesma maneira que para o caso de respostas dicotômicas. As suposições de independência entre determinados parâmetros feitas anteriormente e o critério de parada do processo iterativo de Newton Raphson também são válidos aqui. O escore bruto padronizado do teste é usado como valor inicial da habilidade $\theta$ e as estimativas das habilidades são padronizadas em média zero e

variância um para facilitar a identificação e a solução de possíveis problemas.

O procedimento de EMVConjunta para este caso de respostas categorizadas foi implementado no programa LOGOG, Kolakowski e Bock (1973 b).

Na seção 4.4 usaremos Estimação de Máxima Verossimilhança Marginal (EMVM) para estimar os parâmetros sob o Modelo de Crédito Parcial (MCP).

## 4.4- EMV MARGINAL DOS PARÂMETROS SOB O MODELO DE CREDITO PARCIAL

No momento, o melhor programa para estimação de parâmetros dos modelos de respostas categóricas disponível, é o MULTILOG, versão 6, (Thissen, 1991). Este é um programa generalizado que emprega Estimação de Máxima Verossimilhança Marginal (EMVM) para obter as estimativas dos parâmetros do item, para o Modelo de Resposta Categórica de Samejima (1969), para o Modelo Nominal de Bock (1972) e para Modelos de Múltipla Escolha, Thissen e Steinberg (1984). Outros modelos também podem ser analisados pelo MULTILOG, adequando os modelos citados.

Quando a EMVM é usada as estimativas das habilidades dos indivíduos são obtidas em um estágio separado.

Como o objetivo desse trabalho é analisar um conjunto de itens onde o Modelo de Crédito Parcial (MCP) é adequado e, o programa disponível para estimar os parâmetros de interesse é o MULTILOG, nessa seção vamos desenvolver o procedimento de EMVM para o MCP.

A máxima verossimilhança marginal aproximada por Bock e Lieberman (1970) integra sobre a distribuição da habilidade $\theta$ e estima os parâmetros do item na distribuição marginal dos dados. No

entanto, a estimação simultânea dos parâmetros de L itens é prático somente para testes com L pequeno.

Uma subsequente aproximação considerada uma reformulação da EMVM de Bock e Lieberman (1970), é dada por Bock e Aitkin (1981) com a estrutura do algoritmo EM.

No caso do modelo de resposta categórica de Samejima os parâmetros de dificuldade são os limites entre as categorias e são baseados em probabilidades acumuladas, por isso não podem ser separados dos parâmetros dos indivíduos no modelo. No MCP os parâmetros de dificuldade são os independentes passos de uma categoria a outra em cada item, e assim, podem ser separados no modelo. Esta separabilidade dos parâmetros no MCP é que resulta em estatísticas suficientes para as habilidades dos indivíduos e para os parâmetros de dificuldade dos passos dos itens. E, os parâmetros dos indivíduos são eliminados da equação para a estimação dos parâmetros dos itens independente da amostra de indivíduos avaliada.

Como o MCP é uma extensão do modelo dicotômico de Rasch, representaremos os créditos parciais recebidos pelo j-ésimo indivíduo no i-ésimo item da seguinte maneira, mostrada no Quadro 4.1 a seguir, para i=1.

| *categorias* ⇒ | 1 | | 2 | | 3 | | 4 | | 5 | | 6 | $escore(x_{ji})$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *passos* ⇒ | | *passo*1 | | *passo*2 | | *passo*3 | | *passo*4 | | *passo*5 | | ⇓ |
| *indivíduos* ⇓ | | | | | | | | | | | | |
| 1 | 1 | → | 1 | → | 1 | → | 1 | → | 1 | → | 1 | $x_{11} = 5$ |
| 2 | 1 | → | 1 | → | | → | | → | | → | | $x_{21} = 1$ |
| 3 | 1 | → | 1 | → | 1 | → | 1 | → | | → | | $x_{31} = 3$ |
| 4 | 1 | → | | → | | → | | → | | → | | $x_{41} = 0$ |
| 5 | 1 | → | 1 | → | 1 | → | 1 | → | 1 | → | | $x_{51} = 4$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| *N* | 1 | → | 1 | → | | → | | → | | → | | $x_{N1} = 1$ |
| $s_{ik}$ | $s_{11}$ | | $s_{12}$ | | $s_{13}$ | | $s_{14}$ | | $s_{15}$ | | $s_{16}$ | |

**QUADRO 4.1**- Representação dos créditos parciais ($x_{ji}$) dos indivíduos e contagem dos indivíduos ($s_{ik}$) que completam cada passo do item i=1.

Teremos então, L quadros representando os créditos parciais dos indivíduos nos itens do teste.

Seja $S_j = \sum_{i=1}^{L} x_{ji}$, j=1,2,...,N, onde $x_{ji}$ é o número de passos completados no item i pelo indivíduo j. Logo, $S_j$ é o escore total do indivíduo j, e este escore total é uma estatística suficiente para $\theta_j$.

Agora seja $s_{ik}$ como a contagem de todos os indivíduos que completam cada passo no item i, onde x=1,2,..,$m_i$-1.

Podemos perceber que o número de indivíduos que completam o passo (x+1) é sempre menor ou igual ao número de indivíduos que completam o passo x, ou seja, $s_{i1} \geq s_{i2} \geq \ldots\ldots \geq s_{i(mi)}$.

Pode ser verificado que $s_{ik}$ é uma estatística suficiente para $\beta_{ik}$, para i=1,2,...,L e k=1,2,...,$m_i$.

Masters(1982) mostra a separabilidade desses parâmetros no MCP e o assegura como membro da família Rasch.

A seguir, vamos ver como estimar os parâmetros dos itens pela Estimação de Máxima Verossimilhança Marginal (EMVM).

## 4.4.1- SOLUÇÃO DE BOCK E AITKIN

A Estimação de Máxima Verossimilhança Marginal (EMVM) é usada porque trabalhamos com uma variável aleatória contínua não observável (habilidade θ). Ao contrário da EMVC, onde estimamos parâmetros dos itens e dos indivíduos simultaneamente, baseado nas respostas dos indivíduos aos itens, aqui consideramos que θ tem uma distribuição de probabilidade contínua conhecida. Assim a probabilidade conjunta (θ,**x**), (onde $\mathbf{x}=x_1,x_2,\ldots,x_L$ representa o padrão de respostas aos itens existentes nos dados) será dado por,

$$P(\theta,\mathbf{x}/\underline{\beta_i})=P(\mathbf{x}/\theta,\underline{\beta_i})\,\phi(\theta), \qquad (4.20)$$

onde $\underline{\beta_i}$ é o vetor de parâmetros dos itens $\underline{\beta_i}=\left(\beta_{i1},\ldots,\beta_{i(m_i-1)}\right)$ e $P(\mathbf{x}/\theta,\underline{\beta_i})=\prod_{i=1}^{L} P_i(x_i)$ é a probabilidade condicional do padrão de resposta **x** dado a habilidade θ e o vetor de parâmetros $\underline{\beta_i}$.

Integrando essa função de densidade de probabilidade conjunta de (4.20) com respeito a variável aleatória não observável θ, obteremos a função densidade marginal dos dados,

$$P(\mathbf{x})=\int_{-\infty}^{\infty}\prod_{i=1}^{L} P_i(x_i)\phi(\theta)d\theta, \qquad (4.21)$$

onde $P_i(x_i)$ é a probabilidade do indivíduo de habilidade $\theta$ obter escore x no item i e, é a probabilidade dada em (3.8),

$$P_i(x_i) = P_{ix}(\theta_j) = \frac{exp\left[\sum_{n=0}^{x}(\theta_j - \beta_{in})\right]}{\sum_{k=0}^{m_i} exp\left[\sum_{n=0}^{k}(\theta_j - \beta_{in})\right]}, \quad x=0,1,2,\ldots,m_i-1,$$

que é a probabilidade do indivíduo de habilidade $\theta_j$ obter escore x no item i sob o Modelo de Crédito Parcial de Masters (1982).

Percebemos que, dessa forma, as estimativas dos parâmetros dos itens estão livres das estimativas das habilidades de cada indivíduo. Elas dependem da distribuição da habilidade $\phi(\theta)$ que é identificada antes de os dados serem coletados. Geralmente $\phi(\theta)$ é escolhida como a função densidade de probabilidade N(0,1), chamada de distribuição de habilidade a priori.

A verossimilhança, considerando todos os padrões de resposta **x** existentes nos dados, assumindo que os itens sejam independentes, os indivíduos sejam independentes e que itens e indivíduos sejam independentes, fica

$$\wp = C\prod_{\mathbf{x}}[P(\mathbf{x})]^{r_{\mathbf{x}}} \qquad (4.22)$$

onde C é uma constante de proporcionalidade e $r_x$ é a quantidade de vezes que o padrão **x** ocorre nos dados. Maximizar (4.22) é equivalente a maximizar o logaritmo da função de verossimilhança,

$$\ell = \sum_{\mathbf{x}} r_{\mathbf{x}} \ln P(\mathbf{x}). \qquad (4.23)$$

A função de verossimilhança marginal (4.22), pode ser resolvida pelo método de Newton-Raphson/Fischer para estimar os $\sum_{i=1}^{L}(m_i - 1)$ parâmetros dos itens como foi feito na seção 4.3 desse Capítulo. No caso da aplicação do Capítulo 6 desse trabalho teremos L=12 e $m_i$=6 para qualquer item. Então teremos, $\sum_{i=1}^{12}(m_i - 1) = 60$ parâmetros a serem estimados. Nesse caso será necessário a inversão de uma matriz de informação 60x60. Por essa dificuldade computacional e pelo fato de que este método necessita que a distribuição da habilidade populacional seja amostrada para ser conhecida antecipadamente, não é recomendável o uso desse método.

Alternativamente, Bock e Aitkin (1981) sugeriram a utilização do "algorítmo EM" (Dempster, Laird & Rubin, 1977). O algorítmo é um procedimento iterativo usado para encontrar estimativas de máxima verossimilhança dos parâmetros dos itens de modelos de probabilidade, na presença de variáveis aleatórias não observáveis ($\theta$).

#### 4.4.1.1- ALGORITMO EM

Sob a suposição de que a população é composta de indivíduos que são membros de Q "classes" discretas com valores $\theta_1$, $\theta_2,\ldots,\theta_Q$ sob a variável latente e, $r^*_{iqx}$ - definido como o número de indivíduos na classe $\theta_q$ obtendo escore x sob o item i - é uma quantidade desconhecida. Dessa forma, o passo E do algoritmo EM, consiste em encontrar o valor esperado de $r^*_{iqx}$ condicionado aos dados e usando "estimativas provisórias" dos parâmetros do item. Assim,

$$E\left(r^*_{iqx} \mid dados; \left\{\hat{\beta}_{i1}, \hat{\beta}_{i2}, \ldots \hat{\beta}_{i(m_i-1)}\right\}\right) = \sum_{\mathbf{x}\in\xi} r_{\mathbf{x}} \left[\frac{P(\mathbf{x};\theta_q)}{\sum_{q=1}^{Q} P(\mathbf{x};\theta_q)}\right] \qquad (4.24)$$

onde $\xi$ representa o conjunto de padrões $\mathbf{x}$ no qual $x_i=x$, $x=0,1,\ldots,m_i-1$ e,

$$P\left(\mathbf{x}\ ;\theta_q\right) = \prod_{i=1}^{L} P_i(x_i)\phi(\theta_q). \qquad (4.25)$$

Por exemplo, para o primeiro item da aplicação do Capítulo 6, onde L=12 e $m_1=6$ teremos a matriz $R_1^*$ dada por,

$$R^*_{1(Q\times m_1)} = \begin{array}{c|cccccc} x \Rightarrow & 0 & 1 & 2 & 3 & 4 & 5 \\ q \Downarrow & & & & & & \\ 1 & r^*_{110} & r^*_{111} & r^*_{112} & r^*_{113} & r^*_{114} & r^*_{115} \\ 2 & r^*_{120} & r^*_{121} & r^*_{122} & \left(r^*_{123}\right) & r^*_{124} & r^*_{125} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Q & r^*_{1Q0} & r^*_{1Q1} & r^*_{1Q2} & r^*_{1Q3} & r^*_{1Q4} & r^*_{1Q5} \end{array}$$

Assim, de (4.24), teremos,

$$E\left(r^*_{123} \mid dados; \left\{\hat{\beta}_{11}, \hat{\beta}_{12}, \ldots \hat{\beta}_{15}\right\}\right) = \sum_{\mathbf{x}\in\xi} r_{\mathbf{x}} \left[\frac{P(\mathbf{x};\theta_2)}{\sum_{q=1}^{Q} P(\mathbf{x};\theta_q)}\right]$$

onde $\xi$ representa o conjunto de padrões onde $x_1=3$, ou seja, todos os padrões em que o escore obtido no item 1 é 3.

O passo E vai produzir L matrizes $R^*_{i(Qxm_i)}$. Os $r^*_{iqx}$ são chamados dados artificiais de contagem e estes são usados no passo M que consiste em maximizar o logaritmo da função de verossimilhança para encontrar as estimativas dos parâmetros do item. Usando os dados obtidos em $R^*_i$, a verossimilhança para o item i é

$$\ell_i \approx \sum_{x=0}^{m_i-1} \sum_{q=1}^{Q} r^*_{iqx} \ln P_{ix}(\theta_q), \qquad (4.26)$$

onde $P_{ix}\left(\theta_q\right)$ é dado em (3.8).

$$\frac{\partial \ell_i}{\partial \underline{\beta_i}} = \sum_{x=0}^{m_i-1} \sum_{q=1}^{Q} \frac{r^*_{iqx}}{P_{ix}(\theta_q)} \frac{\partial P_{ix}(\theta_q)}{\partial \underline{\beta_i}} = \mathbf{0}, \quad \text{onde } \underline{\beta_i} \text{ é o vetor de}$$

parâmetros do item i.

Precisamos então da derivada de $P_{ix}(\theta_q)$ em relação ao vetor de parâmetros $\underline{\beta_i}$, ou seja, em relação a $\beta_{ix}$ para $x=1,2,\ldots,m_i-1$.

Logo, teremos $m_i-1$ equações para serem resolvidas iterativamente.

$$\sum_{x=0}^{m_i-1} \sum_{q=1}^{Q} \frac{r^*_{iqx}}{P_{ix}(\theta_q)} \frac{\partial P_{ix}(\theta_q)}{\partial \beta_{i1}} = 0$$

$$\sum_{x=0}^{m_i-1} \sum_{q=1}^{Q} \frac{r^*_{iqx}}{P_{ix}(\theta_q)} \frac{\partial P_{ix}(\theta_q)}{\partial \beta_{i2}} = 0$$

$$\sum_{x=0}^{m_i-1} \sum_{q=1}^{Q} \frac{r^*_{iqx}}{P_{ix}(\theta_q)} \frac{\partial P_{ix}(\theta_q)}{\partial \beta_{i3}} = 0$$

.....

$$\sum_{x=0}^{m_i-1} \sum_{q=1}^{Q} \frac{r^*_{iqx}}{P_{ix}(\theta_q)} \frac{\partial P_{ix}(\theta_q)}{\partial \beta_{i,m_i-1}} = 0,$$

onde as $\frac{\partial P_{ix}(\theta_q)}{\partial \beta_{ix}}$ são as derivadas de 3.8 em relação aos parâmetros de dificuldade dos passos.

As estimativas $\hat{\beta}_{ix}$, para x=1,2,...$m_i$-1, obtidas aqui serão usadas no próximo ciclo do algoritmo EM, assim os ciclos EM são repetidos até que os parâmetros se estabilizem ou até que um número de ciclos seja alcançado (o programa MULTILOG usa 25 ciclos, podendo ser alterado de acordo com as necessidades do usuário).

Dessa forma serão encontrados os parâmetros de dificuldade dos passos para cada item i, i=1,2,...,L, sob o MCP de Masters.

## ESTIMATIVAS DOS PARÂMETROS DOS INDIVÍDUOS

Na EMVM os valores de $\theta$ para cada indivíduo que produz um certo padrão de resposta **x**, são estimados separadamente e os parâmetros dos itens (já estimados anteriormente pela EMVM/EM) são usados como valores verdadeiros. Dado então $\hat{\beta}_{i1},\hat{\beta}_{i2},\ldots,\hat{\beta}_{i,m_i-1}$, para i=1,2,...,L a densidade posterior de $\theta$ será

$$P(\mathbf{x};\theta)=\prod_{i=1}^{L} P_i(x_i)\phi(\theta) \qquad (4.30)$$

onde $P_i(x_i)$ é dado em 3.8. Assim, (4.30) descreve a distribuição da habilidade do indivíduo que responde com padrão **x**.

Como o MCP é do "tipo-Rasch", padrões de resposta com a mesma soma total de escores tem a mesma distribuição posterior de $\theta$. Isto significa que para L=12 itens com $m_i$=6 categorias de resposta para qualquer item, indivíduos que obtém padrões com a mesma soma, por exemplo:

INDIVÍDUO A→ x 1 3 1 4 5 5 1 2 1 soma 23
INDIVÍDUO B→ x 1 4 1 1 2 5 5 1 3 soma 23

obterão estimativas iguais de $\theta$ e seus respectivos desvios padrões também serão iguais. Isto não ocorre com o MRC de Samejima, mesmo quando consideramos que todos os itens tem o mesmo parâmetro de discriminação ($\alpha$).

Maximizando o logaritmo da função de verossimilhança dada em (4.30), ou seja, maximizando a distribuição á posteriori de $\theta$ vamos obter uma estimativa chamada modal Bayesiana ou "máximo posterior"(MAP).

$$\frac{\partial \log P(x;\theta)}{\partial \theta} = \sum_{i=1}^{L} \frac{\partial \log P_i(x_i)}{\partial \theta} + \frac{\partial \log \phi(\theta)}{\partial \theta} = 0$$

Com variância aproximada pela inversa da matriz $E\left[\frac{\partial^2 \log P(x;\theta)}{\partial \theta^2}\right]$, desta forma a primeira e segunda derivada de $P(\mathbf{x};\theta)$ em relação a $\theta$ são necessárias. Como consideramos que $\phi(\theta)$ é a função

densidade da distribuição normal padrão a estimativa modal é facilmente computada para cada valor de $\theta_q$.

Dado um valor inicial $\theta_q^0$ a estimativa MAP ($\hat{\theta}_q$) para os indivíduos na classe de habilidade $\theta_q$ é obtida resolvendo a seguinte equação iterativamente:

$$\left[\hat{\theta}_q\right]_{t+1} = \left[\hat{\theta}_q\right]_t - \left[E\frac{\partial^2 \log P(x;\theta)}{\partial\theta^2}\right]_t^{-1}\left[\frac{\partial \log P(x;\theta)}{\partial\theta}\right]_t \qquad (4.31)$$

Por outro lado, integrar (4.30) para obter sua média e variância não é difícil e tem a vantagem de não ser necessário derivar $\phi(\theta)$, pois a estimativa modal Bayesiana só será prática quando $\phi(\theta)$ for uma função densidade razoável. Esta média é chamada de estimativa "esperada posterior"(EAP), dada por Bock e Mislevy(1982).

Também podemos estimar a habilidade ($\theta$) dos indivíduos através da Máxima Verossimilhança (MV) como vimos em 4.3.1, considerando as estimativas obtidas pela Máxima Verossimilhança Marginal como valores verdadeiros dos parâmetros dos itens.

As estimativas MAP e EAP, existem para qualquer padrão de resposta **x**. Mesmo para indivíduos com padrão de resposta

x, onde x representa um padrão onde o indivíduo acertou ou errou todos os itens. Mas, as estimativas de habilidade encontradas por MV não são possíveis quando existem tais padrões de resposta x. Thissen (1982) desenvolve a estimação de máxima verossimilhança marginal para o modelo logístico de um parâmetro ou modelo de Rasch.

# CAPÍTULO 5

# FUNÇÃO DE INFORMAÇÃO

A Estimativa de Máxima Verossimilhança (EMV) $\hat{\theta}$ da habilidade do indivíduo, vista no Capítulo 4, converge a uma distribuição normal assintótica com média $\theta$ e variância dada por:

$$\sigma^2_{\hat{\theta}/\theta} = \frac{1}{-E\left[\frac{\partial^2}{\partial^2\theta}\ln P(\theta)\right]} = \left[I(\theta)\right]^{-1}, \qquad (5.1)$$

ou seja, o inverso da função de informação. Dessa forma, podemos encontrar o erro padrão da estimativa de cada habilidade, $\sigma_{\hat{\theta}/\theta}$, e conseqüentemente, um intervalo de confiança para cada parâmetro de habilidade dos indivíduos.

Para facilitar a extensão do caso de respostas dicotômicas para o caso de respostas com mais de duas categorias, que será visto posteriormente, é importante ressaltar que todas as categorias de resposta contribuem individualmente para a quantidade de informação do item e, conseqüentemente para a quantidade de informação do teste. Samejima (1969) assume que no caso

dicotômico, ambas as categorias de resposta proporcionam informação sobre a estimativa da habilidade do indivíduo[1].

(1) Para simplificar a notação excluiremos os índices j e i, nesse Capítulo.

## 5.1- FUNÇÃO DE INFORMAÇÃO DO ITEM - DICOTÔMICO

Para encontrar a função de informação do item precisamos da quantidade de informação dada por cada categoria k, e nesse caso k=1,2. Assim $I_k(\theta)$, a função de informação para a categoria k, é

$$I_k(\theta) = -\frac{\partial^2 \log P_k(\theta)}{\partial \theta^2} = -\frac{\partial}{\partial \theta}\left[\frac{P_k'(\theta)}{P_k(\theta)}\right] = \frac{\left[P_k'(\theta)\right]^2 - P_k(\theta)\,P_k''(\theta)}{\left[P_k(\theta)\right]^2}, \quad (5.2)$$

onde $P_k'(\theta)$ e $P_k''(\theta)$ são, respectivamente, a primeira e segunda derivadas da curva característica da categoria de resposta do item (CCCRI) em relação a $\theta$.

Considerando:

$P_1(\theta)$: a probalidade do indivíduo de habilidade $\theta$ escolher a resposta correta do item i, e $P_1(\theta)=P(\theta)$;

$P_2(\theta)$: a probalidade do indivíduo de habilidade $\theta$ escolher a resposta incorreta do item i, e $P_2(\theta)=Q(\theta) = 1-P(\theta)$.

A contribuição da categoria de resposta correta para a quantidade de informação do item é $I_1(\theta)P_1(\theta)$, e a contribuição da categoria de resposta incorreta é

$I_2(\theta)\ P_2(\theta)=I_2(\theta)Q(\theta)$, logo

$$I_k(\theta)\ P_k(\theta),$$

é a quantidade de informação proporcionada pela k-ésima categoria de resposta. Portanto,

$$I_i(\theta) = \sum_{k=1}^{2} I_k(\theta)P_k(\theta).$$

De (5.2) teremos,

$$= \sum_{k=1}^{2}\left\{\frac{\left[P_k'(\theta)\right]^2 - P_k(\theta)\,P_k''(\theta)}{\left[P_k(\theta)\right]^2}\right\}P_k(\theta),$$

sendo $P_1(\theta)=P(\theta)$ e $P_2(\theta)=Q(\theta)$

$$= \left\{\frac{\left[P'(\theta)\right]^2 - P(\theta)\,P''(\theta)}{\left[P(\theta)\right]^2}\right\}P(\theta) + \left\{\frac{\left[Q'(\theta)\right]^2 - Q(\theta)\,Q''(\theta)}{\left[Q(\theta)\right]^2}\right\}Q(\theta),$$

sendo[1] $P'(\theta)=-Q'(\theta)$ e $P''(\theta)=-Q''(\theta)$

---

(1) Pela derivada do quociente e de (2.7) onde: $P_i(\theta)=e^W/1+e^W$, sendo $W=D\alpha_i(\theta-\beta_i)$.

$P_i'(\theta)=\partial P_i(\theta)/\partial\theta=\left[e^W(1+e^W)-e^W(e^W)\right]/(1+e^W)^2=e^W/(1+e^W)^2$ e

$Q_i(\theta)=1-P_i(\theta)=1-e^W/1+e^W=1/1+e^W$, logo $Q_i'(\theta)=\partial Q_i(\theta)/\partial\theta=-e^W/(1+e^W)^2$.

Portanto $P_i'(\theta)=-Q_i'(\theta)$ e, da mesma forma verificamos que $P_i''(\theta)=-Q_i''(\theta)$.

$$= \frac{Q(\theta)\left[P'(\theta)\right]^2 - Q(\theta)\,P(\theta)\,P''(\theta) + \left[-P'(\theta)\right]^2 P(\theta) + P(\theta)\,Q(\theta)\,P''(\theta)}{P(\theta)\;Q(\theta)}$$

$$= \frac{\left[P'(\theta)\right]^2 \left[Q(\theta) + P(\theta)\right]}{P(\theta)\,Q(\theta)}$$

$$= \frac{\left[P'(\theta)\right]^2}{P(\theta)\,Q(\theta)}, \text{ é a função de informação do item.} \quad (5.3)$$

Este resultado é idêntico à expressão da função de informação do item definida por Birnbaun (1968) para escores de itens dicotômicos.

Samejima (1973), adverte que sob o modelo de três parâmetros de Birnbaum, a quantidade de informação da categoria de resposta do item pode ser negativa sob certas circunstâncias, mesmo que a quantidade total de informação do item seja sempre positiva.

## 5.1.1 - FUNÇÃO DE INFORMAÇÃO DO TESTE

A função de informação de um teste fornece uma medida do grau de precisão com que os L itens do teste estão estimando a habilidade, para qualquer ponto ao longo da escala de habilidade. Esta função tem a mesma interpretação na TRI que a fidedignidade na TCT.

O coeficiente de fidedignidade na TCT é simplesmente uma medida total da consistência da medição. A vantagem da função de informação do teste é que esta nos fornece uma medida da precisão da estimativa para cada nível de habilidade de interesse, em vez de uma medida global. A função de informação do teste é dada pela soma das funções de informação de todos os itens do teste e é denotada por I(θ).

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

De (5.3) teremos,

$$I(\theta) = \sum_{i=1}^{L} \frac{[P_i'(\theta)]^2}{P_i(\theta)\,Q_i(\theta)} \qquad (5.4)$$

Da expressão (5.4) podemos obter a função de informação do teste para vários modelos de CCI, como abaixo:

$$\text{NORMAL} \rightarrow I(\theta) = \sum_{i=1}^{L} \frac{a_i^2 [h_i(\theta)]^2}{P_i(\theta) Q_i(\theta)},$$

onde $h_i(\theta)$ é a ordenada da densidade normal padrão, correspondente a $P_i(\theta)$.

$$\text{LOGÍSTICO} \rightarrow I(\theta) = \sum_{i=1}^{L} a_i^2 P_i(\theta) Q_i(\theta),$$

Birnbaum (1968) e Lord (1980).

$$\text{RASCH} \rightarrow I(\theta) = \sum_{i=1}^{L} P_i(\theta) Q_i(\theta)$$

LOGÍSTICO de três parâmetros de Birnbaum (1968),

$$\rightarrow I(\theta) = \sum_{i=1}^{L} \frac{a_i^2 Q_i(\theta) \left(P_i(\theta) - \psi_i\right)^2}{P_i(\theta) \left(1 - \psi_i\right)^2}.$$

Veremos mais adiante que se tivermos mais do que duas categorias de resposta, ou seja, $k=1,2,\ldots,m_i$, a função de informação do teste é dada pela soma da contribuição de cada categoria de resposta dentro de cada item, para todo item, isto é,

$$I(\theta)=\sum_{i=1}^{L} I_i(\theta)=\sum_{i=1}^{L}\sum_{k=1}^{m_i} I_k(\theta)\,P_k(\theta),$$

onde L é o número de itens.

Se calcularmos a I(θ) para todos os níveis de habilidade θ obteremos uma curva de função de informação do teste e, é de grande importância calcularmos as funções de informação dos itens, pois estas, constituem um poderoso instrumento para a análise dos itens. Por exemplo, na Figura 5.1, o item 1 dá informação máxima para valores próximos de θ=-1.5, o item 2 dá informação máxima para valores próximos de θ=0 e o item 3 para valores próximos de θ=2.

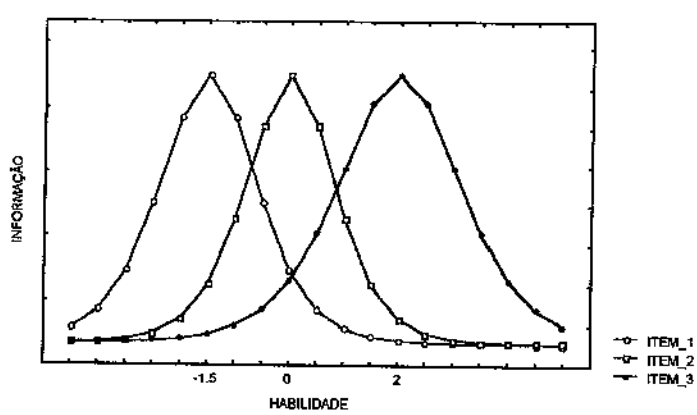



**FIGURA 5.1**- Funções de informação de três itens, sendo $\alpha_1=1, \alpha_2=1.1, \alpha_3=0.8, \beta_1=-1.5, \beta_2=0, \beta_3=2$ e $\psi=0$.

Se estamos interessados em medir θ para valores baixos, por exemplo, para $-2 \leq \theta \leq -1$ o item 1 dará mais informação. Desta forma, podemos usar a função de informação do item para obter

testes ajustados às nossas necessidades escolhendo os itens que mais interessarem aos nossos objetivos.

Quando construímos testes, é comum definirmos uma "função de informação alvo", que depende dos objetivos do teste, e a partir desta, selecionar os itens que dão mais informação em direção a nosso objetivo, Lord (1980).

A estrutura básica da função de informação dos modelos: logístico e normal, são as mesmas. Uma propriedade importante da função de informação do item que é de considerável interesse prático é a quantidade máxima de informação disponível. Quando utilizamos os modelos logístico e normal de dois parâmetros o máximo ocorre quando $\alpha_i(\theta-\beta_i) = 0$, ou seja, quando $\theta=\beta_i$.

O máximo sob estes modelos, segundo foram obtidos por Birnbaum (1968), são:

$$\text{NORMAL} \Rightarrow \max\left[I_i(\theta)\right]=\alpha_i^2\frac{\left[0.3989^2\right]}{(0.5)(0.5)}=0.64\,\alpha_i^2,$$

onde 0.3989 é a ordenada da densidade normal padrão quando Z = 0.

$$\text{LOGÍSTICO} \Rightarrow \max\,[I_i(\theta)] = \alpha_i^2\,(0.5)\,(0.5) = 0.25\,\alpha_i^2$$

$$\text{RASCH} \Rightarrow \text{máx}\,[I_i(\theta)] = (0.5)\,(0.5) = 0.25$$

Para que o modelo logístico esteja na mesma base que o modelo normal é preciso multiplicar a função de informação por $D^2$, onde D=1.7, logo máx $[I_i(\theta)] = 0.7225\,\alpha_i^2$ para o modelo logístico que, desta forma será ligeiramente maior do que o máximo no modelo normal.

O máximo da função de informação para o modelo logístico de três parâmetros de Birnbaum, não pode ser obtido quando $\theta=\beta_i$ porque sua média não ocorre em $\beta_i$. Birnbaum (1968) obteve a seguinte expressão para o nível de habilidade para o qual o máximo é obtido:

$$\theta_{max.} = \beta_i + \frac{1}{\alpha_i}\log\left[\frac{1+\sqrt{1+8\psi_i}}{2}\right],$$

assim, a quantidade máxima de informação para este modelo é uma função de $\theta_{max}$ e do parâmetro de "adivinhação" $\psi_i$ e, pode ser obtido substituindo $\theta_{max.}$ em $I_i(\theta)$.

Como resultado, a função de informação para o modelo logístico de dois parâmetros é um limite superior para a função de informação do modelo logístico de três parâmetros quando $\alpha_i$ e $\beta_i$ são os mesmos para os dois modelos.

## 5.2 - FUNÇÃO DE INFORMAÇÃO DO ITEM PARA RESPOSTAS CATEGORIZADAS

A quantidade de informação de um item aumenta quando são utilizadas respostas categóricas ao invés de respostas dicotômicas, visto que, cada categoria contribui para a informação do item, o aumento no número de categorias resultará em um aumento na soma de informação do item.

Vimos que a função de informação para itens dicotômicos pode ser estendida para itens com mais de duas categorias de resposta, considerando que $\sum_{k=1}^{m_i} P_k(\theta) = 1$.

A quantidade de informação, $I_K(\theta)$, devida a uma categoria de resposta do item, é uma medida de quão bem as respostas nesta categoria estimam a habilidade do indivíduo. A porção de informação de uma categoria de resposta, $I_k(\theta)P_k(\theta)$, é a quantidade de informação contribuída pela categoria para a informação do item.

Assim, considerando $k=1,2,\ldots,m_i$, a função de informação do item será dada por:

$$I_i(\theta) = \sum_{k=1}^{m_i} I_k(\theta) P_k(\theta), \qquad (5.5)$$

onde o produto $I_k(\theta)P_k(\theta)$ é a quantidade de informação da categoria k.

Sendo $P_k(\theta) = P_{ik}(\theta_j)$, a probabilidade definida em (3.3) para o Modelo de Resposta Categórica de Samejima (MRCS). A função de informação do item é definida por:

$$I_i(\theta) = \sum_{k=1}^{m_i} \frac{\left[P_k^{*\prime}(\theta) - P_{k+1}^{*\prime}(\theta)\right]^2}{\left[P_k^{*}(\theta) - P_{k+1}^{*}(\theta)\right]}. \qquad (5.6)$$

Vejamos,

a função de informação da k-ésima categoria de resposta de um item, é definida como sendo:

$$I_k(\theta) = -\frac{\partial^2 \log P_k(\theta)}{\partial \theta^2} = -\frac{\partial}{\partial \theta}\left[\frac{P_k'(\theta)}{P_k(\theta)}\right]$$

$$I_k(\theta) = \frac{\left[P_k'(\theta)\right]^2 - P_k(\theta)P_k''(\theta)}{\left[P_k(\theta)\right]^2}$$

A porção de informação da categoria k de resposta do item é

$$I_k(\theta)P_k(\theta) = \frac{\left[P_k'(\theta)\right]^2}{P_k(\theta)} - P_k''(\theta),$$

que substituído em (5.5) nos dá a função de informação do item:

$$I_i(\theta) = \sum_{k=1}^{mi}\left\{\frac{\left[P_k'(\theta)\right]^2}{P_k(\theta)} - P_k''(\theta)\right\} \qquad (5.7)$$

De (3.1) temos que

$$P_k(\theta) = P_k^*(\theta) - P_{k+1}^*(\theta),$$

$$P_k'(\theta) = P_k^{*'}(\theta) - P_{k+1}^{*'}(\theta) \text{ e}$$

$$P_k''(\theta) = P_k^{*''}(\theta) - P_{k+1}^{*''}(\theta), \text{ logo}$$

$$I_i(\theta) = \sum_{k=1}^{m_i}\frac{\left[P_k^{*'}(\theta) - P_{k+1}^{*'}(\theta)\right]^2}{\left[P_k^*(\theta) - P_{k+1}^*(\theta)\right]} - \sum_{k=1}^{m_i}\left[P_k^{*''}(\theta) - P_{k+1}^{*''}(\theta)\right] \qquad (5.8)$$

O somatório no segundo termo do lado direito de (5.8) desaparece.

Desenvolvendo o somatório em k, teremos:

$$\sum_{k=1}^{m_i}\left[P_k^{*''}(\theta) - P_{k+1}^{*''}(\theta)\right] = \left[P_1^{*''}(\theta) - P_2^{*''}(\theta)\right] + \left[P_2^{*''}(\theta) - P_3^{*''}(\theta)\right] + \left[P_3^{*''}(\theta) - P_4^{*''}(\theta)\right] + \ldots +$$

$$\left[P_{m_i-2}^{*''}(\theta) - P_{m_i-1}^{*''}(\theta)\right] + \left[P_{m_i-1}^{*''}(\theta) - P_{m_i}^{*''}(\theta)\right] + \left[P_{m_i}^{*''}(\theta) - P_{m_i+1}^{*''}(\theta)\right]$$

Os termos iguais de sinais opostos se cancelam,

$$\sum_{k=1}^{m_i}\left[P_k^{*\prime}(\theta)-P_{k+1}^{*\prime}(\theta)\right]=P_1^{*\prime}(\theta)-P_{m_i+1}^{*\prime}(\theta)$$

Sendo $\frac{\partial}{\partial\theta}\left(\frac{\partial}{\partial\theta}P_k^*(\theta)\right)=P_k^{*\prime}(\theta)$ ,

$$\sum_{k=1}^{m_i}\left[P_k^{*\prime}(\theta)-P_{k+1}^{*\prime}(\theta)\right]=\frac{\partial}{\partial\theta}\left(\frac{\partial}{\partial\theta}P_1^*(\theta)\right)-\frac{\partial}{\partial\theta}\left(\frac{\partial}{\partial\theta}P_{m_i+1}^*(\theta)\right),$$

do Capítulo 3, seção 3.1, $P_1^*(\theta)=1$ e $P_{m_i+1}^*(\theta)=0$.

$$=\frac{\partial}{\partial(\theta)}\left(\frac{\partial}{\partial(\theta)}1\right)-\frac{\partial}{\partial(\theta)}\left(\frac{\partial}{\partial(\theta)}0\right)=0$$

Portanto (5.8) pode ser escrita como a função de informação do item expressa em (5.6). □

**RESULTADO 5.1**: No caso de usarmos a distribuição normal, a função de informação do item para um dado nível de habilidade θ, será:

$$\mathrm{I}_i(\theta)=\sum_{k=1}^{m_i}\frac{\left[\alpha_i h_k^*(\theta)-\alpha_i h_{k+1}^*(\theta)\right]^2}{P_k^*(\theta)-P_{k+1}^*(\theta)}, \qquad (5.9)$$

onde $\frac{\partial P_k^*(\theta)}{\partial\theta}=\alpha_i h_k^*(\theta)$ e $\frac{\partial^2 P_k^*(\theta)=}{\partial\theta^2}\alpha_i^2 h_k^*(\theta)$.

**RESULTADO 5.2**: No caso de usarmos a distribuição logística a função de informação do item será (Baker,1992):

$$I_i(\theta) = \sum_{k=1}^{m_i} \frac{\alpha_i^2 \left(w_k - w_{k+1}\right)^2}{P_k^*(\theta) - P_{k+1}^*(\theta)}, \qquad (5.10)$$

sendo $\frac{\partial P_{K+1}^*(\theta)}{\partial \theta} = \alpha_i w_{k+1}$, $\frac{\partial P_k^*(\theta)}{\partial \theta} = \alpha_i w_k$, $\frac{\partial^2 P_k^*(\theta)}{\partial \theta^2} = \alpha_i^2 w_k \left[Q_k^*(\theta) - P_k^*(\theta)\right]$ e

$\frac{\partial^2 P_{k+1}^*(\theta)}{\partial \theta^2} = \alpha_i^2 w_{k+1} \left[Q_{k+1}^*(\theta) - P_{k+1}^*(\theta)\right]$,

onde $w_k = P_k^*(\theta) Q_k^*(\theta)$ *e* $w_{k+1} = P_{k+1}^*(\theta) Q_{k+1}^*(\theta)$.

# CAPÍTULO 6

# APLICAÇÃO: TEORIA DE RESPOSTA AO ITEM

O objetivo, nesse Capítulo, é analisar e interpretar os parâmetros dos itens (questões) da Prova de Inglês do Vestibular/97 da UNICAMP quando aplicamos os modelos estudados no Capítulo 3, em especial, o interesse está nos parâmetros dos itens estimados através do Modelo de Crédito Parcial de Masters (1982) - MCPM.

Desde que foi implantado em 1987, o vestibular da UNICAMP se diferencia dos demais vestibulares do País por ser um exame dissertativo que requer do candidato uma certa capacidade de organizar idéias e de estabelecer hipóteses e relações.

A UNICAMP vem, através dos anos, elevando seu prestígio a tal ponto, que hoje é uma das universidades públicas mais procuradas do País com cursos de graduação e pós-graduação conceituados dentro e fora do País. Aproximadamente 35.000 candidatos se inscreveram no vestibular de 1997 pleiteando vagas para os 42 cursos de graduação que ela oferece. O vestibular se estrutura em duas fases:

A **primeira fase,** realizada em 01/12/96, consistiu de uma prova de 12 questões de conhecimentos gerais (sobre as disciplinas obrigatórias do segundo grau: matemática, física, química, biologia, história e geografia) e de uma redação, ambas valendo 50% do escore final. O candidato que não "zerar" em nenhuma das provas passa para a segunda fase. A nota da primeira fase contribui com peso 2 (dois) para a nota final.

A **segunda fase,** realizada nos dias 12, 13, 14, 15/01/97, consistiu de quatro exames: Língua Portuguesa e Literatura de Língua Portuguesa e Ciências Biológicas; Química e História; Física e Geografia; Matemática e Língua Estrangeira (Inglês ou Francês), cada uma contendo doze questões dissertativas das quais quatro são de conceituação básica, quatro de complexidade média e quatro de alta complexidade. Nessa fase, as provas consideradas prioritárias para o candidato tem peso 2 (dois) no cômputo geral, (sendo a nota mínima não inferior a 30%) e as demais peso 1 (um), estabelecendo um maior equilíbrio entre o conhecimento específico e as habilidades gerais.

Para sua realização, o vestibular da UNICAMP envolve uma grande quantidade de profissionais competentes, sendo que, aproximadamente 200 deles são para a correção das provas,

dividindo-se em 9 bancas de corretores, cada uma correspondendo a uma disciplina.

Cada prova, na segunda fase, contém 12 questões (itens) e são corrigidas por no mínimo dois examinadores. Na existência de divergência entre esses dois, o presidente da banca é consultado.

## INFORMAÇÕES SOBRE A PROVA DE INGLÊS

Foram aprovados na primeira fase mais de 12.912 candidatos, dos quais 12.912 fizeram a prova de inglês na segunda fase, temos então, N=12.912 indivíduos.

Os itens da prova são examinados de tal forma que os indivíduos obtêm escores parciais, dependendo da organização e conteúdo de sua resposta, que variam de 0 a 5, ou seja, os indivíduos podem obter os escores 0, 1, 2, 3, 4 ou 5. A primeira categoria - obter escore zero - significa que a resposta está completamente inadequada; por outro lado - obter escore cinco - significa que resposta está adequada, e os escores intermediários significam que o indivíduo acertou parcialmente o item.

A Prova de Inglês do vestibular/97 pode ser encontrada no Apêndice B desse trabalho. Como exemplo, citamos, a seguir, o item 1 e sua grade de correção para compreendermos como são atribuídos os escores aos indivíduos:

ITEM1 - Por que Jane está nessa posição?

LAZY JANE

Lazy
lazy
lazy
lazy
lazy
lazy
Jane,
she
wants
a
drink
of
water
so
she
waits
and
waits
and
waits
and
waits
and
waits
for
it
to
rain.

(Silberstein, S. *Where the sidewalk ends.* Harper & Row Publishers, New York, 1974, p.87)

"Este item busca recuperar o funcionamento do texto.

**Escore 5** → Para obtenção do escore 5 é necessário incluir, na resposta, os seguintes elementos, assim como as relações corretas entre eles: a) Jane é preguiçosa (preguiça); b) quer beber água ou está com sede (sede); c) está esperando chover (chuva). Como exemplos de respostas com escore 5, citamos:

"Por que Jane está esperando que chova, pois quer beber água. Faz isso porque é devagar (lazy), preguiçosa";

"Porque ela é preguiçosa, então vai esperar chover para tomar água";

"Porque Jane é extremamente preguiçosa e espera que, até o copo de água que ela quer, caia do céu".

**Escore 3** → Para a obtenção desse escore era necessário que a resposta incluísse dois dos elementos acima descritos (preguiça, sede ou chuva). Como exemplo de escore 3 em que está faltando o elemento preguiça, podemos citar:

"Porque ela quer beber água e para isso está esperando chover".

**Escore 0** → As possíveis respostas para escore 0 (zero) podem ser:

"Ela está conversando com Lazy que está no céu";

"Porque ela está bêbada";

"Jane está nessa posição porque teve um desejo enorme de beber água até que esperou tanto que a água veio da chuva e ela acabou se afogando nela".

Como você pode notar, são respostas que não levaram em conta informações do texto ou que, apesar de incluir dois dos elementos (sede) e (chuva) estão comprometidas por adendos problemáticos" (Caderno de Questões: A UNICAMP Comenta suas Provas, 1998).

Como as categorias de respostas dos itens da prova são ordenadas, de modo que - para um indivíduo obter escore x ele deve já ter obtido o escore x-1 - os modelos que podem ser utilizados nesse caso são: o MCPM- Modelo de Crédito Parcial de Masters (1982), o MCPGM- Modelo de Crédito Parcial Generalizado de Muraki (1992), e também o MRCS- Modelo de Resposta Categórica de Samejima (1972), todos descritos no Capítulo 3.

O programa utilizado para estimar os parâmetros dos itens nesses modelos foi o programa MULTILOG versão 6.0 (Thissem, 1991). Este programa usa MVM/algoritmo EM (Bock e Aitkin,1981) para estimar os parâmetros dos itens, ver Capítulo 4.

O MULTILOG é um programa específico que facilita a análise de itens dentro da estrutura de TRI. Nas próximas seções apresentamos as estimativas dos parâmetros dos itens obtidas com o programa MULTILOG para os três modelos citados acima. Os programas para o uso do MULTILOG estão no Apêndice C.

## 6.1- ANÁLISE DESCRITIVA

No Quadro 6.1, apresentamos na primeira coluna a média e o desvio padrão (valor entre parênteses) dos escores obtidos em cada item e, no interior do Quadro as freqüências observadas (F.OBS.), as proporções observadas (P.OBS.) e as proporções esperadas (P.ESP.) de respostas dos indivíduos em cada categoria de resposta. Nesse Quadro podemos ver que a categoria com maior porcentagem de respostas foi a primeira categoria (obter escore zero) do item 8, com 64% das respostas dos indivíduos. Depois seguem as primeiras categorias (obter escore zero) dos itens 9 e 2 com 58% e 56%, respectivamente.

Observando esse fato poderíamos considerar os itens 2, 8, e 9 como os mais difíceis dessa prova por apresentarem uma grande porcentagem de indivíduos com escore zero e, consequentemente, as menores médias: $m_2$=1.133; $m_9$=1.189; $m_8$=1.982. Apesar disso, o item 8 tem o maior desvio padrão (2.333) dos escores e a terceira maior porcentagem de indivíduos com escore 5 (26%), ficando abaixo somente dos itens 1 e 11, cujas porcentagens de escore 5 foram 31% e 32%, respectivamente. Os itens 1 e 11, quando analisados pela TCT, foram apontados como os itens mais

fáceis da prova com as maiores médias: $m_1$=3.496 e $m_{11}$=3.847, respectivamente.

**QUADRO 6.1**- Dados gerais sobre os itens da Prova de Inglês.

| Categorias | (k) | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| Escores | (x) | 0 | 1 | 2 | 3 | 4 | 5 |
| **ITEM 1** | F.OBS. | 995 | 1239 | 1174 | 5192 | 162 | 4023 |
| m=3.496 | P.OBS. | 0.08 | 0.10 | 0.09 | 0.41 | 0.01 | **0.31** |
| (1.445) | P.ESP. | 0.08 | 0.10 | 0.09 | 0.41 | 0.01 | 0.31 |
| **ITEM 2** | F.OBS. | 6808 | 1973 | 755 | 1768 | 690 | 139 |
| m=1.133 | P.OBS. | **0.56** | 0.16 | 0.06 | 0.15 | 0.06 | 0.01 |
| (1.450) | P.ESP. | 0.57 | 0.16 | 0.06 | 0.14 | 0.06 | 0.01 |
| **ITEM 3** | F.OBS. | 2861 | 2106 | 3486 | 2488 | 624 | 841 |
| m=2.117 | P.OBS. | 0.23 | 0.17 | 0.28 | 0.20 | 0.05 | 0.07 |
| (1.402) | P.ESP. | 0.24 | 0.17 | 0.28 | 0.19 | 0.05 | 0.07 |
| **ITEM 4** | F.OBS. | 1613 | 335 | 1625 | 4062 | 3294 | 808 |
| m=2.987 | P.OBS. | 0.14 | 0.03 | 0.14 | 0.35 | 0.28 | 0.07 |
| (1.322) | P.ESP. | 0.15 | 0.03 | 0.14 | 0.34 | 0.27 | 0.07 |
| **ITEM 5** | F.OBS. | 2278 | 2256 | 2071 | 2064 | 2118 | 1825 |
| m=2.868 | P.OBS. | 0.18 | 0.18 | 0.16 | 0.16 | 0.17 | 0.14 |
| (1.618) | P.ESP. | 0.18 | 0.18 | 0.17 | 0.16 | 0.16 | 0.14 |
| **ITEM 6** | F.OBS. | 2842 | 1532 | 1434 | 2344 | 3182 | 887 |
| m=2.698 | P.OBS. | 0.23 | 0.13 | 0.12 | 0.19 | 0.26 | 0.07 |
| (1.598) | P.ESP. | 0.24 | 0.13 | 0.12 | 0.19 | 0.25 | 0.07 |
| **ITEM 7** | F.OBS. | 4299 | 1937 | 1110 | 1933 | 788 | 1546 |
| m=2.068 | P.OBS. | 0.37 | 0.17 | 0.10 | 0.17 | 0.07 | 0.13 |
| (1.816) | P.ESP. | 0.39 | 0.17 | 0.09 | 0.16 | 0.06 | 0.13 |
| **ITEM 8** | F.OBS. | 7891 | 213 | 493 | 122 | 438 | 3229 |
| m=1.982 | P.OBS. | **0.64** | 0.02 | 0.04 | 0.01 | 0.04 | **0.26** |
| (2.333) | P.ESP. | 0.65 | 0.02 | 0.04 | 0.01 | 0.03 | 0.25 |
| **ITEM 9** | F.OBS. | 6723 | 1386 | 2208 | 734 | 298 | 316 |
| m=1.189 | P.OBS. | **0.58** | 0.12 | 0.19 | 0.06 | 0.03 | 0.03 |
| (1.529) | P.ESP. | 0.59 | 0.12 | 0.18 | 0.06 | 0.02 | 0.03 |
| **ITEM 10** | F.OBS. | 1881 | 3378 | 4289 | 966 | 1084 | 528 |
| m=1.983 | P.OBS. | 0.16 | 0.28 | 0.35 | 0.08 | 0.09 | 0.04 |
| (1.275) | P.ESP. | 0.16 | 0.28 | 0.35 | 0.08 | 0.09 | 0.04 |
| **ITEM 11** | F.OBS. | 1031 | 505 | 442 | 2509 | 3548 | 3749 |
| m=3.847 | P.OBS. | 0.09 | 0.04 | 0.04 | 0.21 | 0.30 | **0.32** |
| (1.324) | P.ESP. | 0.09 | 0.05 | 0.04 | 0.22 | 0.30 | 0.31 |
| **ITEM 12** | F.OBS. | 2831 | 1851 | 1426 | 1468 | 1753 | 1695 |
| m=2.460 | P.OBS. | 0.26 | 0.17 | 0.13 | 0.13 | 0.16 | 0.15 |
| (1.822) | P.ESP. | 0.29 | 0.18 | 0.13 | 0.13 | 0.14 | 0.14 |

Antes de estimar os parâmetros dos modelos CPM, CPGM e RCS, é necessário verificar as suposições de unidimensionalidade e independência condicional. A unidimensionalidade é verificada utilizando a técnica de Análise Fatorial Multivariada apresentada no Apêndice A desse trabalho. Como a unidimensionalidade é satisfeita, então a suposição de independência condicional também é satisfeita (Hambleton,1979), isso significa que os 12 itens da prova de inglês estão medindo a mesma habilidade nos indivíduos, ou seja leitura (inglês).

## 6.2- ESTIMAÇÃO DOS PARÂMETROS DOS ITENS SOB O MODELO DE CRÉDITO PARCIAL DE MASTERS

Como vimos na teoria do Capítulo 3, nos modelos de Masters e Muraki, a probabilidade de um indivíduo com habilidade $\theta_j$ obter um certo escore no item i é dada por $P_{ix}(\theta_j)$, onde x pode ser interpretado como a quantidade de passos completados com sucesso.

Por exemplo, para seis categorias ($m_i$=6), a probabilidade de alcançar a categoria 4 (escore 3) é dada pela probabilidade de completar o terceiro passo, $P_{i3}(\theta_j)$.

$$
\begin{array}{lccccccc}
k & \Rightarrow & 1 & 2 & 3 & (4) & 5 & 6 \\
x & \Rightarrow & 0 & 1 & 2 & (3) & 4 & 5 \\
 & & \rightarrow & \rightarrow & \rightarrow & & \rightarrow & \rightarrow \\
passos & \Rightarrow & 1 & 2 & (3) & & 4 & 5
\end{array}
$$

A dificuldade dos passos é dada por $\beta_{ix}$ - chamado parâmetro de dificuldade individual do passo x do i-ésimo item. Por exemplo no Quadro 6.2, onde são apresentadas as estimativas dos parâmetros dos passos individuais em cada item ($\beta_{ix}$ para x=0,1,2,3,4,5 e i= 1,2,...,12), o valor $\beta_{81}$= 3.49 é a dificuldade individual do passo 1, ou seja, de passar de escore zero para escore um no item 8. Para interpretar esse valor de $\beta_{81}$, que é maior que as dificuldades

individuais dos outros passos do item 8, podemos dizer que é difícil obter 1 ponto no item 8, porém, tendo conseguido obter 1 ponto nesse item, o indivíduo conseguiu também os outros passos e, consequentemente, obter mais pontos nesse item, porque os passos posteriores eram mais fáceis. Isso também pode ser facilmente verificado no Quadro 6.1 pois 64% dos indivíduos obtiveram escore zero e 26% obtiveram escore 5, restando apenas 10% dos indivíduos nos escores intermediários. Citamos também, como exemplo, o item 8 e sua grade de correção: O item 8 diz respeito ao trecho abaixo:

STRICTLY SPEAKING, this story should not be written or told at all. To write it or to tell it is to spoil it. This is because the man who had the strange experience we are going to talk about never mentioned it to anybody, and the fact that he kept his secret and sealed it up completely in his memory is the whole point of the story. Thus we must admit that handicap at the beginning – that it is absurd for us to tell the story, absurd for anybody to listen to it and unthinkable that anybody should believe it.

We will, however, do this man one favour. We will refrain from mentioning him by his complete name. This will enable us to tell his secret and permit him to continue looking his friends in the eye. But we can say that his surname is Duffy. There are thousands of these Duffys in the world; even at this moment there is probably a new Duffy making his appearance in some corner of it. We can even go so far as to say that he is John Duffy's brother. We do not break faith in saying so, because if there are only one hundred John Duffys in existence, and even if each one of them could be met and questioned, no embarrassing enlightenments would be forthcoming. That is because the John Duffy in question never left his house, never left his bed, never talked to anybody in his life and was never seen by more than one man. That man's name was Gumley. Gumley was a doctor. He was present when John Duffy was born and also when he died, one hour later.

(O'Brien, F. *Stories and Plays*. Penguin Books, 1974, p.91.)

ITEM8 – Quem é o personagem principal da estória que o texto vai contar?

"Para responder adequadamente o item 8, o candidato deveria ser capaz de seguir uma cadeia referencial ou uma seqüência de pronomes para identificar o seu referente. Esse item ilustra claramente a importância da função e não apenas do significado da palavra no texto. Neste caso, não é suficiente saber que "he" significa "ele" mas é necessário identificar a quem esse pronome se refere no texto. O objetivo

colocado pela pergunta restringe as possibilidades de resposta, uma vez que somente três pessoas são mencionadas e, portanto, poderiam ser, em princípio, o protagonista da história, "fechando" a grade naturalmente e facilitando a correção.

**Escore 5** → Para obter escore 5, a resposta deveria conter apenas um elemento, isto é, a identificação do personagem. As seguintes respostas foram consideradas adequadas:

"O irmão de John Duffy";

"Alguem de sobrenome Duffy, irmão de John Duffy";

"É um homem de sobrenome Duffy, irmão de John Duffy";

"O personagem do texto é um homem de sobrenome Duffy que tinha um irmão, John Duffy, que morreu uma hora após ao nascer".

**Escore 0** → As seguintes respostas obtiveram escore 0 (zero), por serem vagas, sem maiores explicações, com adendos comprometedores ou identificação errada do personagem:

"Um homem";

"Um homem que teve estranhas experiências";

"O personagem principal da estória que o texto vai contar é o irmão de John Duffy: o médico Gumley";

"Gumley";

"John Duffy" " (Caderno de Questões: A UNICAMP Comenta suas Provas, 1998).

No Quadro 6.2 estão as estimativas dos parâmetros dos itens e seus respectivos desvios padrões, entre parênteses. Nesse Quadro verificamos que os passos mais difíceis da prova toda foram os necessários para obter escore 4 no item 1 e para obter escore 1 no item 8. Na seção 6.2.1. utilizaremos esses valores para construir as Curvas Características das Categorias de Respostas dos Itens (CCCRI).

**QUADRO 6.2**- Estimativas dos Parâmetros dos Itens - Masters.

| ITEM | i=1 | i=2 | i=3 | i=4 | i=5 | i=6 | i=7 | i=8 | i=9 | i=10 | i=11 | i=12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\beta_{i1}$ | -0.74 (0.06) | 1.18 (0.04) | 0.05 (0.04) | 1.18 (0.08) | -0.33 (0.04) | 0.30 (0.04) | 0.62 (0.04) | **3.49** (0.09) | 1.54 (0.04) | -0.84 (0.04) | 0.14 (0.07) | 0.20 (0.04) |
| $\beta_{i2}$ | -0.25 (0.06) | 1.10 (0.06) | -0.54 (0.04) | -1.77 (0.08) | -0.05 (0.04) | -0.05 (0.05) | 0.57 (0.05) | -0.79 (0.11) | -0.30 (0.05) | -0.26 (0.03) | -0.24 (0.09) | 0.23 (0.05) |
| $\beta_{i3}$ | -1.59 (0.04) | -0.54 (0.06) | 0.49 (0.04) | -0.91 (0.04) | 0.05 (0.04) | -0.42 (0.05) | -0.38 (0.05) | 1.59 (0.14) | 1.44 (0.06) | 1.67 (0.05) | -1.91 (0.07) | 0.10 (0.05) |
| $\beta_{i4}$ | **3.55** (0.11) | 1.41 (0.06) | 1.71 (0.06) | 0.39 (0.03) | 0.19 (0.04) | -0.07 (0.04) | 1.22 (0.06) | -0.96 (0.14) | 1.40 (0.09) | 0.24 (0.06) | -0.32 (0.04) | 0.10 (0.05) |
| $\beta_{i5}$ | -2.95 (0.11) | 2.21 (0.13) | 0.18 (0.07) | 1.76 (0.05) | 0.52 (0.04) | 1.67 (0.05) | -0.21 (0.06) | -1.57 (0.07) | 0.58 (0.11) | 1.23 (0.07) | 0.16 (0.03) | 0.45 (0.05) |

Os valores em negrito no Quadro 6.3 correspondem ao ponto na escala da habilidade $\theta$ onde o item acusa maior informação. Essa é uma das vantagens de se usar a TRI, pois ela fornece informação para cada valor da habilidade $\theta$. Por exemplo: de acordo com a teoria de Função de Informação do Capítulo 5 - como desejamos obter mais informação para valores altos da habilidade, pois deseja-se selecionar os melhores indivíduos (candidatos) - os itens que fornecem mais informação nesse sentido são os itens de números 1, 4 e 11. Esses itens dão mais informação para indivíduos com habilidade entre 0.5 e 1.5, aproximadamente (ver Figura 6.1).

**QUADRO 6.3**- Funções de Informação dos Itens e Total - Masters.

| $(\theta)$ | **-2.0** | **-1.5** | **-1.0** | **-0.5** | **0.0** | **0.5** | **1.0** | **1.5** | **2.0** |
|---|---|---|---|---|---|---|---|---|---|
| $I_1(\theta)$ | 0.20 | 0.28 | 0.37 | 0.44 | 0.50 | **0.53** | **0.53** | 0.50 | 0.42 |
| $I_2(\theta)$ | 0.51 | 0.62 | **0.63** | 0.52 | 0.36 | 0.23 | 0.14 | 0.08 | 0.05 |
| $I_3(\theta)$ | 0.45 | **0.49** | **0.49** | 0.47 | 0.43 | 0.38 | 0.32 | 0.25 | 0.19 |
| $I_4(\theta)$ | 0.18 | 0.21 | 0.25 | 0.31 | 0.41 | 0.53 | **0.60** | 0.55 | 0.42 |
| $I_5(\theta)$ | 0.27 | 0.38 | 0.51 | 0.62 | **0.65** | 0.57 | 0.44 | 0.32 | 0.22 |
| $I_6(\theta)$ | 0.22 | 0.30 | 0.42 | 0.56 | **0.66** | 0.64 | 0.50 | 0.34 | 0.21 |
| $I_7(\theta)$ | 0.40 | 0.60 | **0.78** | 0.81 | 0.66 | 0.45 | 0.28 | 0.17 | 0.10 |
| $I_8(\theta)$ | 0.24 | 0.63 | 1.27 | **1.45** | 0.87 | 0.35 | 0.13 | 0.05 | 0.02 |
| $I_9(\theta)$ | **0.70** | 0.68 | 0.56 | 0.42 | 0.29 | 0.20 | 0.13 | 0.08 | 0.06 |
| $I_{10}(\theta)$ | 0.48 | **0.50** | 0.47 | 0.40 | 0.33 | 0.28 | 0.24 | 0.20 | 0.17 |
| $I_{11}(\theta)$ | 0.12 | 0.15 | 0.20 | 0.27 | 0.39 | 0.57 | 0.74 | **0.78** | 0.62 |
| $I_{12}(\theta)$ | 0.29 | 0.44 | 0.63 | **0.76** | 0.73 | 0.57 | 0.38 | 0.24 | 0.15 |
| $I(\theta)$ | 5.1 | 6.3 | 7.6 | **8.0** | 7.3 | 6.3 | 5.4 | 4.6 | 3.7 |
| $S(\theta)$ | 0.44 | 0.40 | 0.36 | 0.35 | 0.37 | 0.40 | 0.43 | 0.47 | 0.52 |

A função de informação total fornece mais informação para valores da habilidade entre -1 e zero. Esse fato poderia nos levar a acreditar que a maior parte dos indivíduos tem habilidade em leitura (inglês) classificada entre -1 e zero, o que não é verdade segundo alguns autores. Carlson (1996) faz um estudo da Função de Informação considerando itens dicotômicos e itens politômicos e diz que itens dissertativos têm uma expectativa um tanto alta dos indivíduos pois, pelo seu estudo de tais itens, eles parecem fornecer medidas precisas além do intervalo de habilidade onde a maioria dos indivíduos está classificado.

Em geral, os itens que fornecem mais informação entre -1 e zero são os itens 2, 3, 5, 6, 7, 8, e 12. O item 9 dá mais informação para valores da habilidade em torno de -2, ou seja, para valores baixos da habilidade.

Nas próximas seções, onde aplicamos modelos que estimam o parâmetro de discriminação dos itens $(\alpha_i)$, vamos verificar que os itens 5 e 12 são os que discriminam mais entre os indivíduos e os itens 2 e 9 são os que discriminam menos entre os indivíduos. O item 8 fornece mais informação para valores de habilidade em torno de -0.5, este fato pode ser visto na Figura 6.1.

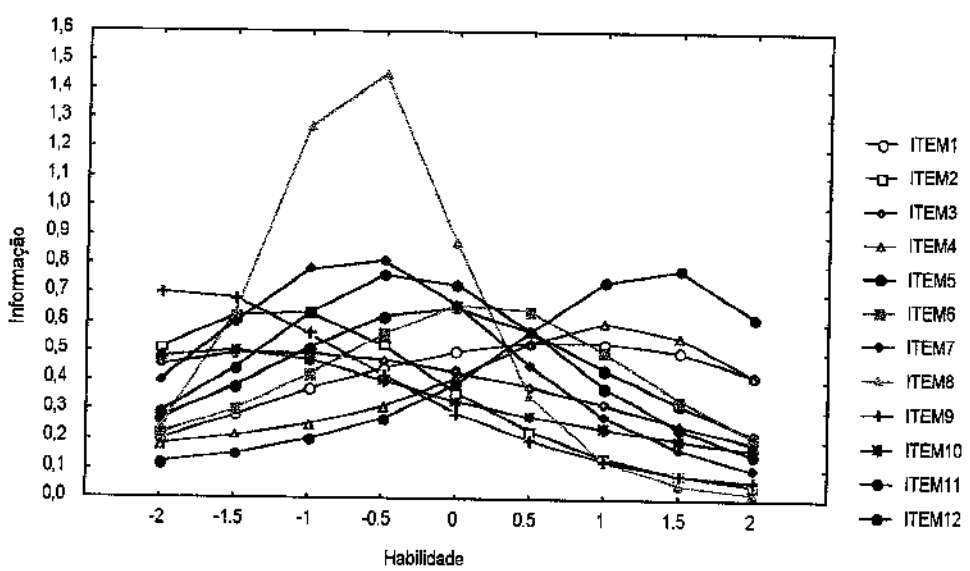



**FIGURA 6.1**- Funções de Informação dos 12 Itens da Prova de Inglês - Quadro 6.3.

## 6.2.1- INTERPRETAÇÃO GRÁFICA DOS ITENS - MCPM

A Teoria de Resposta ao Item (TRI) é uma teoria voltada para a análise individual dos itens de uma prova (ou teste). Dessa forma, aqui analisaremos individualmente os 12 itens da Prova de Inglês com o objetivo de interpretar o comportamento dos itens através de seus parâmetros $(\beta_{ix})$ e fornecer ao COMVEST subsídios para o aprimoramento das grades de correção utilizadas pelos examinadores na atribuição dos escores de 0, 1, 2, 3, 4 e 5 às respostas dos indivíduos aos itens.

Para analisar o comportamento dos itens, construímos os gráficos das Curvas Características das Categorias de Respostas dos Itens (CCCRI), para os 12 itens, usando o programa STATISTICA, versão 5.0. As estimativas dos parâmetros dos itens estão disponíveis no Quadro 6.2.

Os valores dos parâmetros $(\beta_{ix})$ correspondem aos pontos onde ocorrem as interseções entre as Curvas Características das Categorias de Respostas do Item (CCCRI) adjacentes e, podem ser localizados no eixo da abscissa nas Figuras 6.2 a 6.13. Esse fato também é uma característica importante da TRI: os parâmetros dos

itens e dos indivíduos podem ser localizados no mesmo eixo, na mesma escala. Interpretaremos os parâmetros de dificuldade dos passos $\beta_{ix}$ (i=1,2,...,12 e x=0,1,2,3,4,5) de alguns itens dados no Quadro 6.2.

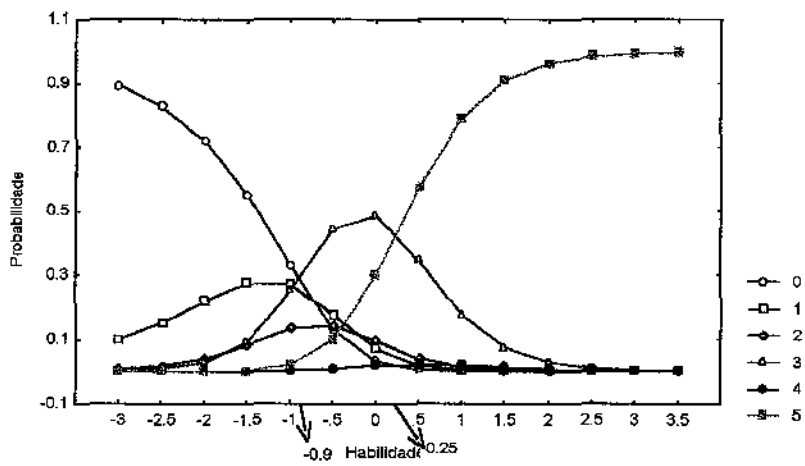



**FIGURA 6.2-** Curvas Características das Categorias de Respostas do Item 1.

O mais provável nesse item é que os indivíduos obtenham escores 3 ou 5, vejamos:

Completar somente o primeiro ou os dois primeiros passos $(\beta_{11} = -0.74 e \beta_{12} = -0.25)$ e obter escores 1 ou 2 são eventos pouco prováveis nesse item (curvas de cor vermelha e verde) para indivíduos com habilidade acima de -0.9, aproximadamente. Isso significa que muitos indivíduos obtiveram escore 3, completando o terceiro passo $(\beta_{13} = -1.59)$. O quarto passo $(\beta_{14} = 3.55)$, obter escore 4, foi o mais difícil nesse item. Se observarmos a Figura 6.2 veremos que a curva de cor rosa se sobressai indicando que indivíduos com habilidade entre −0.9 e

0.25, aproximadamente, tem boa probabilidade de obter escore 3 nesse item. Observamos também que obter escore 4 nesse item tem probabilidade quase zero (curva de cor preta), porque o quinto passo $(\beta_{15} = -2.95)$ foi muito fácil para quem alcançou escore 4.

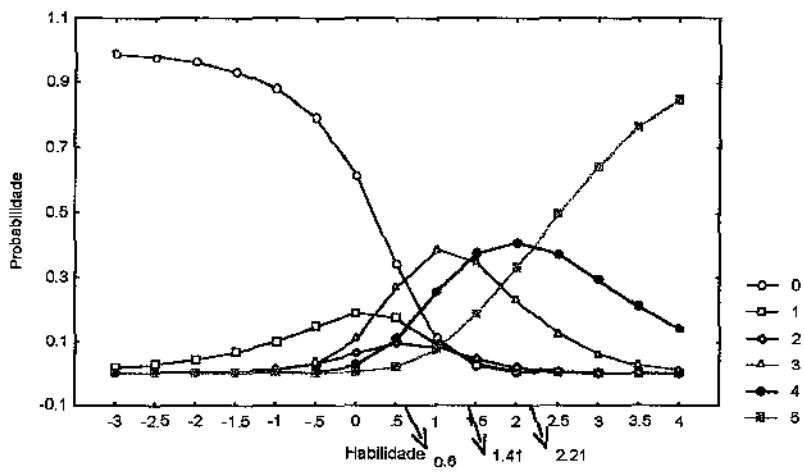



**FIGURA 6.3-** Curvas Características das Categorias de Respostas do Item 2.

No item 2, os dois primeiros passos $(\beta_{21} = 1.18 e \beta_{22} = 1.10)$ são razoavelmente difíceis e tem aproximadamente a mesma dificuldade, por isso o evento mais provável para indivíduos com habilidade menor que 0.6 é obter escore zero (curva de cor azul). Os indivíduos que conseguiram efetuar os dois primeiros passos, conseguiram também o terceiro passo $(\beta_{23} = -0.54)$ e obter escore 3, esse foi o passo mais fácil nesse item. Os indivíduos com habilidade maior que 1.41 $(\beta_{24} = 1.41)$ tiveram boa probabilidade de obter escore 4 (curva de cor preta) e, apenas os indivíduos com habilidade maior que 2.21

$(\beta_{25} = 2.21)$ tiveram boa probabilidade de obter escore 5 (curva de cor cinza).

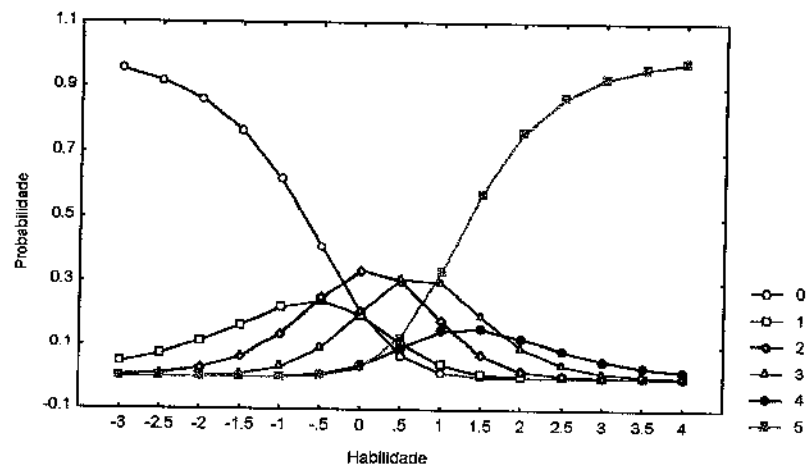



**FIGURA 6.4-** Curvas Características das Categorias de Respostas do Item 3.

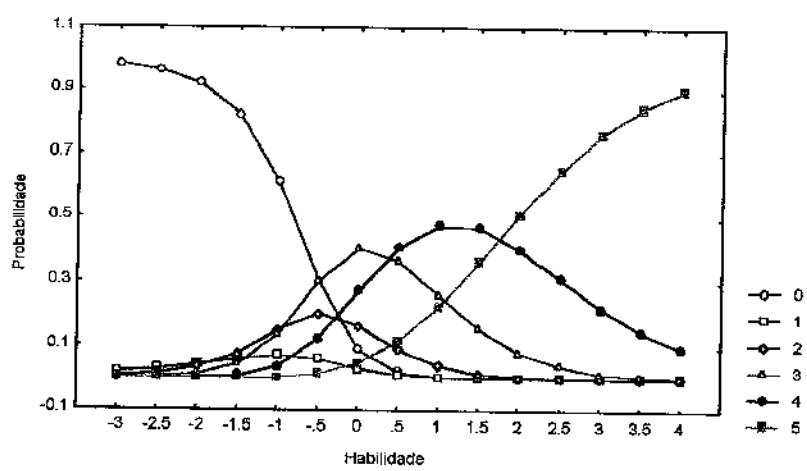



**FIGURA 6.5-** Curvas Características das Categorias de Respostas do Item 4.

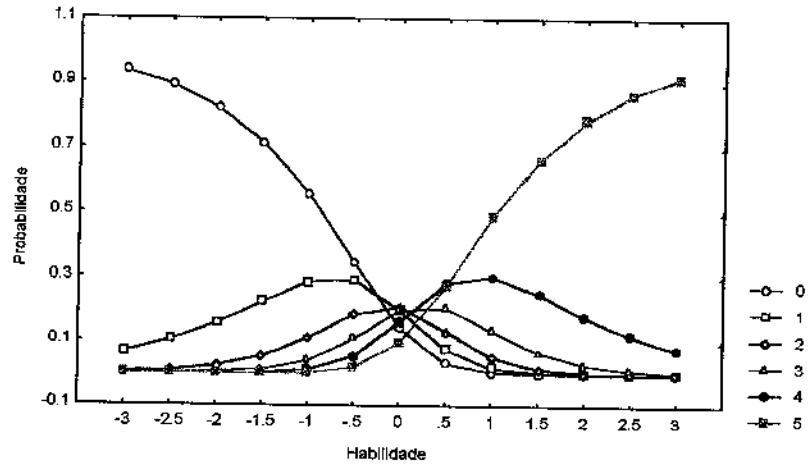



**FIGURA 6.6-** Curvas Características das Categorias de Respostas do Item 5.

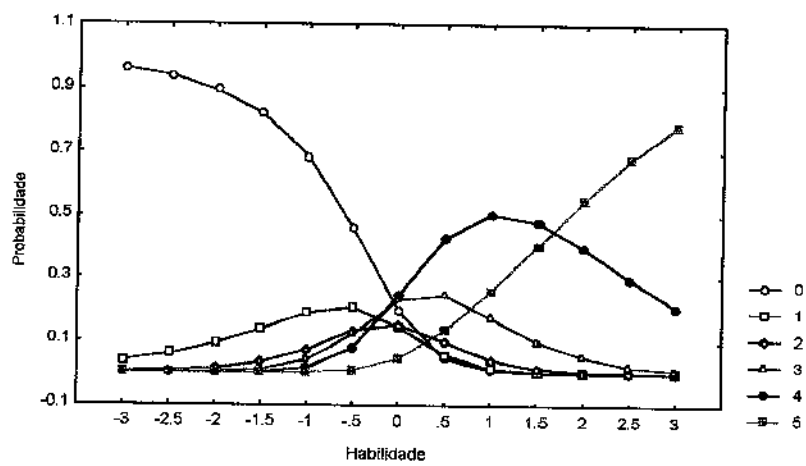



**FIGURA 6.7-** Curvas Características das Categorias de Respostas do Item 6.

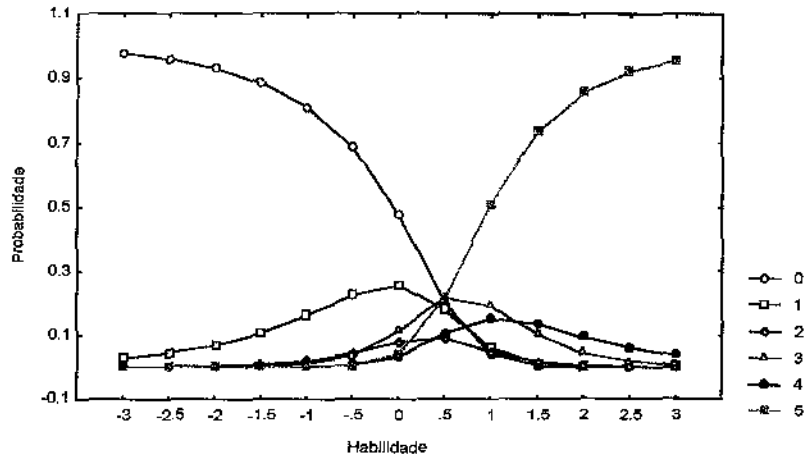



**FIGURA 6.8-** Curvas Características das Categorias de Respostas do Item 7.

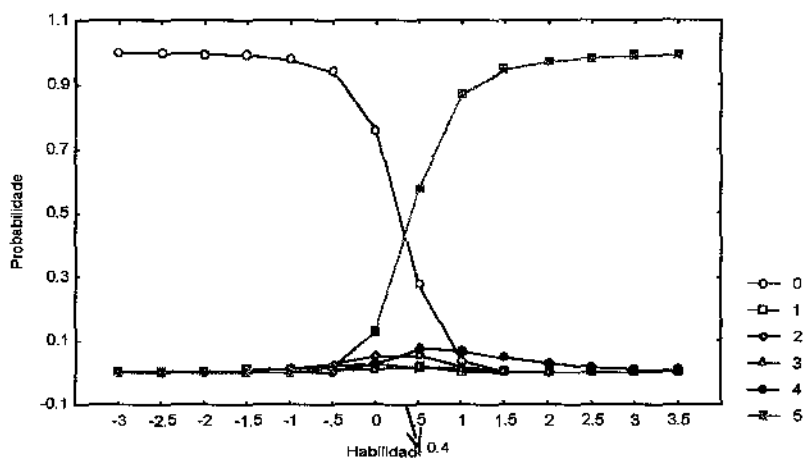



**FIGURA 6.9-** Curvas Características das Categorias de Respostas do Item 8.

Nesse item, obter escores 1, 2, 3 e 4 (curvas de cores: vermelha, verde, rosa e preta) são eventos pouco prováveis para quem conseguiu alcançar o escore 1 (primeiro passo, $\beta_{81} = 3.49$). Veja na Figura 6.9 que indivíduos com habilidade maior que 0.4 tiveram boa probabilidade de obter escore 5 (curva de cor cinza). Isso se deve ao fato de que o passo mais difícil nesse item foi o primeiro

$(\beta_{81} = 3.49)$ e, os demais foram muito fáceis em relação ao primeiro $(\beta_{82} = -0.79, \beta_{83} = 1.59, \beta_{84} = -0.96 e \beta_{85} = -1.57)$.

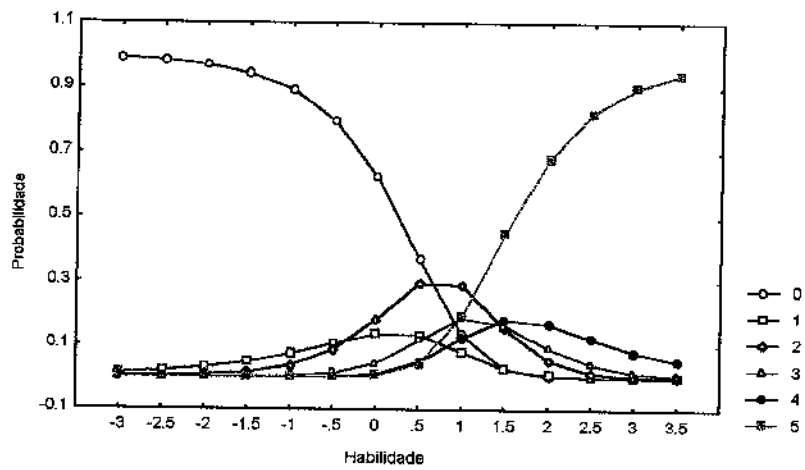



**FIGURA 6.10-** Curvas Características das Categorias de Respostas do Item 9.

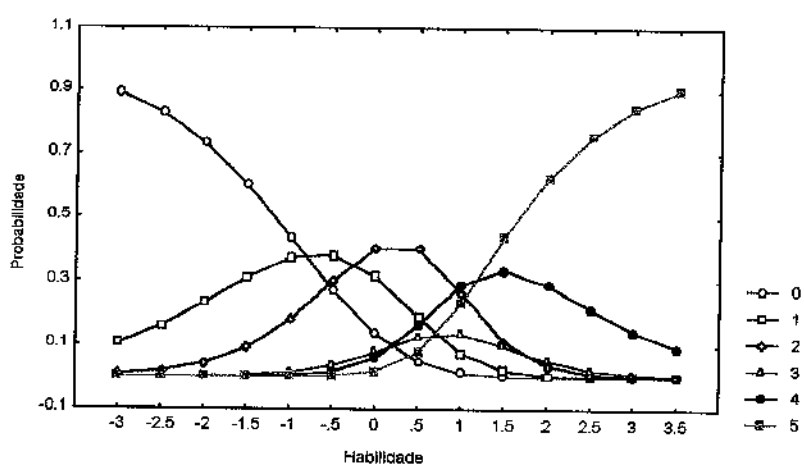



**FIGURA 6.11-** Curvas Características das Categorias de Respostas do Item 10.

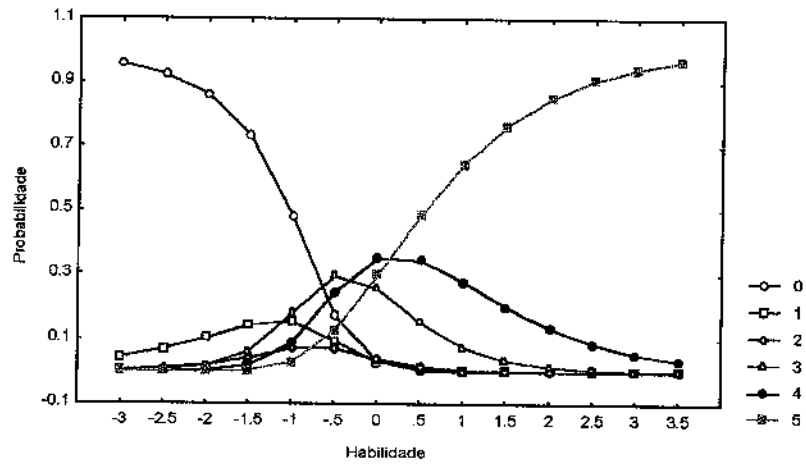



**FIGURA 6.12**- Curvas Características das Categorias de Respostas do Item 11.

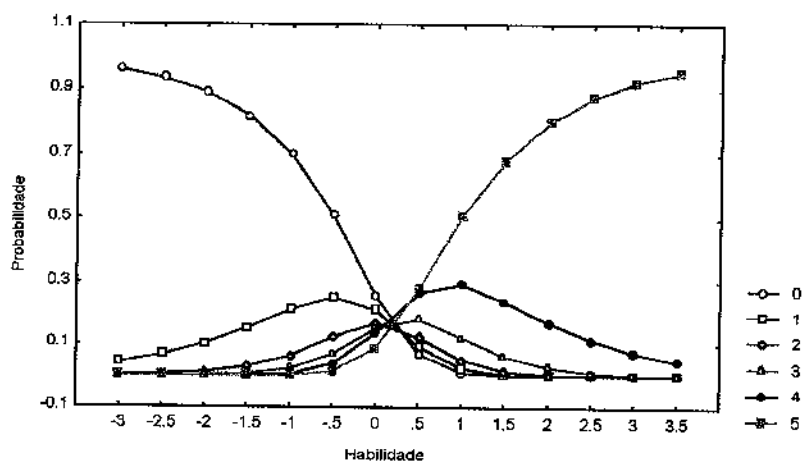



**FIGURA 6.13**- Curvas Características das Categorias de Respostas do Item 12.

## 6.3- ESTIMAÇÃO DOS PARÂMETROS DOS ITENS SOB O MODELO DE CRÉDITO PARCIAL GENERALIZADO DE MURAKI

Quando consideramos, além dos parâmetros de dificuldade individual dos passos nos itens, o parâmetro de discriminação $(\alpha_i)$ no modelo de CPM, estaremos utilizando o modelo de CPGM. Dessa forma, no Quadro 6.4 estão as estimativas dos 72 parâmetros dos itens $\rightarrow$ 5x12=60 parâmetros de dificuldade individual dos passos e 12 parâmetros de discriminação.

As estimativas dos parâmetros obtidas com o modelo CPGM apresentadas no Quadro 6.4 abaixo nos levam às mesmas interpretações feitas anteriormente para as estimativas obtidas com o modelo CPM, seções 6.2 e 6.2.1.

**QUADRO 6.4**-Estimativas dos Parâmetros dos Itens - Muraki.

| ITEM | i=1 | i=2 | I=3 | i=4 | i=5 | i=6 | i=7 | i=8 | i=9 | i=10 | i=11 | i=12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha_i$ | **0.71** (0.01) | **0.23** (0.01) | **0.46** (0.01) | **0.48** (0.01) | **1.10** (0.02) | **0.60** (0.01) | **0.41** (0.01) | **0.27** (0.01) | **0.33** (0.01) | **0.74** (0.02) | **0.65** (0.02) | **0.75** (0.02) |
| $\beta_{i1}$ | -1.02 (0.06) | 1.23 (0.04) | 0.10 (0.04) | 1.24 (0.09) | -0.93 (0.05) | 0.23 (0.05) | 0.67 (0.04) | 3.56 (0.09) | 1.56 (0.04) | -1.03 (0.04) | -0.05 (0.08) | 0.04 (0.04) |
| $\beta_{i2}$ | -0.43 (0.06) | 1.00 (0.06) | -0.54 (0.04) | -1.75 (0.08) | -0.30 (0.04) | -0.09 (0.05) | 0.56 (0.05) | -0.83 (0.11) | -0.38 (0.05) | -0.29 (0.03) | -0.38 (0.09) | 0.19 (0.05) |
| $\beta_{i3}$ | -1.67 (0.04) | -0.77 (0.06) | 0.46 (0.04) | -0.92 (0.04) | 0.11 (0.04) | -0.42 (0.05) | -0.43 (0.05) | 1.46 (0.14) | 1.27 (0.06) | 1.79 (0.05) | -1.99 (0.07) | 0.17 (0.05) |
| $\beta_{i4}$ | 3.59 (0.11) | 1.07 (0.06) | 1.66 (0.06) | 0.37 (0.03) | 0.54 (0.05) | -0.02 (0.04) | 1.13 (0.06) | -1.16 (0.14) | 1.16 (0.10) | 0.50 (0.06) | -0.33 (0.04) | 0.27 (0.05) |
| $\beta_{i5}$ | -2.80 (0.11) | 1.77 (0.13) | 0.12 (0.07) | 1.72 (0.06) | 1.19 (0.05) | 1.77 (0.06) | -0.33 (0.06) | -1.82 (0.07) | 0.28 (0.11) | 1.61 (0.08) | 0.23 (0.04) | 0.73 (0.05) |

**QUADRO 6.5**- Funções de Informação dos Itens e Total - Muraki.

| $(\theta)$ | **-2.0** | **-1.5** | **-1.0** | **-0.5** | **0.0** | **0.5** | **1.0** | **1.5** | **2.0** |
|---|---|---|---|---|---|---|---|---|---|
| $I_1(\theta)$ | 0.23 | 0.39 | 0.57 | 0.71 | 0.78 | 0.82 | **0.85** | 0.78 | 0.60 |
| $I_2(\theta)$ | **0.13** | **0.13** | 0.12 | 0.10 | 0.09 | 0.07 | 0.06 | 0.05 | 0.04 |
| $I_3(\theta)$ | 0.37 | **0.39** | **0.39** | 0.38 | 0.35 | 0.31 | 0.26 | 0.22 | 0.17 |
| $I_4(\theta)$ | 0.16 | 0.18 | 0.22 | 0.28 | 0.36 | 0.45 | **0.50** | 0.48 | 0.39 |
| $I_5(\theta)$ | 0.45 | 0.77 | 1.28 | 1.84 | **2.02** | 1.62 | 1.04 | 0.61 | 0.35 |
| $I_6(\theta)$ | 0.25 | 0.33 | 0.48 | 0.68 | **0.83** | 0.80 | 0.60 | 0.38 | 0.22 |
| $I_7(\theta)$ | 0.33 | 0.44 | **0.51** | **0.51** | 0.44 | 0.32 | 0.22 | 0.15 | 0.10 |
| $I_8(\theta)$ | 0.26 | 0.35 | **0.41** | 0.39 | 0.31 | 0.22 | 0.14 | 0.08 | 0.05 |
| $I_9(\theta)$ | **0.31** | 0.28 | 0.24 | 0.19 | 0.15 | 0.11 | 0.08 | 0.06 | 0.05 |
| $I_{10}(\theta)$ | 0.76 | **0.90** | 0.84 | 0.68 | 0.54 | 0.45 | 0.39 | 0.33 | 0.27 |
| $I_{11}(\theta)$ | 0.14 | 0.19 | 0.26 | 0.35 | 0.52 | 0.80 | **1.09** | **1.09** | 0.75 |
| $I_{12}(\theta)$ | 0.36 | 0.61 | 1.00 | **1.36** | 1.30 | 0.89 | 0.51 | 0.29 | 0.17 |
| $I(\theta)$ | 4.7 | 6.0 | 7.3 | 8.5 | **8.7** | 7.9 | 6.7 | 5.5 | 4.1 |
| $S(\theta)$ | 0.46 | 0.41 | 0.37 | 0.34 | 0.34 | 0.36 | 0.39 | 0.43 | 0.49 |

Dos Quadros 6.3 e 6.5, as Funções de Informação Total obtidas com os modelos de Masters e Muraki dão mais informação para indivíduos entre -1 e 0 e entre -0.5 e 0.5, respectivamente. Vendo a Figura 6.14, os modelos de Masters e Muraki dão a mesma informação para valores da habilidade próximos de -0.85. Para valores da habilidade maiores que -0.85 o modelo de Muraki fornece mais informação.

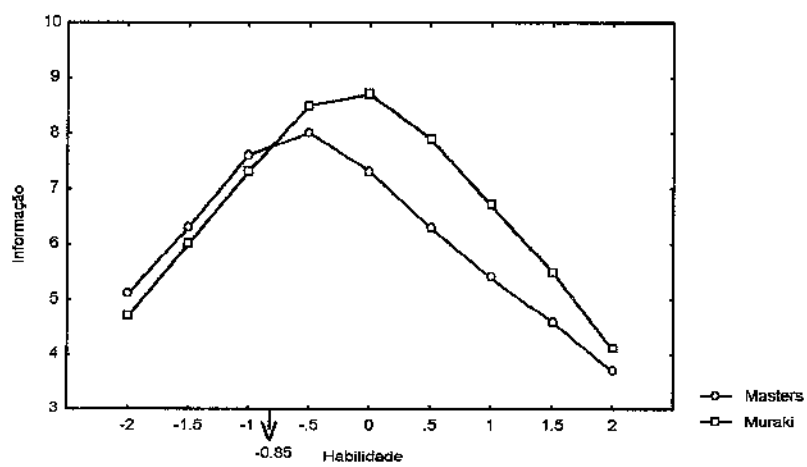



**FIGURA 6.14**- Funções de Informação Total para os modelos de Masters e Muraki.

## 6.4- ESTIMAÇÃO DOS PARÂMETROS DOS ITENS SOB O MODELO DE RESPOSTA CATEGÓRICA DE SAMEJIMA

Nesse modelo trabalhamos com probabilidades acumuladas: $P_{ik}(\theta_j) = P^*_{ik}(\theta_j) - P^*_{i,k+1}(\theta_j)$, onde $P^*_{ik}(\theta_j)$ representa a probabilidade de escolher uma categoria maior que o limite $\lambda_{k-1}$.

$$
\begin{array}{lccccccc}
k \Rightarrow & 1 & 2 & 3 & 4 & 5 & 6 & \\
 & & \lambda_{i1} & \lambda_{i2} & \lambda_{i3} & \lambda_{i4} & \lambda_{i5} & \\
 & & \downarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow P^*_{i3}(\theta_j) \\
 & & & \downarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow \quad P^*_{i4}(\theta_j)
\end{array}
$$

Portanto, para k=3, temos

$$P_{13}(\theta_j) = P^*_{13}(\theta_j) - P^*_{14}(\theta_j)$$

No modelo de RCS são estimados 72 parâmetros dos itens (Quadro 6.6), sendo $\lambda_{is}$, s=1,2,3,4,5, os parâmetros de locação associados aos limites entre as categorias. Nesse caso, como temos 6 categorias temos 5 limites entre as categorias, s=k-1.

Por exemplo, $\lambda_{13}$=-0.91 é o parâmetro de locação associado ao limite entre as categorias 3 e 4, e $\lambda_{12}$=-1.37 é o parâmetro de locação associado ao limite entre as categorias 2 e 3 do item 1.

De (3.2) temos que: a probabilidade de um indivíduo de habilidade $\theta_j$ responder na categoria 4 ou maior do item 1, é dado por,

$$P_{14}^{*}(\theta_j)=\frac{\exp\left[\alpha_1(\theta_j-\lambda_{13})\right]}{1+\exp\left[\alpha_1(\theta_j-\lambda_{13})\right]}, \qquad (6.1)$$

e a probabilidade do indivíduo responder na categoria 3 ou maior é dado por,

$$P_{13}^{*}(\theta_j)=\frac{\exp\left[\alpha_1(\theta_j-\lambda_{12})\right]}{1+\exp\left[\alpha_1(\theta_j-\lambda_{12})\right]}. \qquad (6.2)$$

Assim, de (6.1) e (6.2), a probabilidade do indivíduo de habilidade $\theta_j$ responder na categoria 3 do item 1, sob o Modelo de Resposta Categórica de Samejima (1972), formula (3.3), é dado por:

$$P_{13}(\theta_j)=P_{13}^{*}(\theta_j)-P_{14}^{*}(\theta_j)=\frac{1}{1+\exp\left[-\alpha_1(\theta_j-\lambda_{12})\right]}-\frac{1}{1+\exp\left[-\alpha_1(\theta_j-\lambda_{13})\right]} \qquad (6.3)$$

No Quadro 6.6 a seguir encontramos as estimativas dos parâmetros dos itens e, no Quadro 6.7 as funções de informação dos itens e a função de informação total obtidas através do modelo de Samejima.

Considerando $\theta_j = 0.5$, do Quadro 6.6, para o item 1 teremos,

$$P_{11}(0.5) = P_{11}^{*}(0.5) - P_{12}^{*}(0.5) = 1 - \frac{1}{1+\exp[-1.64(0.5+2.12)]} = 0.0134,$$

a probabilidade de um indivíduo de habilidade 0.5 responder na categoria 1 **(obter escore 0)** do item 1 é de **1.34%**;

$$P_{12}(0.5) = P_{12}^{*}(0.5) - P_{13}^{*}(0.5) = \frac{1}{1+\exp[-1.64(0.5+2.12)]} - \frac{1}{1+\exp[-1.64(0.5+1.37)]} = 0.0311,$$

a probabilidade de um indivíduo de habilidade 0.5 responder na categoria 2 **(obter escore 1)** do item 1 é de **3.11%**;

$$P_{13}(0.5) = P_{13}^{*}(0.5) - P_{14}^{*}(0.5) = \frac{1}{1+\exp[-1.64(0.5+1.37)]} - \frac{1}{1+\exp[-1.64(0.5+0.91)]} = 0.0456,$$

a probabilidade de um indivíduo de habilidade 0.5 responder na categoria 3 **(obter escore 2)** do item 1 é de **4.56%**;

$$P_{14}(0.5) = P_{14}^{*}(0.5) - P_{15}^{*}(0.5) = \frac{1}{1+\exp[-1.64(0.5+0.91)]} - \frac{1}{1+\exp[-1.64(0.5-0.66)]} = 0.4751,$$

a probabilidade de um indivíduo de habilidade 0.5 responder na categoria 4 **(obter escore 3)** do item 1 é de **47.51%**;

$$P_{15}(0.5) = P_{15}^{*}(0.5) - P_{16}^{*}(0.5) = \frac{1}{1+\exp[-1.64(0.5-0.66)]} - \frac{1}{1+\exp[-1.64(0.5-0.71)]} = 0.0200,$$

a probabilidade de um indivíduo de habilidade 0.5 responder na categoria 5 **(obter escore 4)** do item 1 é de **2%**, e

$$P_{16}(\theta_j) = P_{16}^{*}(\theta_j) = \frac{1}{1+\exp\left[-1.64\left(\theta_j - 0.71\right)\right]} - 0 = 0.4147,$$

a probabilidade de um indivíduo de habilidade 0.5 responder na categoria 6 **(obter escore 5)** do item 1 é de **41.47%.**

Do Capítulo 3 temos que, para um item i e um certo nível de habilidade $\sum_{k=1}^{m_i} P_{ik}(\theta_j) = 1$. Assim, para $\theta_j = 0.5$ a soma das probabilidades de resposta nas 6 categorias do primeiro item deve ser igual um,

$$\sum_{k=1}^{6} P_{1k}(0.5) = 0.0134 + 0.0311 + 0.0456 + 0.4751 + 0.02 + 0.4147 = 0.9999 \cong 1. \square$$

No item 1 o indivíduo de habilidade 0.5 tem maior probabilidade de responder nas categorias 4 e 6 **(escores 3 e 5)**, com 47.51 e 41.47%, respectivamente. Isso concorda com a interpretação gráfica dos itens na seção 6.2.1, Figura 6.2.

**QUADRO 6.6**- Estimativas dos Parâmetros dos Itens - Samejima.

| ITEM | i=1 | i=2 | i=3 | i=4 | i=5 | i=6 | i=7 | i=8 | i=9 | i=10 | i=11 | i=12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha_i$ | **1.64** (0.03) | **0.66** (0.03) | **1.06** (0.02) | **1.04** (0.03) | **2.20** (0.03) | **1.42** (0.03) | **1.18** (0.03) | **1.14** (0.03) | **0.75** (0.03) | **1.37** (0.03) | **1.50** (0.03) | **1.79** (0.03) |
| $\lambda_{i1}$ | -2.12 (0.04) | 0.49 (0.04) | -1.33 (0.04) | -1.97 (0.05) | -1.20 (0.02) | -1.13 (0.03) | -0.50 (0.03) | 0.65 (0.03) | 0.51 (0.04) | -1.64 (0.04) | -2.06 (0.04) | -0.79 (0.02) |
| $\lambda_{i2}$ | **-1.37** (0.03) | 1.68 (0.07) | -0.39 (0.03) | -1.71 (0.05) | -0.46 (0.02) | -0.53 (0.02) | 0.25 (0.02) | 0.73 (0.03) | 1.28 (0.05) | -0.26 (0.02) | -1.69 (0.04) | -0.13 (0.02) |
| $\lambda_{i3}$ | **-0.91** (0.02) | 2.21 (0.09) | 0.94 (0.03) | -0.82 (0.03) | 0.10 (0.01) | -0.04 (0.02) | 0.68 (0.03) | 0.73 (0.03) | 3.04 (0.11) | 1.30 (0.03) | -1.43 (0.03) | 0.33 (0.02) |
| $\lambda_{i4}$ | **0.66** (0.02) | 4.28 (0.17) | 2.28 (0.06) | 0.82 (0.03) | 0.67 (0.02) | 0.74 (0.02) | 1.55 (0.04) | 0.99 (0.03) | 4.25 (0.15) | 1.83 (0.04) | -0.43 (0.02) | 0.81 (0.02) |
| $\lambda_{i5}$ | 0.71 (0.02) | 7.14 (0.31) | 2.91 (0.07) | 2.96 (0.07) | 1.39 (0.02) | 2.35 (0.05) | 2.03 (0.05) | 1.18 (0.04) | 5.20 (0.19) | 2.85 (0.06) | 0.75 (0.02) | 1.54 (0.03) |

**QUADRO 6.7**- Funções de Informação dos Itens e Total - Samejima.

| $(\theta)$ | **-2.0** | **-1.5** | **-1.0** | **-0.5** | **0.0** | **0.5** | **1.0** | **1.5** | **2.0** |
|---|---|---|---|---|---|---|---|---|---|
| $I_1(\theta)$ | 0.79 | **0.84** | 0.82 | 0.76 | 0.73 | 0.74 | 0.66 | 0.46 | 0.26 |
| $I_2(\theta)$ | 0.06 | 0.07 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.13 | **0.14** |
| $I_3(\theta)$ | 0.26 | 0.31 | 0.34 | **0.35** | **0.35** | **0.35** | **0.35** | **0.35** | **0.35** |
| $I_4(\theta)$ | 0.31 | 0.33 | **0.34** | **0.34** | 0.33 | 0.33 | 0.32 | 0.31 | 0.31 |
| $I_5(\theta)$ | 0.60 | 1.12 | 1.42 | 1.50 | **1.53** | 1.52 | 1.47 | 1.28 | 0.81 |
| $I_6(\theta)$ | 0.36 | 0.50 | 0.60 | **0.64** | **0.64** | 0.63 | 0.60 | 0.57 | 0.56 |
| $I_7(\theta)$ | 0.18 | 0.26 | 0.34 | 0.40 | 0.44 | **0.45** | **0.45** | 0.44 | 0.41 |
| $I_8(\theta)$ | 0.06 | 0.10 | 0.15 | 0.22 | 0.29 | 0.35 | **0.36** | 0.33 | 0.27 |
| $I_9(\theta)$ | 0.06 | 0.08 | 0.11 | 0.13 | 0.15 | 0.16 | 0.17 | 0.17 | **0.18** |
| $I_{10}(\theta)$ | 0.46 | 0.53 | 0.55 | 0.56 | 0.55 | 0.55 | 0.56 | 0.58 | **0.59** |
| $I_{11}(\theta)$ | 0.66 | **0.71** | **0.71** | 0.70 | 0.68 | 0.65 | 0.58 | 0.42 | 0.26 |
| $I_{12}(\theta)$ | 0.30 | 0.55 | 0.83 | 0.98 | **1.03** | **1.03** | 1.00 | 0.92 | 0.70 |
| $I(\theta)$ | 5.1 | 6.4 | 7.3 | 7.7 | 7.8 | **7.9** | 7.6 | 7.0 | 5.8 |
| $S(\theta)$ | 0.44 | 0.40 | 0.37 | 0.36 | 0.36 | 0.36 | 0.36 | 0.38 | 0.41 |

# PARÂMETROS DE DISCRIMINAÇÃO DA TCT E DA TRI

Os índices de discriminação obtidos com a TCT (fornecidos pela COMVEST/UNICAMP) e, com os modelos de Muraki e Samejima dão a seguinte seqüência aos itens por ordem decrescente de discriminação:

| | 1º | 2º | 3º | 4º | 5º | 6º | 7º | 8º | 9º | 10º | 11º | 12º |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $DIS_{TCT}$ | 12 | 5 | 11 | 8 | 6 | 7 | 1 | 4 | 10 | 3 | 2 | 9 |
| $DIS_{MURAKI}$ | 5 | 12 | 10 | 1 | 11 | 6 | 4 | 3 | 7 | 9 | 8 | 2 |
| $DIS_{SAMEJIMA}$ | 5 | 12 | 1 | 11 | 6 | 10 | 7 | 8 | 3 | 4 | 9 | 2 |

Verificamos que, pela TCT, ou por modelos da TRI utilizados que os itens que discriminam mais entre os indivíduos são os itens 5 e 12. Esses itens dão mais informação para valores de habilidade em torno de zero para o item 5, considerando os modelos de Masters, Muraki e Samejima (Figura 6.17). E o item 12, em torno de -0.5 para os modelos de Masters e Muraki e entre zero e 0.5 para o modelo de Samejima (Figura 6.20). O item 8 não foi classificado com alto parâmetro de discriminação pela TRI mas é considerado um dos itens mais discriminativos da prova pela TCT, ele fornece mais informação para indivíduos de habilidade em torno de -0.5 para o modelo de Masters em torno de -1.0 para o modelo de Muraki e em torno de 1.0 para o modelo de Samejima (Figura 6.18).

Os itens 2 e 9, também em concordância entre a TCT e a TRI, são os que discriminam menos entre os indivíduos. Das Figuras 6.16 e 6.19 vemos que eles dão mais informação para indivíduos de baixa habilidade para os modelos de Masters e Muraki e de alta habilidade para o modelo de Samejima. Verificamos que o modelo de RCS tende a dar mais informação para indivíduos com valores altos da habilidade.

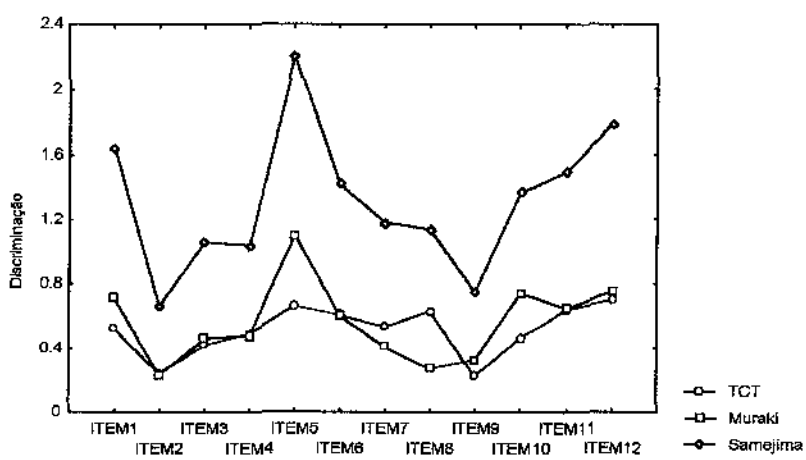



**FIGURA 6.15-** Índice de Discriminação dos Itens obtidos pela TCT e Parâmetros de Discriminação dos Itens obtidos pelos modelos de Muraki e Samejima - TRI.

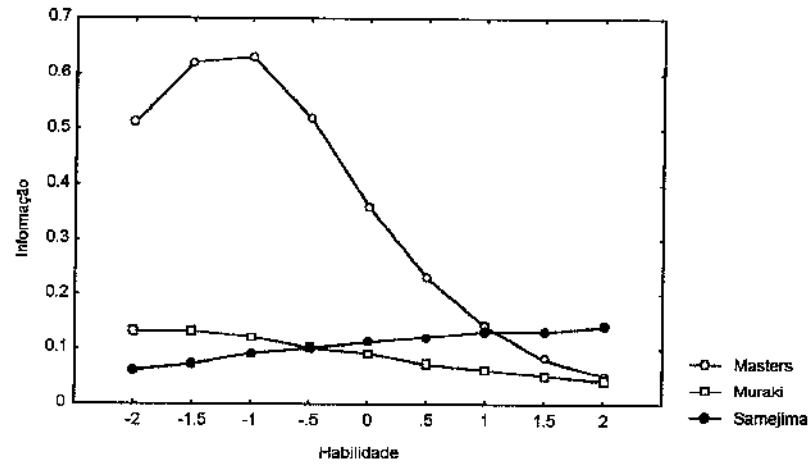



**FIGURA 6.16-** Funções de Informação do Item 2 para os modelos de Masters, Muraki e Samejima.

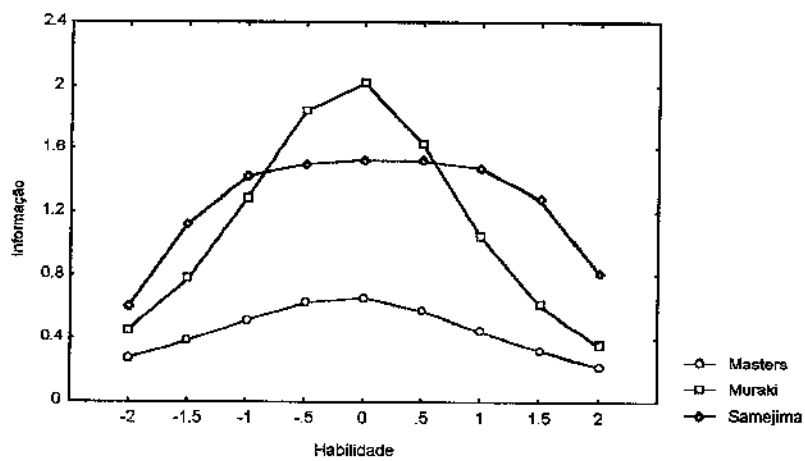



**FIGURA 6.17-** Funções de Informação do Item 5 para os modelos de Masters, Muraki e Samejima.

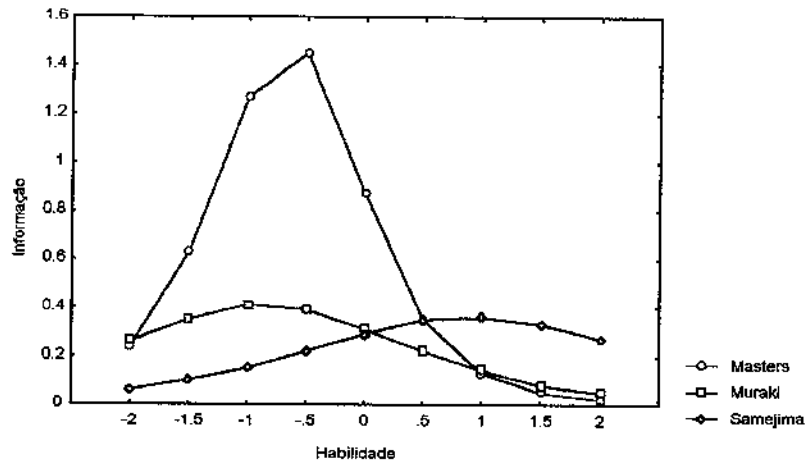



**FIGURA 6.18-** Funções de Informação do Item 8 para os modelos de Masters, Muraki e Samejima.

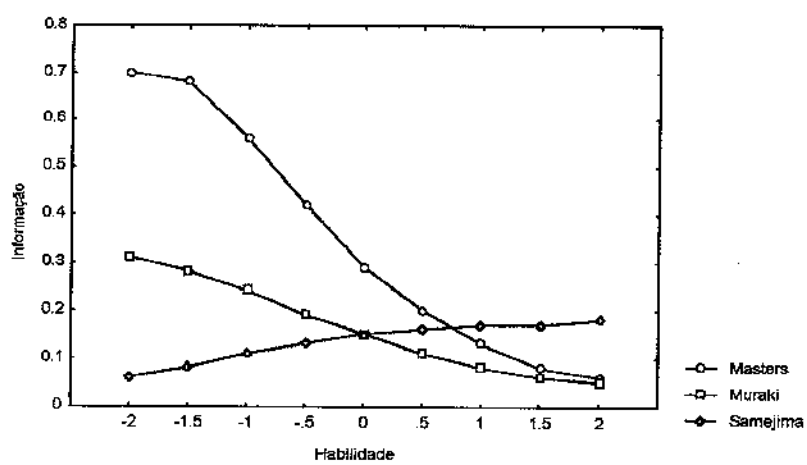



**FIGURA 6.19-** Funções de Informação do Item 9 para os modelos de Masters, Muraki e Samejima.

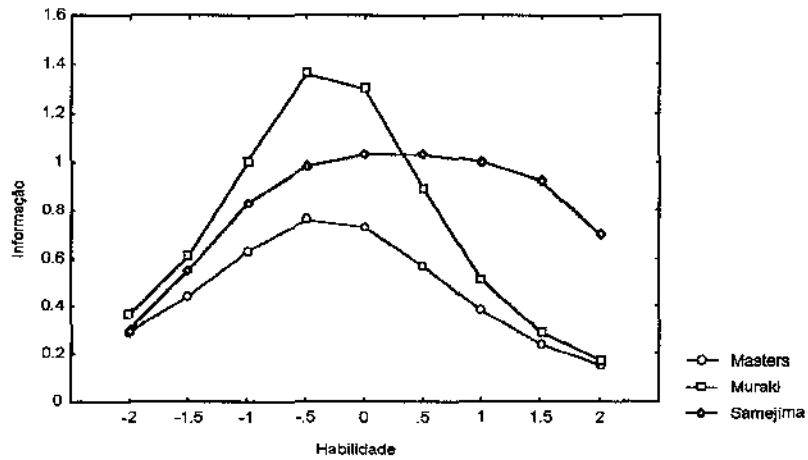



**FIGURA 6.20-** Funções de Informação do Item 12 para os modelos de Masters, Muraki e Samejima.

**CONCLUSÕES**: A Teoria de Resposta ao Item, através dos parâmetros de dificuldade dos modelos estudados nesse trabalho, permite-nos uma análise e compreensão do comportamento dos itens da Prova de Inglês. Como sugestão, alguns itens poderiam ter suas grades de correção de forma que houvesse mais equilíbrio entre as dificuldades dos 5 passos individuais em cada item. Entretanto, por se tratar de um exame sem pré-testagem (dadas suas características) isto não seria possível, uma vez que a informação só é obtida a partir dos resultados.

A Teoria de Resposta ao Item (TRI) não substitui a Teoria Clássica de Testes (TCT), ela vem como complemento, dando informações mais detalhadas sobre o comportamento dos itens de um teste.

# APÊNDICE A

## ANÁLISE FATORIAL MULTIVARIADA

### I- Introdução

A Análise Fatorial é uma técnica pertencente ao corpo em expansão das técnicas voltadas à variáveis latentes que apareceram e cresceram quase que exclusivamente nas ciências sociais e comportamento. Muitas áreas da ciência que se utilizam de dados multivariados, tem como objetivo a obtenção de estimativas de escores de indivíduos em poucas variáveis (chamadas fatores) que não podem ser observadas diretamente. O estudo da Análise Fatorial baseia-se na obtenção de fatores e escores através da estrutura da matriz de variâncias e covariâncias (ou correlações). Ou seja, a proposta é descrever as inter-relações entre variáveis diretamente observáveis e que estão relacionadas às quantidades latentes de interesse mais erro. Em outras palavras, obtidas as observações de p variáveis realizadas sobre N indivíduos, desejamos descrevê-las linearmente em função das estimativas dos escores desses fatores. Em suma, essas variáveis não-observáveis das variáveis observadas são identificadas como possíveis causas determinantes do comportamento manifestado nas observações originais. Isto significa um grande avanço na redução das informações contidas nas muitas variáveis observadas, porque essas

informações, na verdade, estão contidas nas variáveis não-observáveis ou fatores.

## II- Considerações Gerais

Geralmente, em análise multivariada trabalhamos com um número muito grande de variáveis, suponha p. Assim, tenta-se obter $m < p$ fatores aos quais são dados nomes para então trabalhar com tais fatores na análise subsequente, implicando na redução de variáveis.

Dessa forma, considere p vetores de observações $\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_p$ associados às variáveis $\mathbf{X}_1$, $\mathbf{X}_2$, ..., $\mathbf{X}_p$, de tal maneira que as componentes de vetor $\mathbf{x}_j$ ($j = 1, 2, \ldots, p$) sejam as N observações $x_{ij}$ ($i = 1, 2, \ldots, n$) da variável $\mathbf{X}_j$ sobre os indivíduos. Assim, o modelo de Análise Fatorial pressupõe a existência de m vetores ou fatores, $\mathbf{f}_1$, $\mathbf{f}_2$, ..., $\mathbf{f}_m$, associados $\mathbf{F}_1$, $\mathbf{F}_2$, ..., $\mathbf{F}_m$ que são as variáveis não-observáveis, além de p vetores de erros $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_p$, associados às $\varepsilon_1$, $\varepsilon_2$, ..., $\varepsilon_p$. Logo, os vetores de observações são gerados linearmente pelos fatores mais os erros, ou seja, cada j-ésimo vetor de observação pode ser escrito como

$$\left(x_{ij} - \bar{x}_j\right) = \sum_{k=1}^{m} \ell_{jk} f_k + e_j \quad , \quad (i = 1, 2, \ldots, N \quad e \quad j = 1, 2, \ldots, p)$$

Em termos populacionais, temos

$$_p(X-\mu)_1 = {}_pL_{mm}F_1 + {}_p\varepsilon_1 \qquad (1)$$

onde $_pX_1$: é o vetor de variáveis aleatórias;

$_p\mu_1$: é o vetor das médias de $_pX_1$;

$_pL_m$: é a matriz de cargas fatoriais;

$_mF_1$: é o vetor de fatores comuns;

$_p\varepsilon_1$: é o vetor de fatores específicos.

Assim, temos que

$$_pX_1 = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix}, \; _p\mu_1 = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix}, \; _pL_m = \begin{pmatrix} \ell_{11} & \ell_{12} & \cdots & \ell_{1m} \\ \ell_{21} & \ell_{22} & \cdots & \ell_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{p1} & \ell_{p2} & \cdots & \ell_{pm} \end{pmatrix}, \; _mF_1 = \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{pmatrix} \; e \; _p\varepsilon_1 = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{pmatrix} \qquad (2)$$

Nota-se que o modelo proposto nada mais é do que um conjunto de p modelos de regressão linear múltipla, mas com a ressalva de que o vetor de variáveis dependentes F é não-observável e, além disso, as m variáveis não-observáveis devem ser comuns às m

regressões. Então, o modelo será explorado de maneira diferente da análise de regressão, mas antes temos de ter algumas suposições:

(i) $E\ ({}_mF_1) = {}_m0_1$, ou seja, os fatores comuns têm média zero;

(ii) $E\ ({}_p\varepsilon_1) = {}_p0_1$, ou seja, os fatores específicos têm média zero;

(iii) $Cov\ ({}_mF_1) = E\ ({}_mF_1{}_mF_1) = {}_mI_m$, ou seja, os fatores são ortogonais ou incorrelatos e padronizados;

(iv) $Cov\ ({}_p\varepsilon_1) = E({}_p\varepsilon_1{}_p\varepsilon'_1) = {}_p\Psi_p = \begin{pmatrix} \psi_1 & 0 & \cdots & 0 \\ 0 & \psi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \psi_p \end{pmatrix}$, ou seja, os fatores específicos são incorrelatos. Denominaremos de matriz de especificidade;

(v) $Cov\ ({}_p\varepsilon_1,\ {}_mF_1) = E\ ({}_p\varepsilon_1{}_mF_1) = {}_pO_m$, ou seja, ausência de correlação entre fatores comuns e fatores específicos.

Seguindo as suposições acima, podemos obter:

(a.1) <u>Decomposição da matriz de variâncias e covariâncias</u>

$$\sum = \mathrm{Cov}(X) = E(X-\mu)(X-\mu)' = E\left[(LF+\varepsilon)(LF+\varepsilon)'\right]$$
$$= E\left(LFF'L' + LF\varepsilon' + \varepsilon F'L' + \varepsilon\varepsilon'\right)$$
$$= LE(FF')L' + LE(F\varepsilon') + E(\varepsilon F')L' + E(\varepsilon\varepsilon')$$
$$= LL' + \psi$$

ou seja

$$\sum = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_{pp} \end{pmatrix} = \begin{pmatrix} \ell_{11} & \ell_{12} & \cdots & \ell_{1m} \\ \ell_{12} & \ell_{22} & \cdots & \ell_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{1p} & \ell_{2p} & \cdots & \ell_{pm} \end{pmatrix} \begin{pmatrix} \ell_{11} & \ell_{12} & & \ell_{1p} \\ \ell_{12} & \ell_{22} & & \ell_{2p} \\ & & & \\ \ell_{1m} & \ell_{2m} & & \ell_{pm} \end{pmatrix} + \begin{pmatrix} \psi_1 & 0 & \cdots & 0 \\ 0 & \psi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \psi_{1p} \end{pmatrix}$$

(2

Assim,

$$\sigma_{ii} = \mathrm{Var}(X_j) = \ell_{j1}^2 + \ell_{j2}^2 + \ldots + \ell_{jm}^2 + \psi_j, \quad (j = 1,2,\ldots,p)_e, \qquad (3)$$

$$\sigma_{jk} = \mathrm{Cov}(X_j, X_k) = \ell_{j1}\ell_{k1} + \ell_{j2}\ell_{k2} + \ldots + \ell_{jm}\ell_{km}, \quad (j \neq K = 1,2,\ldots,p)$$

(a.2) Covariâncias entre variáveis originais e variáveis não-observáveis

$$\mathrm{Cov}(X,F) = E(X-\mu)(F-E(F))' = E\left[(X-\mu)F'\right] = E\left[(LF+\varepsilon)F'\right] = E\left(LFF'\right) + E\left(\varepsilon F'\right) = LE(FF') = L$$

ou

$$Cov(X_j, F_k) = \ell_{jk}, \quad (j \neq k = 1,2,...,p)$$

Na decomposição da variância dada em (3) definimos os seguintes termos $h_j^2 = \ell_{j1}^2 + \ell_{j2}^2 + ... + \ell_{jm}^2$ que será denominado de comunalidade (termo que vem de fator comum) e $\Psi_j$ de especificidade (termo proveniente de fatores específicos). Desta forma (3) pode ser reescrito como

$$\sigma_{jj} = Var(X_j) = h_j^2 + \psi_j, \quad (j = 1,2,...,p)$$

Na prática, as variáveis originais geralmente são padronizadas, ou seja, têm média zero e variância unitária. Desse modo, a matriz de covariâncias $\Sigma$ transforma-se na matriz de correlações P. Daqui em diante, adotando P como a matriz a ser decomposta temos:

(b.1) Decomposição da matriz de correlações

$$P = LL' + \Psi,$$

ou seja,

$$P=\begin{pmatrix}\rho_{11} & \rho_{12} & \cdots & \rho_{1p}\\ \rho_{12} & \rho_{22} & \cdots & \rho_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ \rho_{1p} & \rho_{2p} & \cdots & \rho_{pp}\end{pmatrix}=\begin{pmatrix}\ell_{11} & \ell_{12} & \cdots & \ell_{1m}\\ \ell_{12} & \ell_{22} & \cdots & \ell_{2m}\\ \vdots & \vdots & \ddots & \vdots\\ \ell_{p1} & \ell_{p2} & \cdots & \ell_{pm}\end{pmatrix}\begin{pmatrix}\ell_{11} & \ell_{12} & \cdots & \ell_{p1}\\ \ell_{12} & \ell_{22} & \cdots & \ell_{p2}\\ \vdots & \vdots & \ddots & \vdots\\ \ell_{1m} & \ell_{2m} & \cdots & \ell_{pm}\end{pmatrix}+\begin{pmatrix}\psi_1 & 0 & \cdots & 0\\ 0 & \psi_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \psi_p\end{pmatrix} \quad (4)$$

Assim

$$\rho_{jj}=Corr(X_j)=1=\ell_{j1}^2+\ell_{j2}^2+\ldots+\ell_{jm}^2+\psi_j,\quad (j=1,2,\ldots,p) \text{ e,} \quad (5)$$

$$\rho_{jk}=\mathrm{Corr}(X_j,X_k)=\ell_{j1}^2\ell_{k1}+\ell_{j2}^2\ell_{k2}+\ldots+\ell_{jm}^2\ell_{km},\quad (j\neq k=1,2,\ldots p)$$

(b.2) <u>Correlação entre variáveis originais e variáveis não-observáveis</u>

Corr(X, F)=L ou $\mathrm{Corr}(X_j,F_k)=\ell_{jk}$, (j=1,2,..., p e k = 1,2,..., m). Como foi visto em (a.2), a decomposição dos elementos da diagonal de P (que é igual a 1) pode ser decomposta da seguinte maneira:

$$1=\mathrm{Corr}(X_j,X_j)-h_j^2+\psi_j,(j=1,2,\ldots,p)$$

Observação: Note que, apesar das notações para Σ e P serem os mesmos, pode-se distinguir no contexto as diferenças dos termos

$\ell_{jk}$, $h_j^2$ e $\psi_j$ quando utilizamos a matriz de variâncias e covariâncias e a matriz de correlações.

## III-Metódo de Estimação - Componentes Principais

Este método tenta recuperar a decomposição de P dada em (4) e (5). Assim, por decomposição espectral, como P que é simétrica e positiva definida com par de autovalores-autovetores ($\lambda_j$, $\mathbf{u}_i$), (j=1, 2, ..., p) onde $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p > 0$ e $\geq \mathbf{u}_i$ (j = 1, 2, ..., p) são autovetores normalizados, pode ser expressa da seguinte forma:

$$P = Q\Lambda Q' = Q\Lambda^{1/2}\Lambda^{1/2}Q' = L_1 L_1'$$

onde $L_1 = \left(\sqrt{\lambda_1 u_1} \quad \sqrt{\lambda_2 u_2} \ldots \sqrt{\lambda_p u_p}\right)$.

Eliminando as colunas correspondentes aos menores autovalores, ou seja, aqueles que contribuem pouco para P, temos

$$P \cong LL'$$

onde $L = \left(\sqrt{\lambda_1 u_1} \quad \sqrt{\lambda_2 u_2} \ldots \sqrt{\lambda_m u_m}\right)$ com m <p

Se incluirmos a variação do fator específico (matriz de especificidade) temos,

$$P \cong LL' + \psi \qquad (6)$$

onde a igualdade é exata na diagonal, valendo apenas as aproximações fora da diagonal principal. A aproximação aumenta na medida em que o número de componentes principais m aproxima-se de p.

A aplicação de (6) na matriz de correlações amostrais R, é conhecido como a solução de componentes principais. Como estamos interessados em obter as estimativas de L, que são as cargas fatoriais, temos as seguintes estimativas,

$$\hat{L} = \left(\sqrt{\hat{\lambda}_1\hat{u}_1} \quad \sqrt{\hat{\lambda}_2\hat{u}_2} \ldots \sqrt{\hat{\lambda}_m\hat{u}_m}\right)$$

e $R \cong \hat{L}\hat{L}' + \hat{\psi}$. Consequentemente, $\hat{\psi}_j = 1 - \hat{h}_j^2, (j = 1,2,\ldots,p)$ que é a estimativa dos elementos da matriz de especificidade.

Para verificar a proporção da variabilidade acumulada pelos m primeiros fatores utilizamos o seguinte critério

$$\begin{cases} \dfrac{\sum_{j=1}^{m} \hat{\lambda}_j}{s_{11} + s_{22} + \ldots + s_{pp}} \text{ para analise fatorial de S} \\ \dfrac{\sum_{j=1}^{m} \hat{\lambda}_j}{p} \quad \text{para analise fatorial de R} \end{cases}$$

Em termos práticos, poder-se-ia reter aqueles fatores associados aos autovalores maiores que 1. Se a proporção da variabilidade explicada pelos autovalores maiores que 1 for baixa, toma-se os demais fatores até que se consiga uma proporção adequada de variabilidade, geralmente entre 80% e 90%.

**PROGRAMA - SAS**

```
DATA FATORIAL;
INPUT NUM 1-5 ITEM1 6-6 ITEM2 7-7 ITEM3 8-8 ITEM4 9-9
              ITEM5 10-10 ITEM6 11-11 ITEM7 12-12
              ITEM8 13-13 ITEM9 14-14 ITEM10 15-15
              ITEM11 16-16 ITEM12 17-17;
PONTOS=SUM(ITEM1,ITEM2,ITEM3,ITEM4,ITEM5,ITEM6,ITEM7,
           ITEM8,ITEM9,ITEM10,ITEM11,ITEM12);
CARDS;
PROC PRINT DATA FATORIAL;
TITLE 'DADOS DA PROVA DE INGLÊS DO VESTIBULAR DA
UNICAMP/97';
VAR ITEM1 ITEM2 ITEM3 ITEM4 ITEM5 ITEM6 ITEM7 ITEM8
    ITEM9 ITEM10 ITEM11 ITEM12 PONTOS;
PROC FACTOR DATA=FATORIAL OUT=FATO N=4 SIMPLE CORR
            RESIDUAL PREPLOT SCORE;
TITLE 'ANÁLISE FATORIAL';
VAR ITEM1 ITEM2 ITEM3 ITEM4 ITEM5 ITEM6 ITEM7 ITEM8
    ITEM9 ITEM10 ITEM11 ITEM12;
PROC PRINT DATA=FATO;
VAR PONTOS FACTOR1-FACTOR4;
PROC CORR DATA=FATO;
VAR FACTOR1-FACTOR4;
RUN;
```

## ANÁLISE FATORIAL

Means and Standard Deviations

| | ITEM1 | ITEM2 | ITEM3 | ITEM4 |
|---|---|---|---|---|
| Mean | 3.4959115 | 1.13323713 | 2.11688312 | 2.98701299 |
| Std Dev | 1.44529218 | 1.44951002 | 1.40184376 | 1.32201604 |

| | ITEM5 | ITEM6 | ITEM7 | ITEM8 |
|---|---|---|---|---|
| Mean | 2.86820587 | 2.6979317 | 2.06806157 | 1.98196248 |
| Std Dev | 1.61837661 | 1.59764397 | 1.81577156 | 2.3334927 |

| | ITEM9 | ITEM10 | ITEM11 | ITEM12 |
|---|---|---|---|---|
| Mean | 1.18879269 | 1.98292448 | 3.84728235 | 2.46031746 |
| Std Dev | 1.5287162 | 1.2748174 | 1.32390034 | 1.82167161 |

Correlations

| | **ITEM1** | **ITEM2** | **ITEM3** | **ITEM4** | **ITEM5** | **ITEM6** |
|---|---|---|---|---|---|---|
| ITEM1 | 1.00000 | 0.17284 | 0.25254 | 0.27947 | 0.46052 | 0.32597 |
| ITEM2 | 0.17284 | 1.00000 | 0.13949 | 0.11865 | 0.20048 | 0.14058 |
| ITEM3 | 0.25254 | 0.13949 | 1.00000 | 0.22836 | 0.29446 | 0.20545 |
| ITEM4 | 0.27947 | 0.11865 | 0.22836 | 1.00000 | 0.32740 | 0.25691 |
| ITEM5 | 0.46052 | 0.20048 | 0.29446 | 0.32740 | 1.00000 | 0.48952 |
| ITEM6 | 0.32597 | 0.14058 | 0.20545 | 0.25691 | 0.48952 | 1.00000 |
| ITEM7 | 0.28908 | 0.14636 | 0.22709 | 0.22935 | 0.38518 | 0.28530 |
| ITEM8 | 0.25936 | 0.10860 | 0.17192 | 0.19292 | 0.30576 | 0.22683 |
| ITEM9 | 0.19519 | 0.06768 | 0.14640 | 0.17262 | 0.22806 | 0.18548 |
| ITEM10 | 0.34184 | 0.13701 | 0.19293 | 0.24866 | 0.40455 | 0.31625 |
| ITEM11 | 0.31831 | 0.12907 | 0.20016 | 0.23720 | 0.42432 | 0.31574 |
| ITEM12 | 0.40666 | 0.16754 | 0.26040 | 0.28487 | 0.50094 | 0.37428 |
| | **ITEM7** | **ITEM8** | **ITEM9** | **ITEM10** | **ITEM11** | **ITEM12** |
| ITEM1 | 0.28908 | 0.25936 | 0.19519 | 0.34184 | 0.31831 | 0.40666 |
| ITEM2 | 0.14636 | 0.10860 | 0.06768 | 0.13701 | 0.12907 | 0.16754 |
| ITEM3 | 0.22709 | 0.17192 | 0.14640 | 0.19293 | 0.20016 | 0.26040 |
| ITEM4 | 0.22935 | 0.19292 | 0.17262 | 0.24866 | 0.23720 | 0.28487 |
| ITEM5 | 0.38518 | 0.30576 | 0.22806 | 0.40455 | 0.42432 | 0.50094 |
| ITEM6 | 0.28530 | 0.22683 | 0.18548 | 0.31625 | 0.31574 | 0.37428 |
| ITEM7 | 1.00000 | 0.20757 | 0.17277 | 0.25771 | 0.25290 | 0.32005 |
| ITEM8 | 0.20757 | 1.00000 | 0.42249 | 0.23069 | 0.21799 | 0.27942 |
| ITEM9 | 0.17277 | 0.42249 | 1.00000 | 0.20631 | 0.16401 | 0.23666 |
| ITEM10 | 0.25771 | 0.23069 | 0.20631 | 1.00000 | 0.29507 | 0.38479 |
| ITEM11 | 0.25290 | 0.21799 | 0.16401 | 0.29507 | 1.00000 | 0.41906 |
| ITEM12 | 0.32005 | 0.27942 | 0.23666 | 0.38479 | 0.41906 | 1.00000 |

Initial Factor Method: Principal Components

Prior Communality Estimates: ONE

Eigenvalues of the Correlation Matrix: Total = 12 Average = 1

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Eigenvalue | 3.9770 | 1.1053 | 0.9363 | 0.8622 |
| Difference | 2.8717 | 0.1690 | 0.0741 | 0.0906 |
| Proportion | 0.3314 | 0.0921 | 0.0780 | 0.0719 |
| Cumulative | 0.3314 | 0.4235 | 0.5015 | 0.5734 |

| | 5 | 6 | 7 | 8 |
|---|---|---|---|---|
| Eigenvalue | 0.7716 | 0.7516 | 0.7077 | 0.6870 |
| Difference | 0.0201 | 0.0438 | 0.0208 | 0.0364 |
| Proportion | 0.0643 | 0.0626 | 0.0590 | 0.0572 |
| Cumulative | 0.6377 | 0.7003 | 0.7593 | 0.8166 |

| | 9 | 10 | 11 | 12 |
|---|---|---|---|---|
| Eigenvalue | 0.6505 | 0.5701 | 0.5408 | 0.4400 |
| Difference | 0.0805 | 0.0293 | 0.1008 | |
| Proportion | 0.0542 | 0.0475 | 0.0451 | 0.0367 |
| Cumulative | 0.8708 | 0.9183 | 0.9633 | 1.0000 |

4 factors will be retained by the NFACTOR criterion.

Factor Pattern

| | FACTOR1 | FACTOR2 | FACTOR3 | FACTOR4 |
|---|---|---|---|---|
| ITEM1 | 0.65118 | -0.10501 | -0.02547 | -0.02273 |
| ITEM2 | 0.31119 | -0.22260 | 0.83314 | -0.37363 |
| ITEM3 | 0.46375 | -0.09528 | 0.29974 | 0.66050 |
| ITEM4 | 0.51533 | -0.05959 | 0.02008 | 0.42157 |
| ITEM5 | 0.77185 | -0.13934 | -0.09267 | -0.06937 |
| ITEM6 | 0.62571 | -0.13747 | -0.17613 | -0.13118 |
| ITEM7 | 0.55403 | -0.10572 | 0.04416 | 0.11486 |
| ITEM8 | 0.51310 | 0.63697 | 0.10078 | -0.07128 |
| ITEM9 | 0.43680 | 0.73715 | 0.08412 | -0.01548 |
| ITEM10 | 0.60277 | -0.05290 | -0.15736 | -0.16626 |
| ITEM11 | 0.60010 | -0.15311 | -0.22399 | -0.16938 |
| ITEM12 | 0.70973 | -0.08506 | -0.13238 | -0.10610 |

Variance explained by each factor

| FACTOR1 | FACTOR2 | FACTOR3 | FACTOR4 |
|---|---|---|---|
| 3.976955 | 1.105302 | 0.936270 | 0.862214 |

Final Communality Estimates: Total = 6.880741

| ITEM1 | ITEM2 | ITEM3 | ITEM4 | ITEM5 | ITEM6 |
|---|---|---|---|---|---|
| 0.436221 | 0.980117 | 0.750240 | 0.447247 | 0.628568 | 0.458644 |

| ITEM7 | ITEM8 | ITEM9 | ITEM10 | ITEM11 | ITEM12 |
|---|---|---|---|---|---|
| 0.333273 | 0.684250 | 0.741501 | 0.418530 | 0.462423 | 0.539727 |

# APÊNDICE B

## PROVA DE INGLÊS DO VESTIBULAR/97 - UNICAMP

# INGLÊS

**Responda a todas as perguntas EM PORTUGUÊS.**

Leia o texto abaixo e responda à questão **13:**

LAZY JANE

Lazy
lazy
lazy
lazy
lazy
lazy
Jane,
she
wants
a
drink
of
water
so
she
waits
and
waits
and
waits
and
waits
and
waits
for
it
to
rain.

(Silberstein, S. *Where the sidewalk ends.* Harper & Row Publishers, New York, 1974, p.87)

**13.** Por que Jane está nessa posição?

Leia o texto abaixo e responda às questões **14, 15 e 16**:

# Sixty million 'missing' women

All things being equal, it is the female of the species who is the stronger of the two. She is more likely to survive birth and to live longer this biological advantage is reflected in countries such as Britain and America in a population containing more females than males (the ratio is about 105 to 100).

But all things are not equal in parts of Asia, northern Africa and China. There the ratio is weighted toward the male. Statistics compiled by the United Nations show that in India there are 93 females per 100 males. In Bangladesh and Afghanistan the ratio is 94 to 100 and in Pakistan it is 92. Converting official Chinese figures show that the ratio there is about 88 females to 100 males.

Every increment on that ratio represents hundreds of thousands and in some cases millions of women and girls. In the case of Pakistan, for instance, there are 6.3 million more men than women. Overall it has been estimated that up to 60 million women are 'missing' in the world.

It is not known how many of those were due to infanticide or to abortion after ultrasound determined the sex of the foetus. 'One of the reasons it is difficult to get information is that infanticide is illegal and so it's almost impossible to monitor,' said Tom Lawson of the UN's children's agency (Unicef).

'Having a son is like having two eyes, having a daughter is like having only one eye,' goes the proverb. And so it seems in many of the countries where the preference among mothers is for sons. The places with such preference include Pakistan, India and Bangladesh. Somehow, as the statistics here show, their wishes do come true.

(*The National Times*, October 1996)

**14.** Explique porque a palavra ***"missing"*** está entre aspas no texto.

**15.** Por que o infanticídio e o aborto são mencionados no texto?

**16.** Explique o sentido do provérbio *"Having a son is like having two eyes; having a daughter is like having only one eye"*, relacionando-o com o restante do parágrafo em que ele aparece.

As questões **17** e **18** dizem respeito ao texto abaixo:

**Alcohol, Driving and Youth**

Drunk driving is the most recent catalyst for public activism against alcohol abuse. At the end of the 1970s, two groups appeared with the goal of combating alcohol-related accidents: Remove Intoxicated Drivers (RID) on the East Coast and Mothers Against Drunk Driving (MADD) in California. Both groups attacked weak drunk-driving laws and judicial laxness, especially in cases where drivers may have been repeatedly arrested for drunk driving – including some who had killed others in crashes – but never imprisoned.

Across the nation RID and MADD have strengthened the drunk-driving laws. Although sometimes at odds with each other, both have successfully lobbied for laws reducing the legal threshold of intoxication, increasing the likelihood of incarceration and suspending drivers' licenses without a hearing if their blood alcohol levels exceed a state's legal limit, typically about 0.1 percent.

In 1981 Students Against Driving Drunk (SADD) was established to improve the safety of high school students. The group promotes a contract between parents and their children in which the children agree to call for transportation if they have been drinking, and the parents agree to provide it. As a result, however, RID and MADD have accused SADD of sanctioning youthful drinking rather than trying to eliminate it.

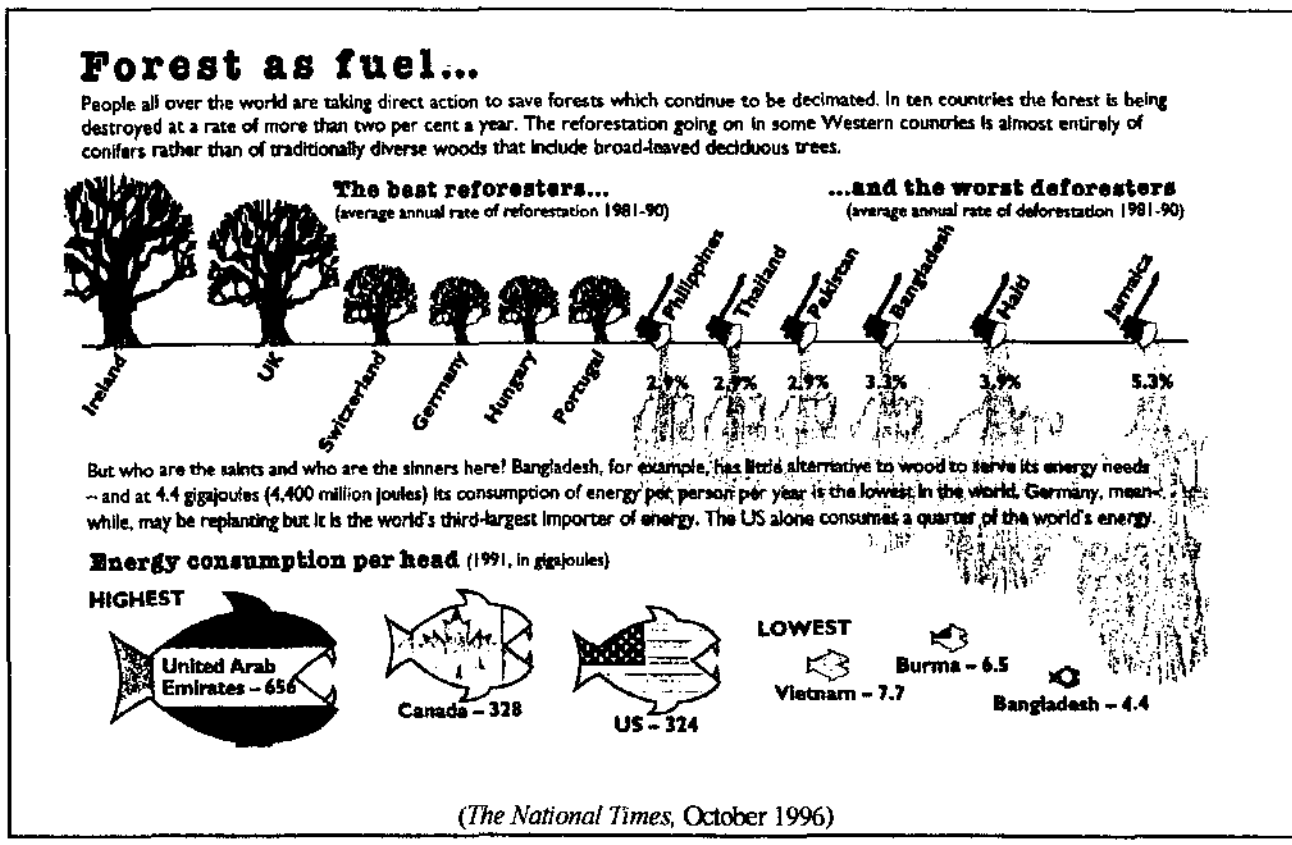
(Alcohol in American History. *Scientific American*, April 1996)

**17.** Transfira o quadro abaixo para seu caderno de respostas, preenchendo-o com informações obtidas a partir do texto:

| ASSOCIAÇÕES | RID | MADD | SADD |
|---|---|---|---|
| a) Origem | | | |
| b) Atuando desde ... | | | |
| c) Forma de atuação | | | |

**18.** O texto explicita um conflito entre associações. Que conflito é esse?

Leia o texto abaixo e responda à questão **19**:




(*The National Times*, October 1996)

**19.** O texto coloca uma pergunta. Ao invés de respondê-la, cita exemplos. Responda à pergunta a partir dos exemplos.

As questões **20** e **21** dizem respeito ao trecho abaixo:

STRICTLY SPEAKING, this story should not be written or told at all. To write it or to tell it is to spoil it. This is because the man who had the strange experience we are going to talk about never mentioned it to anybody, and the fact that he kept his secret and sealed it up completely in his memory is the whole point of the story. Thus we must admit that handicap at the beginning – that it is absurd for us to tell the story, absurd for anybody to listen to it and unthinkable that anybody should believe it.

We will, however, do this man one favour. We will refrain from mentioning him by his complete name. This will enable us to tell his secret and permit him to continue looking his friends in the eye. But we can say that his surname is Duffy. There are thousands of these Duffys in the world; even at this moment there is probably a new Duffy making his appearance in some corner of it. We can even go so far as to say that he is John Duffy's brother. We do not break faith in saying so, because if there are only one hundred John Duffys in existence, and even if each one of them could be met and questioned, no embarrassing enlightenments would be forthcoming. That is because the John Duffy in question never left his house, never left his bed, never talked to anybody in his life and was never seen by more than one man. That man's name was Gumley. Gumley was a doctor. He was present when John Duffy was born and also when he died, one hour later.

(O'Brien, F. *Stories and Plays*. Penguin Books, 1974, p.91.)

**20.** Quem é o personagem principal da estória que o texto vai contar?

**21.** Que tipo de estória vai ser contada? Justifique sua resposta.

Leia os dois trechos a seguir e responda às questões 22, 23 e 24:

# Heat of Lotus Attracts Insects And Scientists

By CAROL KAESUK YOON

THE sacred lotus, which gracefully raises its elegant flower above the murky waters where it grows, has long been revered by both Buddhists and Hindus. Now a study has recruited biologists as well as the flower's devotees. Researchers report that the lotus has the remarkable ability to regulate the temperature of its flowers to within a narrow range just as humans and other warmblooded animals do.

Dr. Roger S. Seymour and Dr. Paul Schultze-Motel, physiologists at the University of Adelaide in Australia, found that lotus flowers blooming in the Adelaide Botanic Gardens maintained a temperature of 86 to 95 degrees Fahrenheit, even when the air temperature dropped to 50 degrees. They suspect the flowers may be turning up the heat for the benefit of their coldblooded insect pollinators.

(*The New York Times*, October 1, 1996)

nature science update

## Hothouse flowers

The sacred lotus (*Nelumbo nucifera*) features strongly in the mythology of many ancient cultures and religions, and it has now excited scientific interest. A report in the 26 September issue of *Nature* shows that the lotus not only generates heat, but can regulate its temperature in much the same was as 'warm-blooded' birds and mammals.

The emergence from the mud of a beautiful white flower has been taken by Buddhists as a symbol of life and rebirth. Not without cause -- research by Dr Jane Shen-Miller at the University of California, Los Angeles and colleagues, published in the *American Journal of Botany*, showed how lotus seeds, preserved in the sediment of Chinese monastery ponds for more than a thousand years, could still germinate. The ancient Egyptians were also captivated by the flower, believing that the sun God Re emerged from a sacred blue lotus that grew out of the mud. And the lotus-eaters described by Homer in the Odyssey, intoxicated by the fruit of the lotus, lost track of the passage of time.

Dr Roger Seymour and Dr Paul Schultze-Motel, botanists at the University of Adelaide, Australia, are equally captivated by the sacred lotus. Although lotuses, like some other plants, are known to generate heat, the researchers were surprised to discover that the lotus could maintain a constant temperature even as ambient air temperature fell.

The researchers speculate on why the lotus should go to the considerable trouble of maintaining a constant 30 °C. It so happens that this is precisely the temperature that insects such as bees and beetles need to reach before they are warm enough to fly. These insects are pollinators of the sacred lotus, so perhaps the flower keeps these insects warm so that they can fly off, without extra effort, to another lotus flower, taking the pollen with them.

©NATURE NEWS SERVICE 1996

**22.** Dentre as descobertas científicas recentes sobre a flor de lotus citadas no texto da *Nature*, uma parece justificar a dimensão simbólica que ela tem para os budistas.
**a)** Qual é a descoberta?
**b)** Qual é a dimensão simbólica que ela tem para os budistas?

**23.** Transfira o quadro abaixo para seu caderno de respostas, preenchendo-o com informações obtidas a partir dos **dois** textos:

| | **Propriedade térmica** |
|---|---|
| **a) Algumas plantas** | |
| **b) Flor de lotus** | |

**24.** O texto do ***The New York Times*** afirma, no último parágrafo: "*They [the phisiologists] suspect the flowers may be turning up the heat for the benefit of their coldblooded insect pollinators*". Use as informações do texto da ***Nature*** para explicar os aspectos grifados da afirmação.

# APÊNDICE C

## PROGRAMAS PARA O USO DO MULTILOG

### MODELO DE CRÉDITO PARCIAL DE MASTERS

```
PRO RA IN NI=12 NG=1 NE=12912;
TEST ALL NO NC=(6(0)12) HI=(6(0)12);
TMATRIX ALL CK TRIANGLE;
TMATRIX ALL AK POLYNOMIAL;
FIX ALL AK=(2,3,4,5) VALUE=0.0;
EQUAL ALL AK=1;
EST NC=25 IT=2 BI;
END;
      6
012345
111111111111
222222222222
333333333333
444444444444
555555555555
666666666666
(5X,12A1)
```

### MODELO DE CRÉDITO PARCIAL GENERALIZADO DE MURAKI

```
PRO RA IN NI=12 NG=1 NE=12912;
TEST ALL NO NC=(6(0)12) HI=(6(0)12);
TMATRIX ALL CK TRIANGLE;
TMATRIX ALL AK POLYNOMIAL;
FIX ALL AK=(2,3,4,5) VA=0.0;
EST NC=25 IT=2 BI;
END;
      6
012345
111111111111
222222222222
333333333333
444444444444
555555555555
666666666666
(5X,12A1)
```

## MODELO DE RESPOSTA CATEGÓRICA DE SAMEJIMA

```
PRO RA IN NI=12 NG=1 NE=12912;
TEST ALL GRADED NC=(6(0)12);
EST NC=25 IT=2 BI;
END;
      6
012345
111111111111
222222222222
333333333333
444444444444
555555555555
666666666666
(5X,12A1)
```

# REFERÊNCIAS BIBLIOGRÁFICAS


- ANDERSEN, E. B. (1972). **The numeral solution of a set of conditional estimation equations**. The Journal of the Royal Statistical Society, series B, 34, 42-54.

- ANDERSEN, E. B. (1977). **Sufficient statistics and latent trait models**. Psychometrika, 42, 69-81.

- ANDERSEN, E. B., e MADSEN, M. (1977). **Estimating the parameters of the latent population distribution.** Psychometrika, 42, 357-374.

- ANDRICH, D. (1982). **An Extension of the Rasch Model for Ratings Providing Both Location and Dispersion Parameters.** Psychometrika, 47, 105-113.

- BAKER, F.B. (1992). **Item Response Theory**: Parameter Estimation Techniques. New York: Marcel Dekker.

- BAKER, R. (1987). **Classical Test Theory and Item Response Theory in Test Analysis**. Special Report n.2: Language Testing Update. University of Edinburgh.

- BIRNBAUM, A. (1957). **Efficient design and use of tests of a mental ability for various desicion-making problems** (series Report 58-16, Project No. 7755-23). USAF School of Aviation Medicine, Randolph Air Force Base, Texas.

- BIRNBAUM, A. (1958). **"On the estimation of mental ability"**. Series report No. 15. USAF School of Aviation Medicine, Randolph Air Force Base, Texas.

- BIRNBAUM, A. (1968). **Parte V. In F. M. Lord e M. R. Novick (Eds.), Statistical Theories of Mental Test Scores.** Reading, Addison-Wesley.

- BOCK, R. D. (1972). **Estimation item parameters and latent ability when responses are scored in two or more nominal categories**. Psychometrika, 37, 29-51.

- BOCK, R. D. e AITKIN, M. (1981). **Marginal Maximum Likelihood Estimation of Item Parameters: Application an EM Algorithm**. Psychometrika, 46, 443-459.

- BOCK, R. D. e LIEBERMAN, M. (1970). **Fitting a Response Model for Dichotomously Scored Items**. Psychometrika, 35, 179-197.

- CARLSON, J.E. (1996). **Information Provided by Polytomous end Dichotomous Items on Certain NAEP Instruments**. Paper presented at the annual meeting of the American Educational Research Association, New York.

- CROCKER, L. E ALGINA, J. (1986). **Introduction to Classical and Modern Test Theory**. Holt Rinehart Winston.

- DE AYALA, R.J.; DODD, B.G. e KOCH, W.R. (1989). **A Comparison of the Graded Response and Partial Credit Models for Assessing Writing Ability**. Paper presented at the annual meeting of the National Council on Measurement in Education. San Francisco.

- DEVELLIS, ROBERTS F. (1991). **Scale Development: Theory and Applications. Applied Social Research Methods Series.** Volume 26. SAGE Publications.

- DEMPSTER, A .P., LAIRD, N. M., RUBIN, D. B. (1977). **Maximum Likelihood from incomplete data via the EM algorithm ( with discussion).** Journal of the Royal Stat. Soc., Series B, 39, 1-38.

- FERNÁNDEZ, J.M. (1990). **Teoria de Respuesta a los Ítens:** Um nuevo enfoque en la evolución psicológica y educativa. Madrid: Ediciones Pirámide.

- FERRARA, S.; WALKER-BARTNICK, L. (1989). **Constructing an Essay Prompt Bank Using the Opartial Credit Modelo**. Paper presented at the annual meeting of the National Council on Measurement in Education. San Francisco.

- FISCHER, G.H. e MOLENAAR, J.W. (1995). **Rasch Models**: Foundations, Recent Developments, and Applications. Springer-verlag, New York.

- GOKHALE, D.V. e KULLBACK, S. (1978). **The Information in Contingency Tables.** Marcel Dekker. New York.

- HAMBLETON, R. K. (1979). **Latent trait models and their applications. In Traub, R (Ed) New Directions for Testing and Measurement**. Methodological Developments. San Francisco: Jossey Bass.

- HAMBLETON, R. K. E SWAMINATHAN, H. (1985). **Item Response Theory. Principles and Applications**, Boston, M. A., Kluwer-Nijhoff Pub.

- HAMBLETON, R. K. E SWAMINATHAN, H. e ROGERS, H. J. (1991). **Fundamentals of Item Response Theory.** Springer.

- HAMBLETON, R. K. E van der LINDEN, W. J. (1996). **Modern Item Response Theory.** Springer.

- JOHNSON, R.A. e WICHERN, D.W. (1992). **Applied Multivariate Statistical Analysis.** 3.ed. Prentice Hall. Englewood Cliffs, New Jersey.

- KOLAKOWSKI, D., and BOCK, R. D. (1973b). LOLOG: **Maximum likelihood item analysis and test scoring: Logistic model for multiple item responses.** Ann arbor, MI: National Educational Resources, Inc.

- LORD, F.M. (1980). **Aplications of Item Response Theory to Practional Testing problems.** Lawrence, Hillsdale.

- LORD, F.M. e NOVICK, M. R. (1968). **Statistical Theories of Mental Test Scores.** Reading, Addison-Wesley.

- LAWLEY, D. N. (1944). **The factorial analysis of multiple item tests.**Proceedings of the Royal Society of Edinburgh, 62A, 74-82.

- LUENBERGER, D.G. Capítulo 10: **Optimization By Vector Space Methods.** Stanford University, California.

- MASTERS, G.N. (1982). **A Rasch Model for Partial Credit Scoring.** Psychometrika, 47, 149-174.

- MASTERS, G.N. (1984). **Constructing an Item Bank Using Partial Credit Scoring.** Journal of Educational Measurement, 21, 19-32.

- MASTERS, G.N. e WRIGHT, B.D. (1984). **The essential Process in a Family of Measurement Models**. Psychometrika, 49, 529-544.

- MISLEVY, R. J. , e BOCK, R. D. (1984). **BILOG I maximum likelihood item analysis and test scoring: Logistic model**. Mooresville, ID: Scientific Software.

- MOLENAAR, I. W. (1983). **Some improved diagnostics for failure of the Rasch model**. Psychometrika, 48 (1), 49-72.

- MURAKI, E. (1990). **Fitting a Polytomous Model to Likert - Type Data**. Applied Psychological Measurement, 14, 59-71.

- MURAKI, E. (1992). **A Generalized Patrial Credit Model: Application of an EM algorithm**. Applied Psychological Measurement, 16, 159-176.

- PIVATTO, M.M. (1992). **Modelos para Testes com Respostas Dicotômicas com Principal Enfoque em Teoria de Resposta ao Item**. Tese de Mestrado - UNICAMP.

- RASCH, G. (1960). **Probabilistic Models for Some Intelligence and Attainment Tests**. Copenhagen: Danish Institute for Educational Research.

- RASCH, G. (1980). **Probabilistic Models for some Intelligence and Attainment Tests**, with forward and afterword by B. D. Wright. Chicago: University of Chicago Press.

- RICHARDSON, M. W. (1936). **The relationship between difficulty and the differential validity of a test**. Psychometrika, 1, 33-39.

- RIBEIRO, S.C.; PESSOA, D.; KEIN, R.; UCHÔA, C.E.F. e FONTANIVE, N.S. (1980). **Flutuação de Critéiros na Avaliação de Redações**. Rio de Janeiro.

- ROSENBAUM, P.R. (1984). **Testing the conditional Independence and Monotonicity Assumptions of Item Response Theory**. Psychometrika, 49, 425-435.

- SAMEJIMA, F. (1969). **Estimation of latent ability using a response pattern of graded scores**. Psychometric Monograph, n.17.

- SAMEJIMA, F. (1972). **A General Model for Free Response Data**. Psychometrika Monograph Supplement, n.18.

- SAMEJIMA, F. (1972). **A Comment on Birnbaum's Three-parameter Logistic Model in the Latent Trait Theorys**.

- SAMEJIMA, F. **Homogeneous Case of the Continuous Response Model.**

- SAMEJIMA, F. (1979). **A new family of models for the multiple choice item**. (Research Report No. 79-4) Knoxville: University of Tennessee, Department of Psychology.

- STOCKING, M. L. e PEARLMAN, M. A. (1989). **Item Response Theory for Test Developers**. Educational Testing Service. Princeton, New Jersey.

- SUEN, H.K. (1990). **Principles of Test Theories**. Lawrence Erlbaum Associates. Hillsdale, New Jarsey.

- TÉCNICAS de análise Multivariada. Apostila elaborada pelos alunos da disciplina de Análise Multivariada sob a orientação da Professora Doutora Regina C. de C. P. Moran. UNICAMP/1996.

- THISSEN, D. (1982). **Marginal Maximum Likelihood Estimation for the One-parameter Logistic Model**. Psychometrika, 47, 175-186.

- THISSEN, D. e STEINBERG, L. (1984). **A Response Model for Multiple Choice Itens**. Psychometrika, 49, 501-519.

- THISSEN, D. e WAINER, H. (1982). **Some Standard Errors in Iten Response Theory**. Psychometrika, 47, 397-412.

- THISSEN, D. (1991). **MULTILOG Version 6 User's Guide**. Mooresville, IN: Scientific Software, Inc.

- TUKER, L. R. (1951). **Academic Ability Test**. Research Memorandum. Educational Testing Service. Princeton, New Jersey.

- WINSGERSKI, M. S. e LORD, F. M. (1973). **A Computer Program for Estimating Examinee Ability and Item Characteristic Curve Parameters**. (RM-76-6) [computer program]. Educational Testing Service, Princeton, New Jersey.

- WOOD, R. L., WINGERSKI, M. S., e LORD, F. M. (1976). **LOGISTIC: A computer program for estimating examinee ability and item characteristic curve parameters** (RM-76-6). Princeton, NJ: Educational Testing Service.

- WRIGHT, B. D. (1977). **Solving Measurement Problems with the Rash Model**. Journal of Educational Measurement, 14, 97-116.

- WRIGHT, B. D., MEAD, R. J. E BELL, S. R. (1979). **BICAL: a Rasch Program for the Analysis of Dichotomus Data**, Chicago, MESA.

- WRIGHT, B. D. e STONE, M. H. (1979). **Best Test Design**. Chicago: MESA.

- YING SO (1997). **A Tutorial on Logistic Regression.** SAS Institute Inc., Cary, NC.