3
Teoria de Resposta ao Item Não Paramétrica

Este capítulo apresenta, com maiores detalhes, a Teoria de Resposta ao Item Não Paramétrica (TRIN) como uma das principais técnicas estatísticas utilizadas para a mensuração de variáveis latentes.

3.1 Abordagem clássica da Teoria de Resposta ao Item não Paramétrica

Como foi mencionada anteriormente, a teoria das variáveis latentes apresenta vários modelos que permitem a mensuração estatística de um construto a partir de itens de um questionário. Neste cenário, a Teoria de Resposta ao Item Não Paramétrica pode ser uma opção bastante flexível e sólida (Mokken, 1971; Sijtsma e Molenaar, 2002) para construção de escalas homogêneas em diversas áreas do conhecimento, particularmente, sob as seguintes condições:

- Os itens são medidos como variáveis dicotômicas.
- Existe a dificuldade do ajuste de modelos paramétricos.
- O conjunto de itens é pequeno (van Schuur, 2003).
- É usada para fins exploratórios. Torna-se uma ferramenta eficiente e rápida para avaliar a qualidade dos itens que posteriormente podem ser empregados na composição de um instrumento medida (pré-teste) ou em outros procedimentos estatísticos, por exemplo, o ajuste de modelos unidimensionais da TRI.
- É aproveitada para investigar a dimensionalidade da estrutura latente de um conjunto de itens.
- Não requer o traço latente apresente distribuição normal na população.
- A escalonabilidade do conjunto de itens (coeficiente H) e o poder de discriminação de um item (coeficiente H_i) não dependem do escore total.

A Teoria da Resposta ao Item não Paramétrica (TRIN) define a Curva Característica do Item (CCI) como a probabilidade de responder positivamente ao item, dado o valor do traço latente β do respondente. Considerando apenas itens de respostas dicotômicas (0: não e 1: sim), pode-se definir formalmente a CCI do item i, de um instrumento de medida com J itens, como:

$$\pi_i(\beta) = \Pr(U_i = 1 \mid \beta),\ i=1,\ 2,\ldots,\ J. \tag{3.1}$$

Ou seja, dado que o indivíduo tem traço latente β, a probabilidade de acertar o item i é igual a $\pi_i(\beta)$. A TRIN não impõe uma expressão algébrica definida para $\pi_i(\beta)$, exige apenas que $\pi_i(\beta)$ seja descrita por uma função monótona não decrescente em β, ou seja:

$$\beta_a < \beta_b \Rightarrow \pi_i(\beta_a) \le \pi_i(\beta_b). \tag{3.2}$$

Desta forma, a curva CCI que satisfaz esta relação de ordem em (3.2) pode não ter uma forma simétrica, mas apresentar intervalos onde a função seja linear (Sijtsma e Molenaar, 2002).

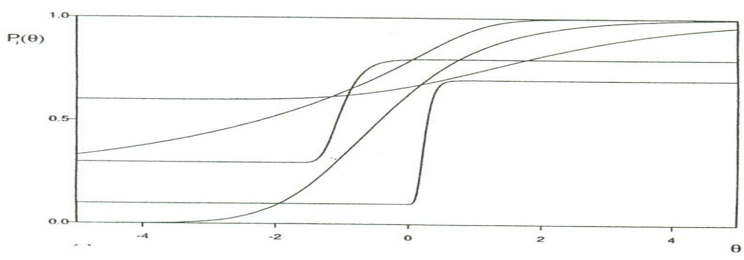

Figura 3.1: Comportamento monótono não decrescente de algumas curvas CCI
Fonte: Sijtsma e Molenaar, 2002.

Uma escala perfeita, chamada *escala de Guttman* (Mokken, 1971; Sijtsma e Molenaar, 2002), é aquela em que a posição do indivíduo está determinada com exatidão, entre dois itens. Tais escalas são raras, mas podem ocorrer, por exemplo, em uma escala de "mobilidade física" onde os itens apresentam popularidades[1] bem distintas. É o caso da escala formada por itens como, "consigo ir ao toalete sozinho", "consigo fazer compras sozinho" e "consigo correr dois quilômetros sozinho"; quem responde sim ao último item, certamente responderá sim aos anteriores.

Os modelos de *Mokken* (Mokken, 1971) da TRIN são de dois tipos: modelo de Homogeneidade Monótona (HM) e o modelo de Dupla Monotonocidade (DM).

[1] É a proporção de indivíduos que respondem positivamente a um item dicotômico.

Estes modelos não impõem uma forma paramétrica para a curva característica do item (CCI), exigindo apenas restrições qualitativas, como: para itens dicotômicos a curva CCI deve ser monótona não decrescente em relação ao traço latente. Tais modelos são, portanto, mais fracos que os modelos paramétricos da Teoria de Resposta ao Item, mas, por outro lado, são capazes de se ajustarem melhor a uma variedade maior de situações, por exemplo, quando o conjunto de itens é pequeno (van Schuur, 2003).

A qualidade do ajuste do modelo HM depende do grau com que os dados satisfazem às hipóteses deste modelo, a saber: *unidimensionalidade*, *independência local* e *monotonicidade* da CCI.

3.1.1 Os coeficientes de escalonabilidade

Mokken (1971) apresentou três coeficientes de escalonabilidade (homogeneidade) com o objetivo de determinar o ajuste do modelo HM aos dados, ou seja, a construção de escalas de Mokken com boas propriedades psicométricas.

Estes coeficientes estão baseados na hipótese de independência marginal. O primeiro deles, H_{ij}, mede o grau de associação entre cada par de itens i e j. O segundo, H_i, mede a homogeneidade de um item i particular referente a um conjunto de itens. O terceiro, denominado de H de Loevinger (Loevinger, 1948), mede a homogeneidade da escala como um todo.

3.1.1.1 O coeficiente de escalonabilidade entre pares de itens - H_{ij}

Sejam os itens i e j num questionário tais que: $i = 1, 2, \ldots, J$ e $j = 2, \ldots, J$. Na literatura existem várias medidas estatísticas para avaliar a dependência entre duas variáveis aleatórias, por exemplo, o coeficiente de correlação de Pearson. Entretanto, no contexto da TRIN, Mokken (1971) propôs o coeficiente de escalonabilidade (homogeneidade) de um par de itens dicotômicos i e j, denotado por H_{ij}, como uma medida de associação entre eles.

Para cada item i, definimos uma variável aleatória U_i que assume apenas os valores: 1, se ocorrer “acerto” ou 0, se ocorrer “erro” na resposta ao item i. Esta variável tem distribuição de Bernoulli com probabilidade de acerto ao item i igual a P_i. De modo análogo, definimos a variável aleatória U_{ij} que assume o valor 1 quando ocorre a resposta correta (acerto) aos itens *i* e *j*, e o valor 0 em caso contrário. P_{ij} denota a probabilidade de acertos simultâneos aos itens *i* e *j*.

A probabilidade atribuída ao evento "acertar o item *i*" pode ser definida como a freqüência relativa de ocorrência (proporção) de acertos ao item *i* (James, 2006). Desta forma, a *popularidade* do item *i* é estabelecida como a proporção de acertos do item *i* e a dificuldade do item *i* é dada por $1 - P_i$.

A relação de ordem $P_i < P_j$ indica que um item *i* é mais difícil que o item *j*. Quando isto ocorrer, a notação "$i < j$" será usada para indicar esta ordenação.

O coeficiente de escalonabilidade (H_{ij}) referente aos itens *i* e *j* é definido pela seguinte expressão:

$$H_{ij} = \frac{P_{ij} - P_i P_j}{P_i - P_i P_j}, \quad i < j. \tag{3.3}$$

Os valores do coeficiente H_{ij} geralmente estão compreendidos no intervalo [0; 1], mas podem ocorrer valores negativos. Segundo Sijtsma e Molenaar (2002), o par de itens (*i, j*) que gera esta situação de valores negativos para H_{ij} deve ser removido da análise.

Sob a abordagem clássica da TRIN, a estimação pontual do coeficiente H_{ij}, considera uma amostra aleatória simples *s* de tamanho *n* na qual são realizadas as operações matemáticas com as freqüências observadas entre as categorias das variáveis U_i e U_j (Quadro 1). Com essa amostra, é possível definir os seguintes estimadores: $\hat{P}_i = \frac{n_{11} + n_{10}}{n}$, $\hat{P}_j = \frac{n_{01} + n_{11}}{n}$ e $\hat{P}_{ij} = \frac{n_{11}}{n}$ para as proporções de acertos P_{i}, P_{j} e P_{ij}, respectivamente.

Assim, a expressão do estimador do coeficiente H_{ij} ($\hat{H}_{ij}$) é dada por:

$$\hat{H}_{ij} = \frac{\hat{P}_{ij} - \hat{P}_i \hat{P}_j}{\hat{P}_i - \hat{P}_i \hat{P}_j}, \quad i < j. \tag{3.4}$$

Quadro 3: Distribuição de freqüências amostrais observadas das respostas aos itens dicotômicos *i* e *j*

	U_j =1	U_j=0
U_i =1	n_{11}	n_{10}
U_i =0	n_{01}	n_{00}

onde n_{ab} representa a freqüência amostral de resposta aos itens *i* e *j*.

3.1.1.2 O coeficiente de escalonabilidade de um item - H_i

Este coeficiente mede a homogeneidade de um item i particular referente a um conjunto de J itens dicotômicos.

Para o item $i = 1, 2, ..., J$, o coeficiente H_i é definido como:

$$H_i = \frac{\sum_{j \neq i}^{J} (P_{ij} - P_i P_j)}{\sum_{i<j}^{J-1} P_i + \sum_{j<i}^{J} P_j - \sum_{j \neq i}^{J} P_i P_j}, \quad 0 \leq H_i \leq 1 \ . \tag{3.5}$$

Sijtsma e Molenaar (2002) afirmam que itens com *coeficiente de escalonabilidade* (H_i) compreendidos no intervalo [0; 0.30] possuem baixo *poder de discriminação* e devem ser removidos do processo de construção da escala.

3.1.1.3 O coeficiente de escalonabilidade da Escala - H

O terceiro coeficiente de escalonabilidade, denominado de H de Loevinger (Loevinger, 1948), mede a homogeneidade da escala como um todo e pode ser definido pela seguinte expressão:

$$H = \frac{\sum_{i} \sum_{j \neq i}^{J} P_{ij} - \sum_{i} \sum_{j \neq i}^{J} P_i P_j}{\sum_{i;i<j}^{J-1} P_i - \sum_{i<j}^{J-1} (P_i P_j) + \sum_{j,j<i} P_j - \sum_{j<i}^{J} (P_i P_j)}, \quad 0 \leq H \leq 1 \,. \tag{3.6}$$

Mokken (1971) e Sijtsma e Molenaar (2002) apresentam os níveis aceitáveis do coeficiente H para aferir a qualidade de uma escala. Este coeficiente tem estreita relação com a qualidade da ordenação estocástica dos respondentes com relação ao seu traço latente baseado no escore total dos itens selecionados. Segundo os autores citados, se $0.30 \leq H < 0.40$, o conjunto de itens forma uma escala fraca; se o valor de H pertence ao intervalo de 0.40 a 0.50 indica uma escala razoável; para valores de H acima de 0.50 indicam que os itens formam uma escala forte.

A *Escala de Mokken* é uma aplicação prática do ajuste do modelo HM a um conjunto de J itens dicotômicos que satisfazem as seguintes condições: $H_{ij} > 0$, $\forall i \forall j$; $H_i > c$ $\forall i$ e $H > c$.

A constante c é atribuída pelo pesquisador, geralmente com valor inicial em 0.30, com intuito de considerar itens com maior poder de discriminação.

A *escala de Guttman* tem um coeficiente de escalonabilidade H igual a um. À medida que a *escala de Mokken* se afasta do padrão de *Guttman*, a escalonabilidade H diminui. Um valor de H igual a 0, indica que os itens não formam uma escala.

A partir da escala final, com o coeficiente H em nível aceitável, é estabelecida uma nova variável que possibilita o ordenamento estocástico confiável dos respondentes. Esta medida, denominada de *escore total*, é construída pela soma dos valores atribuídos às respostas dos itens que compõem esta escala. Melhor dizendo, para um indivíduo respondente k, o respectivo escore total S_k pode ser definido com a soma das respostas aos J itens dicotômicos. Além disso, o escore total desse indivíduo varia entre 0 e J.

3.1.2 Ordenação inversa

A ordenação das *popularidades* dos itens na população pode ser invertida pela seleção de amostras. Quando este fenômeno ocorre em pelo menos um par de itens (i, j) de alguma amostra s, Sijtsma e Molenaar (2002) afirmam que ocorreu uma *ordenação inversa* entre estes itens. Isto pode acontecer quando a distância entre as *popularidades* dos itens i e j é inferior a 0.02. Além disso, estes autores recomendam como uma medida preventiva de possível inversão considerar amostras [aleatórias simples sem reposição] com tamanhos superiores a 400 respondentes. Para planos amostrais complexos, talvez seja necessário considerar tamanhos maiores de amostra.

Quando a ordenação das popularidades dos itens é conhecida em cada população de referência (ver Capítulo 6), é possível através dos estudos de simulação utilizados nesta tese, observar o fenômeno da ordenação inversa. Os resultados dessa análise serão apresentados nos Capítulos 7 e 8.

3.1.3 Limitações da TRIN

A Teoria de Resposta ao Item Não Paramétrica apresenta algumas limitações, a saber:

- O instrumento de medida (teste, questionário, etc.) elaborado sob o enfoque desta teoria é dependente dos itens que os compõem. Deste modo, a validade do instrumento de medida é prejudicada ou limitada.
- Não é possível posicionar o item na escala obtida pelo escore total.
- Quando os itens i e j apresentam a mesma proporção de acertos, o valor do coeficiente de escalonabilidade H_{ij} é indeterminado.
- Os parâmetros descritivos dos itens (dificuldade, popularidade e poder de discriminação) e do teste (coeficiente H e escore total) são afetados pelo grupo de sujeitos examinados.
- Pouco empregada na literatura para a mensuração de variáveis latentes em comparação com a Teoria de Resposta ao Item e a Teoria Clássica dos Testes.