

# CONSIDERAÇÕES SOBRE O ÍNDICE DE DISCRIMINAÇÃO DE ITENS EM TESTES EDUCACIONAIS

Fernando Lang da Silveira\*

## 1.0 – O quê é o índice de discriminação de um item?

Define-se o índice de discriminação de um item em um teste como a diferença entre as médias do item calculadas no grupo superior e no grupo inferior. O grupo superior é constituído usualmente por 27% dos indivíduos que responderam ao teste: aqueles que obtiveram os maiores escores totais. O grupo inferior também é usualmente constituído por 27% dos indivíduos: aqueles que obtiveram os menores escores totais. Em testes com itens que geram uma variável dicotômica, com valores possíveis zero e um (por exemplo, itens de escolha múltipla em testes de rendimento escolar), o índice de discriminação poderá assumir valor desde  $-1,00$  até  $+1,00$ . Caso o item gere uma variável multinomial com valor máximo  $M$ , o índice de discriminação poderá assumir valor desde  $-M$  até  $+M$ .

O índice de discriminação de um item é uma medida de consistência do item com todo o teste (Silveira, 1980). Um alto índice de discriminação indica que os escores gerados pelo item são coerentes com os escores totais no teste. Um índice de discriminação nulo evidencia que os escores gerados pelo item são incoerentes com os escores totais no teste. Quanto mais homogêneo for um teste, tanto maiores serão os índices de discriminação de seus itens.

## 2.0 – É o índice de discriminação a melhor medida da consistência do item com o escore total?

Antes de responder a essa questão, notemos algumas propriedades do índice de discriminação:

1. O índice de discriminação depende da extensão dos grupos extremos. Se a proporção de

---

\* Instituto de Física da UFRGS; Faculdade de Educação e Instituto de Física da PUCRS.

indivíduos nos grupos extremos for diminuída (por exemplo, ao invés de usar 27% dos indivíduos nos grupos extremos, se usar 15%), o índice de discriminação tenderá a aumentar.

2. O índice de discriminação depende do escore máximo no item. Já notamos que o índice de discriminação pode variar entre  $-M$  e  $+M$  ( $M$  é o escore máximo do item). Isso implica o fato de que dois itens no mesmo teste podem ter índices de discriminação diferentes simplesmente porque os escores máximos são diferentes. Mesmo quando os escores máximos são idênticos, não são necessariamente comparáveis os índices de discriminação conforme veremos na próxima propriedade.

3. O índice de discriminação depende do desvio padrão dos escores no item (Silveira, 1980). Se dois itens forem igualmente consistentes com o escore total, mas possuírem diferentes desvios padrão, apresentarão diferentes índices de discriminação; o que tiver o maior desvio-padrão, apresentará o maior índice de discriminação. Ebel (1977) nota que o índice de discriminação beneficia os itens de dificuldade mediana (itens com maior desvio padrão) em testes de aproveitamento escolar; tal fato, antes de ser uma vantagem, é uma restrição ao uso do índice de discriminação como medida da consistência do item com o escore total, pois poderá determinar que entre dois itens julgue erroneamente qual seja o melhor (o mais consistente com o escore total).

Realmente, o índice de discriminação não é a melhor medida da consistência do item com o escore total. Conforme demonstrou Richardson (1936), o coeficiente de fidedignidade de um teste depende dos coeficientes de correlação item-total; portanto, o coeficiente de correlação item-total é a medida mais apropriada da consistência do item com o escore total. Tal posição é apoiada por Nunnally (1967).

Flanagan (1939) preparou uma tábua que permite estimar o coeficiente de correlação item-total a partir das médias do item nos grupos extremos; entretanto, tal tábua somente pode ser utilizada para itens que geram uma variável dicotômica. Silveira (1980) deduz a relação que existe entre o coeficiente de correlação item-total e o índice de discriminação; essa relação é válida mesmo quando o item gera uma variável multinomial. A relação é a seguinte:

$$r_{iT} = \frac{D_i}{K_{SI}s_i} \quad (\text{equação 1})$$

$r_{iT}$  — coeficiente de correlação item-total.

$D_i$  — índice de discriminação do item.

$s_i$  — desvio padrão dos escores do item.

$K_{SI}$  — constante que depende da distribuição do escore total e da extensão dos grupos extremos.

Silveira (1980) apresenta os valores  $K_{SI}$  para diversas extensões dos grupos extremos, quando o escore total possui distribuição normal. Para os usuais 27% nos grupos extremos a constante vale 2,45.

Pode-se utilizar a equação 1 para calcular o coeficiente de correlação item-total e, assim, evitar as possíveis restrições ao uso do índice de discriminação. É interessante notar que não há necessidade de tomar a mesma proporção de indivíduos nos grupos superior e inferior para calcular o índice de discriminação. A constante  $K_{SI}$  depende dessas proporções. Sugerimos ao leitor consultar o trabalho "Relações entre Duas Estatísticas Utilizadas em Análise de Consistência Interna de Testes Psicométricos" (Silveira, 1980).

### 3.0 — Qual a relação entre os índices de discriminação dos itens e coeficientes de fidedignidade (coeficiente alfa de Cronbach) do escore total?

Esta relação foi deduzida por Silveira (1981) e é apresentada a seguir.

$$\alpha = \frac{k}{k-1} \left[ 1 - \frac{k_{SI} \sum s_i^2}{\left( \sum D_i \right)^2} \right] \quad (\text{equação 2})$$

- $\alpha$  – coeficiente alfa de Cronbach.  
 $k$  – número de itens no teste.

O coeficiente alfa de Cronbach reduz-se à fórmula KR-20 quando os itens geram variáveis dicotômicas com valores possíveis 0 e 1. Isto é, a fórmula KR-20 é um caso particular de coeficiente alfa de Cronbach, sendo o último empregável mesmo quando os itens do teste geram variáveis multinomiais.

Pela equação 2 nota-se que quanto maior forem os índices de discriminação dos itens, maior será o coeficiente de fidedignidade.

#### 4.0 – Qual o valor mínimo aceitável para o índice de discriminação de um item em um teste?

Silveira (1981) demonstra que esta pergunta não possui uma única resposta, ou melhor, o problema do valor mínimo aceitável para o índice de discriminação de um item está relacionado a uma série de aspectos. Ele demonstrou que o valor mínimo depende:

- do coeficiente de fidedignidade desejado para o escore total do teste,
- do número de itens no teste,
- do desvio padrão dos itens,
- da extensão dos grupos extremos.

Como ilustração, apresentamos qual deve ser o *valor médio* dos índices de discriminação dos itens em testes com coeficiente de fidedignidade igual a 0,70 (primeira coluna) e 0,90 (segunda coluna), variando o número de itens. Esses valores referem-se a itens que geram uma variável dicotômica e a grupos extremos que perfazem 27% do total de indivíduos cada um; além disso, os itens têm dificuldade mediana, ou seja, apresentam índice de dificuldade de 0,50. Os valores foram calculados através de uma equação deduzida por Silveira (1981).

Apresentamos, ainda, os valores médios dos índices de discriminação para itens com dificuldade de 0,80 ou 0,20 (quando o índice de dificuldade se afasta de 0,50 o desvio padrão do item diminui). Os coeficientes de fidedignidade continuam os mesmos que os anteriores.

Nº de itens	$r_{XX} = 0,70$	$r_{XX} = 0,90$	No de itens	$D = 0,80$	$D = 0,20$
5	0,83	1,00	5	0,66	0,80
10	0,64	0,89	10	0,51	0,71
15	0,54	0,79	15	0,43	0,63
20	0,47	0,72	20	0,38	0,58
25	0,43	0,66	25	0,34	0,53
30	0,39	0,62	30	0,31	0,50
50	0,31	0,50	50	0,25	0,40
100	0,22	0,37	100	0,18	0,30
200	0,16	0,27	200	0,13	0,22

É importante notar que quando o índice de dificuldade se afasta de 0,50 (isso significa que o desvio padrão do item diminui), o índice de discriminação será menor para se obter o mesmo coeficiente de fidedignidade. Esse aspecto enfatiza a não-comparabilidade de índices de discriminação em itens com diferentes desvios padrão.

Os valores relacionados anteriormente são valores médios para os índices de discriminação. Isto significa que em um teste pode haver itens com índice de discriminação menor do que aqueles apresentados, se existirem itens com índice de discriminação maior.

### 5.0 – É o índice de discriminação uma média de consistência do item com os demais itens?

Como o índice de discriminação é uma medida da consistência do item com o escore total no teste e como o escore total depende de todos os itens, o índice de discriminação é uma medida da consistência do item com os demais. Entretanto, o escore total também depende do próprio item em questão e por isso ele é uma medida espuriamente alta da consistência com os demais itens (Silveira, 1982).

É possível fazer-se uma correção no índice de discriminação para que ele se torne uma medida da consistência do item com os demais itens, apenas. A equação que permite fazer tal correção foi deduzida por Silveira (1982) e apresentamos a seguir.

$$D_i' = \frac{D_i s_T - k_{SI} s_i^2}{\sqrt{s_i^2 + s_T^2 - \frac{2s_T D_i}{k_{SI}}}} \quad (\text{equação 3})$$

$D_i'$  – índice de discriminação corrigido.

$s_T$  – desvio padrão do escore total no teste.

Conforme Silveira (1982), a correção merece ser feita quando o teste é constituído por poucos itens, quando o índice de discriminação é pequeno e quando os desvios padrão são muito diferentes.

Quando o teste é constituído por muitos itens, a correção produzirá um efeito irrelevante, principalmente se os índices de discriminação forem grandes. Caso os desvios padrão dos itens sejam iguais, a correção não alterará a ordenação dos itens segundo o índice de discriminação; ou seja, aquele que apresentar o maior índice de discriminação antes da correção, também apresentará o maior após a correção e assim por diante.

### 6.0 – Considerações finais

Apesar das restrições aqui levantadas contra o índice de discriminação, ele pode ser utilizado em análise de consistência interna de testes desde que se tome a devida cautela. O índice de discriminação possui a vantagem de ser facilmente calculado e o seu significado entendido por pessoas que não conhecem estatística. Lembremos, novamente, que, através da equação 1, pode-se, sem muito trabalho, encontrar a estatística mais apropriada à análise de consistência interna: o coeficiente de correlação item-total. O desvio padrão do item que gera uma variável dicotômica é dado pela seguinte equação:

$$s_i = \sqrt{p_i (1 - p_i)} \quad (\text{equação 4})$$

$p_i$  – índice de dificuldade do item.

Caso o item gere uma variável multinomial, o desvio padrão é um pouco mais trabalhoso de ser calculado. Sugerimos ao leitor consultar qualquer texto de estatística, em caso de dúvida. Entretanto, ainda é mais fácil obter o coeficiente de correlação item-total a partir do índice de discriminação do que através das usuais equações para o cálculo do coeficiente de correlação de Pearson.

## BIBLIOGRAFIA

- EBEL, R.L. (1977) – *Fundamentos de la Medición Educacional* – Editorial Guadalupe – Buenos Aires – 710 pág.
- FLANAGAN, J. C. (1939) – General Considerations in the Selection of Test Item and a Short Method of Estimating the Product Moment Coefficient from Data at the Tails of the Distribution – *Journal of Educational Psychology* 30, 674-680.
- NUNNALLY, J. C. (1967) – *Psychometric Theory* – McGraw-Hill Inc. – New York – 640 pág.
- RICHARDSON, M. W. (1936) – Notes on the Rationale of Item Analysis – *Psychometrika*, 1, 69-79.
- SILVEIRA, F. L. (1980) – Relação entre Duas Estatísticas Utilizadas em Análise de Consistência Interna de Testes Psicométricos – *Ciência e Cultura*, 32 (2), 214-216.
- SILVEIRA, F. L. (1981) – Relação entre os Índices de Discriminação de Itens em Testes Psicométricos e Duas Outras Estatísticas: Variância do Escore Total e Coeficiente de Fidedignidade – *Ciência e Cultura*, 33 (2), 246-248).
- SILVEIRA, F. L. (1982) – Correção no Índice de Discriminação de Itens em Testes Psicométricos – Trabalho ainda não publicado; aceito para publicação em *Ciência e Cultura*.