

FIDEDIGNIDADE, UM ENFOQUE ESTATÍSTICO

Maria Emília Camargo

Departamento de Estatística. Centro de Ciências Naturais e Exatas. UFSM. Santa Maria, RS.

José Elias Rigueira

Departamento de Educação Física. Centro de Ciências Biológicas e da Saúde. UFV. Viçosa, MG.

RESUMO

Define-se fidedignidade em termos de erro de mensuração. Descreve-se a teoria da fidedignidade, delineando-se o modelo de segregação da variância total em variância verdadeira e variância do erro ($S_0^2 = S_v^2 + S_e^2$).

Descreve-se o método das duas metades (Spearman-Brown), aplicado para determinar a fidedignidade do instrumento Deslizador Manual utilizado para o controle temporal na área de Educação Física.

SUMMARY

CAMARGO, M.E., and RIGUEIRA, J.E., 1982. Statistical focus on reliability. *Ciência e Natura* (4):27-31.

Reliability is defined as error of measurement. Reliability theory is described and a model is proposed that divides the total variance in two parts: the variance due to error and the real variance.

The method of split-half is described and to determine the reliability of the instrument "Deslizador Manual" used for time control in the area of Physical Education.

INTRODUÇÃO

O desenvolvimento de qualquer atividade científica depende da perfeição dos seus instrumentos de medida (8). Pois, desconhecendo-se a precisão com que está sendo mensurada determinada variável, não se terá segurança em apresentar qualquer conclusão a respeito dos resultados alcançados (5). Fidedignidade é, portanto, componente fundamental a ser considerado, em qualquer instrumento de medida, por um pesquisador que deseje desenvolver um trabalho científico (2,6). Assim, através de técnicas estatísticas pode-se estimar a fidedignidade de um instrumento (3).

DESENVOLVIMENTO

Fundamentação teórica

Toda e qualquer mensuração está sujeita a um erro de medida, assim pode-se dizer que o escore obtido (X_0) num teste é consti

tuído de dois elementos: o escore verdadeiro do indivíduo (X_v) e o escore correspondente ao erro (X_e)

$$X_0 = X_v + X_e \quad [1]$$

Os erros existentes nos escores de um teste, e que afetam sua consistência, são predominantemente erros acidentais. Sua característica principal é que, se várias mensurações são feitas, os erros tendem a cancelar-se e assim, o escore obtido se aproxima do verdadeiro. Mas como normalmente, as medidas que se têm de indivíduos são poucas e a influência do erro, quando não controlada, pode afetar, sensivelmente, o grau de fidedignidade dos escores.

A fidedignidade de um teste (r_{xx}) refere-se à estabilidade dos seus resultados, isto é, quanto maior o erro, menos fidedigno é o instrumento. O importante é minimizar a influência do erro para que o escore obtido (X_0) se aproxime, cada vez mais, do escore verdadeiro (X_v). Isso nem sempre é possível na prática, daí a necessidade de estimar o grau de fidedignidade do teste para saber a confiança que os escores merecem.

Assim, como se decompõe o escore obtido [1] nos seus elementos, pode-se, também, admitindo-se que o escore verdadeiro (X_v) e o escore relativo ao erro (X_e) seja nulo, decompor a sua variância:

$$S_0^2 = S_v^2 + S_e^2 \quad [2]$$

Logo a variância do escore verdadeiro (S_v^2) é:

$$S_v^2 = S_0^2 - S_e^2 \quad [3]$$

A fidedignidade é a relação entre a variância do escore verdadeiro e a do escore obtido, ou seja:

$$r_{xx} = \frac{S_v^2}{S_0^2} \quad [4]$$

Substituindo-se o valor S_v^2 na equação [4], tem-se:

$$r_{xx} = \frac{S_0^2 - S_e^2}{S_0^2} \quad [5]$$

Se o teste for perfeitamente fidedigno, isto é, $S_e^2 = 0$, não haverá erro nos escores, dessa forma:

$$r_{xx} = \frac{S_0^2}{S_0^2} = 1 \quad [6]$$

representando assim uma fidedignidade perfeita para um instrumento igualmente perfeito. Dividindo-se os elementos da equação [2] pela variância dos escores obtidos (S_0^2) temos:

$$\frac{S_0^2}{S_0^2} = \frac{S_v^2}{S_0^2} + \frac{S_e^2}{S_0^2} \quad [7]$$

$$1 = \frac{S_v^2}{S_0^2} + \frac{S_e^2}{S_0^2} \quad [8]$$

$$\frac{S_v^2}{S_0^2} = 1 - \frac{S_e^2}{S_0^2} \quad [9]$$

$$r_{xx} = 1 - \frac{S_e^2}{S_0^2} \quad [10]$$

Assim, do ponto de vista estatístico, a fidedignidade \bar{e} é igual a 1 menos a parte da variância total que \bar{e} a variância do erro. Operacionalmente, a fidedignidade pode ser definida como sendo o coeficiente de correlação entre dois conjuntos de escores obtidos independentemente, em formas paralelas do teste, para um mesmo grupo (EBEL, 1).

Existem vários métodos para determinar a fidedignidade de um instrumento, tais como: Aplicações repetidas da mesma escala, Aplicações de formas paralelas da escala; Divisão da escala em duas metades; Método de Kuder-Richardson; Método de Horst-Richardson; Método de Rulan; Coeficiente de Hoyt; Coeficiente de Cromback; porém, neste trabalho será dado um enfoque ao método das duas metades.

Método das duas metades (Spearman-Brown)

Consiste em construir um teste e aplicá-lo somente uma vez a um grupo de indivíduos numa única sessão. Divide-se o teste em duas metades, correlacionando-se os resultados das duas partes para estimar o coeficiente de fidedignidade. As técnicas de correlação mais empregadas têm sido a linear de Pearson e a ordinal de Spearman. O coeficiente obtido traduz a correlação de um teste que foi reduzido à metade de sua extensão original. Em consequência, para estimar a fidedignidade do teste inteiro deve-se aplicar a fórmula de predição de Spearman-Brown. Essa fórmula permite calcular nova fidedignidade de um teste quando ele é aumentado (ou diminuído) certo número de vezes. Sua expressão geral é:

$$r_{nn} = \frac{nr_{xx}}{1 + (n-1)r_{xx}} \quad [11]$$

onde:

r_{nn} = fidedignidade para o teste aumentado n vêzes

n = número de vêzes que o teste é aumentado

r_{xx} = fidedignidade da metade do teste, calculada através da correlação de Pearson ou Spearman.

Assim, para se encontrar a fidedignidade do teste inteiro temos:

$$r_{2n} = \frac{2 \cdot r_{xx}}{1 + (2-1) \cdot r_{xx}} = \frac{2r_{xx}}{1 + r_{xx}} \quad [12]$$

Esse método só deve ser utilizado, quando as duas metades do teste possuem equivalência de mensuração.

Material

Os dados experimentais foram extraídos de RIGUEIRA (1982) onde vinte e quatro sujeitos foram submetidos a um teste de controle temporal, sendo utilizado um instrumento denominado Deslizador Manual.

O instrumento, composto de duas tábuas paralelas, estava acoplado a um cronômetro digital, que media em segundos o tempo gasto pelos sujeitos para efetuarem uma volta completa num circuito fechado, traçado na sua superfície.

O teste de controle temporal, consistia em realizar toda a volta proposta no circuito, obedecendo a uma duração estabelecida pelo pesquisador. O sujeito executava 12 tentativas, distribuídas aleatoriamente, sendo 4 em cada uma das 3 durações previstas para a realização do teste (10", 15" e 20").

RESULTADOS E DISCUSSÃO

Dos resultados obtidos no teste de controle temporal calculou-se o valor do erro absoluto, com relação a duração prevista, sendo então calculada a média das execuções pares e ímpares. (Tabela I).

TABELA I. MÉDIAS DOS ERROS ABSOLUTOS OBTIDOS PELOS SUJEITOS, CALCULADAS A PARTIR DAS EXECUÇÕES PARES E ÍMPARES.

Ordem de execução	Médias dos erros absolutos	
	Tentativas Pares	Tentativas Ímpares
1	1,5833	1,4167
2	2,0833	2,0833
3	3,6250	3,3333
4	4,7083	3,4583
5	6,3750	5,5417
6	7,5833	7,0833
\bar{X}	4,3264	3,8194

Foi então calculada a correlação linear de Pearson entre as execuções pares e ímpares para a qual foi encontrado 0,98, sendo que esta correlação representa a estimativa da fidedignidade para ca da metade do teste. Finalmente, aplicou-se a equação [12] com a qual foi encontrada uma fidedignidade para o teste inteiro de 0,99.

CONCLUSÃO

Pelo método das duas metades (Spearman-Brown), pode-se concluir que o instrumento Deslizador Manual apresenta uma fidedignidade de 0,99, considerada por MATHEWS (4), como excelente, pois 99% da variância nas execuções dependem da variância verdadeira e de apenas 1% da variância do erro.

REFERÊNCIAS BIBLIOGRÁFICAS

1. EBEL, R.L. *Measuring educational achievement*. Englewood Cliffs, N. I.: Prentice-Hall, Inc. 1965.
2. GARRETT, Henry E. *A estatística na psicologia e na educação*. Rio de Janeiro, Fundo de Cultura. 1962.
3. GULLFORD, I.P. *Fundamental statistics in psychology and education*. 2 ed. New York, McGraw-Hill. 1965.
4. MATHEWS, D.K. *Medida e avaliação em educação física*. Rio de Janeiro, Intêramericana. 1980.
5. NICK, Eva. *Estatística e psicometria*. Rio de Janeiro, I. Ozon Editor. 1963.
6. NUNNALLY, Jum C. *Educational measurement and evaluation*. 2 ed. New York, McGraw-Hill. 1972.
7. RIGUEIRA, J.E. Projeto de dissertação: *A influência do sexo e da lateralidade no controle de uma tarefa motora temporal*. Santa Maria. 1982.
8. WAINERMAN, C.H. *Escalas de medición en ciencias sociales*. Buenos Aires, Nueva Vision. 1976.

Recebido em novembro, 1982; aceito em novembro, 1982.

