PSYCHOMETRIKA—VOL. 43, NO. 4
DECEMBER, 1978

A SERIES OF LOWER BOUNDS TO THE RELIABILITY OF A TEST

JOS. M. F. TEN BERGE AND FRITS E. ZEGERS

UNIVERSITY OF GRONINGEN

Two well-known lower bounds to the reliability in classical test theory, Guttman's λ_2 and Cronbach's coefficient alpha, are shown to be terms of an infinite series of lower bounds. All terms of this series are equal to the reliability if and only if the test is composed of items which are essentially tau-equivalent. Some practical examples, comparing the first 7 terms of the series, are offered. It appears that the second term (λ_2) is generally worth-while computing as an improvement of the first term (alpha) whereas going beyond the second term is not worth the computational effort. Possibly an exception should be made for very short tests having widely spread absolute values of covariances between items. The relationship of the series and previous work on lower bound estimates for the reliability is briefly discussed.

Key words: internal consistency, homogeneity.

Coefficient alpha [Cronbach, 1951] is the most popular lower bound to the reliability of a test in classical test theory, based on a single test administration. The coefficient was originally derived by Kuder and Richardson [1937] as Formula 20. Guttman [1945] called it λ_3 and showed that it was a lower bound. In addition to λ_3, Guttman derived five other lower bounds. One of these, λ_2, is always at least as good as alpha. The formulas of alpha and λ_2 are

$$\alpha = \frac{n}{n-1} \frac{\sum_{i \neq j} \sigma_{ij}}{\sigma_X^2} \tag{1}$$

and

$$\lambda_2 = \frac{\sum_{i \neq j} \sigma_{ij} + \left(\frac{n}{n-1} \sum_{i \neq j} \sigma_{ij}^2\right)^{1/2}}{\sigma_X^2} \tag{2}$$

where σ_{ij} is the covariance between items i and j, σ_X^2 is the variance of the test, and n the number of items. The superiority of λ_2 over α can be expressed in the inequality

$$\alpha \leq \lambda_2 \leq \rho^2(X, T) \tag{3}$$

where $\rho^2(X, T)$ is the reliability of the test. Clearly, λ_2 owes its superiority to the fact that it uses the sum of squares of the covariances between items in addition to the sum of these covariances. In this paper it will be shown how λ_2 can be improved in turn by using the sum of the fourth powers of the item covariances, and so on. In fact a series of lower

The authors are obliged to Henk Camstra for providing a computer program that was used in this study.

Requests for reprints should be sent to Jos M. F. ten Berge, Department of Psychology, University of Groningen, Oude Boteringestraat 34, Groningen, The Netherlands.

bounds can be derived, satisfying

(4) $$\mu_0 \le \mu_1 \le \cdots \le \mu_r \le \cdots \le \rho^2(X, T).$$

The basic definition which generates (4) is

(5) $$\mu_r = \frac{1}{\sigma_X^2}(p_0 + (p_1 + (p_2 + \cdots (p_{r-1} + (p_r)^{1/2})^{1/2} \cdots)^{1/2})^{1/2}), \qquad r = 0, 1, 2, \cdots$$

where

$$p_h = \sum_{i \ne j} \sigma_{ij}^{(2^h)}, \qquad h = 0, 1, 2, \cdots r-1$$

and

$$p_h = \frac{n}{n-1} \sum_{i \ne j} \sigma_{ij}^{(2^h)}, \qquad h = r.$$

Formula (5) contains precisely r square roots. For $r = 0$ we have

(6) $$\mu_0 = \frac{1}{\sigma_X^2}(p_0) = \frac{n}{n-1} \frac{\sum_{i \ne j} \sigma_{ij}}{\sigma_X^2} = \alpha.$$

For $r = 1$ we have

(7) $$\mu_1 = \frac{1}{\sigma_X^2}(p_0 + (p_1)^{1/2}) = \frac{1}{\sigma_X^2}\left(\sum_{i \ne j} \sigma_{ij} + \left(\frac{n}{n-1} \sum_{i \ne j} \sigma_{ij}^2\right)^{1/2}\right) = \lambda_2.$$

For $r = 2$ we obtain

(8) $$\mu_2 = \frac{1}{\sigma_X^2}(p_0 + (p_1 + (p_2)^{1/2})^{1/2}) = \frac{1}{\sigma_X^2}\left(\sum_{i \ne j} \sigma_{ij} + \left(\sum_{i \ne j} \sigma_{ij}^2 + \left(\frac{n}{n-1} \sum_{i \ne j} \sigma_{ij}^4\right)^{1/2}\right)^{1/2}\right)$$

and so on.

Proof. Formula (4) will be proven in two steps. First it will be shown that

(9) $$\mu_r \le \mu_{r+1}, \qquad r = 0, 1, 2, \cdots$$

and next it will be proven that

(10) $$\mu_r \le \rho^2(X, T), \qquad r = 1, 2, 3, \cdots$$

The inequality (9) is equivalent to

(11) $$\frac{n}{n-1} \sum_{i \ne j} \sigma_{ij}^{(2^r)} \le \sum_{i \ne j} \sigma_{ij}^{(2^r)} + \left(\frac{n}{n-1} \sum_{i \ne j} \sigma_{ij}^{(2^{r+1})}\right)^{1/2}$$

or

(12) $$\left(\frac{1}{n(n-1)} \sum_{i \ne j} \sigma_{ij}^{(2^{r+1})}\right)^{1/2} \ge \frac{1}{n(n-1)} \sum_{i \ne j} \sigma_{ij}^{(2^r)}.$$

Since the mean square deviation of the $\sigma_{ij}^{(2^r)}$ is nonnegative we have

(13) $$\frac{1}{n(n-1)} \sum_{i \ne j} \sigma_{ij}^{(2^{r+1})} - \left(\sum_{i \ne j} \frac{\sigma_{ij}^{(2^r)}}{n(n-1)}\right)^2 \ge 0$$

or

(14) $$\frac{1}{n(n-1)} \sum_{i \ne j} \sigma_{ij}^{(2^{r+1})} \ge \left(\sum_{i \ne j} \frac{\sigma_{ij}^{(2^r)}}{n(n-1)}\right)^2.$$

For $r \geq 1$ both members of (12) are nonnegative, hence (14) is equivalent to (12). For $r = 0$ (14) implies (12). This completes the proof of (9). Equality in (9) holds for $r \geq 1$ if and only if the $\sigma_{ij}^{(2^r)}$ are equal, hence if and only if the $|\sigma_{ij}|$ are equal; for $r = 0$ equality holds if and only if the σ_{ij} are equal and nonnegative.

Finally we shall prove (10). From $(\sigma^q(T_i) - \sigma^q(T_j))^2 \geq 0$ and $\sigma^q(T_i)\sigma^q(T_j) \geq \sigma^q(T_i, T_j)$ we have

$$(15) \qquad \sigma^{2q}(T_i) + \sigma^{2q}(T_j) \geq 2\,\sigma^q(T_i)\,\sigma^q(T_j) \geq 2\,\sigma^q(T_i,T_j) = 2\sigma_{ij}^q\,, q = 1, 2, 4, 8 \cdots$$

where T_i is the true score of item i. Summing yields

$$(16) \qquad \sum_{i \neq j} (\sigma^{2q}(T_i) + \sigma^{2q}(T_j)) \geq 2 \sum_{i \neq j} \sigma_{ij}^q$$

or

$$(17) \qquad 2(n - 1) \sum_{i=1}^{n} \sigma^{2q}(T_i) \geq 2 \sum_{i \neq j} \sigma_{ij}^q.$$

Setting $q = 2^r$ and dividing by $2(n - 1)$ yields

$$(18) \qquad \sum_{i=1}^{n} \sigma^{(2^{r+1})}(T_i) \geq \frac{1}{n-1} \sum_{i \neq j} \sigma_{ij}^{(2^r)}, \qquad r = 0, 1, 2, \cdots$$

In addition, $\sigma^2(T)$, the true-score variance of the test, can be written as

$$(19\text{-}0) \qquad \sigma^2(T) = \sum_{i \neq j} \sigma_{ij} + \sum \sigma^2(T_i)$$

$$= \sum_{i \neq j} \sigma_{ij} + \left(\left(\sum \sigma^2(T_i)\right)^2\right)^{1/2}$$

$$(19\text{-}1) \qquad = \sum_{i \neq j} \sigma_{ij} + \left(\sum_{i \neq j} \sigma^2(T_i)\sigma^2(T_j) + \sum \sigma^4(T_i)\right)^{1/2}$$

$$= \sum_{i \neq j} \sigma_{ij} + \left(\sum_{i \neq j} \sigma^2(T_i)\sigma^2(T_j) + \left(\left(\sum \sigma^4(T_i)\right)^2\right)^{1/2}\right)^{1/2}$$

$$(19\text{-}2) \qquad = \sum_{i \neq j} \sigma_{ij} + \left(\sum_{i \neq j} \sigma^2(T_i)\sigma^2(T_j) + \left(\sum_{i \neq j} \sigma^4(T_i)\sigma^4(T_j) + \sum \sigma^8(T_i)\right)^{1/2}\right)^{1/2}$$

$$\vdots$$

$$(19\text{-}r) \qquad = \sum_{i \neq j} \sigma_{ij} + \left(\sum_{i \neq j} \sigma^2(T_i)\sigma^2(T_j) + \left(\sum_{i \neq j} \sigma^4(T_i)\sigma^4(T_j) + \cdots \right.\right.$$

$$\left.\left. \cdots + \left(\sum_{i \neq j} \sigma^{(2^r)}(T_i)\sigma^{(2^r)}(T_j) + \sum \sigma^{(2^{r+1})}(T_i)\right)^{1/2} \cdots\right)^{1/2}\right)^{1/2}$$

where (19-r) contains precisely r square roots. Applying (18) to $\sum \sigma^{(2^{r+1})}(T_i)$ in (19-r) and reducing the remaining terms except $\sum_{i=j} \sigma_{ij}$ by a part of (15):

$$(20) \qquad \sigma^q(T_i)\sigma^q(T_j) \geq \sigma^q(T_i, T_j) = \sigma_{ij}^q, \qquad q = 2, 4, 8, 16, \ldots,$$

we have, at least for $r > 0$,

$$(21) \qquad \sigma^2(T) \geq \sum_{i \neq j} \sigma_{ij} + \left(\sum_{i \neq j} \sigma_{ij}^2 + \left(\sum_{i \neq j} \sigma_{ij}^4 \right.\right.$$

$$\left.\left. + \cdots + \left(\sum_{i \neq j} \sigma_{ij}^{(2^r)} + \sum_{i \neq j} \frac{\sigma_{ij}^{(2^r)}}{n-1}\right)^{1/2} \cdots\right)^{1/2}\right)^{1/2}.$$

Dividing (21) by σ_X^2 completes the proof of (10). Equality in (10) holds if and only if (16) holds as an equality; that is, if and only if the items are essentially tau-equivalent [Lord & Novick, 1974]. Equality in (9) is necessary, but not sufficient for equality to hold in (10). For example, let T_1, T_2 and T_3 represent variables with $\sigma^2(T_i) = c_1$ and $\sigma(T_i, T_j) = c_2$ $(i \neq j)$ with $c_1 > c_2$. Then three variables X_1, X_2 and X_3 can be constructed with true scores $\theta_i = T_1 + T_2 + T_3 - T_i$, $i = 1, 2, 3$. For X_1, X_2 and X_3 we have (9) holding as an equality, but not (10). In addition, these variables can be used to show that μ_r need not converge to $\rho^2(X, T)$ as r tends to infinity.

It was shown above [see the proof of (9)] that the μ_r are strictly increasing if the $|\sigma_{ij}|$ are not all equal. Nevertheless, it may not be worthwhile computing μ_r for a large value of r. In order to investigate this, values of μ_0 to μ_6 were computed for eight college tests on various topics of a psychology curriculum. The results are in the rows 1 to 8 of Table 1.

Several conclusions can be drawn from these results. First, the increments $\mu_{r+1} - \mu_r$ appear to be rapidly converging to zero. Secondly, computing μ_1 rather than μ_0 seems to be rewarding, but going beyond μ_1 does not seem to pay off. Apparently the spread of the $|\sigma_{ij}|$ tends to be too small to produce large differences between μ_2 and μ_1. This conclusion, however, needs modification. Large differences between μ_2 and μ_1 do not require a large spread of the $|\sigma_{ij}|$ but also tests containing very few items. This can be seen if we consider a perfect Guttman-scale consisting of three items with p-values .9, .2, and .1. The scale has a μ_1 of .6498 and a μ_2 of .6747. As more items are added, the discrepancy between μ_2 and μ_1 decreases rapidly, even though the variance of the $|\sigma_{ij}|$ increases. This shows that having a small number of items is favourable to $\mu_2 - \mu_1$. Accordingly, the first three and first four

TABLE 1

Values of μ_0 to μ_6 for 10 sets of data.

data-sets	μ_0	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
1	.6627	.7184	.7248	.7260	.7262	.7263	.7263
2	.4743	.5037	.5043	.5044	.5044	.5044	.5044
3	.5805	.6152	.6160	.6161	.6161	.6161	.6161
4	.6033	.6790	.6808	.6809	.6809	.6809	.6809
5	.6843	.7029	.7031	.7031	.7031	.7031	.7031
6	.6716	.7151	.7157	.7158	.7158	.7158	.7158
7	.7656	.7853	.7858	.7858	.7858	.7858	.7858
8	.6551	.6961	.6966	.6967	.6967	.6967	.6967
9	.3971	.4469	.4659	.4711	.4724	.4728	.4728
10	.4359	.4602	.4673	.4692	.4697	.4699	.4699

items of the eight tests were examined separately, and for the test showing the best results for μ_2 (Test 7) the values of μ_0 to μ_6 are reported in row 9 (three items) and 10 (four items). Clearly, for very short tests having widely spread absolute values of σ_{ij}, it can be recommended to compute μ_2 or even μ_3. Such tests will rarely occur in actual practice. Therefore, it is in general sufficient to compute μ_1 (Guttman's λ_2) as a lower bound to the reliability of a test.

Relation to Previous Work

The series of lower bounds μ_r, $r = 0, 1, 2, \cdots$ includes Guttman's well-known lower bounds λ_2 (μ_1) and λ_3 (μ_0). However, these are not the only lower bound indices available. First, Guttman's coefficients λ_1, λ_4, λ_5 and λ_6 are not included in the series. Secondly, Jackson and Agunwamba [1977] developed three additional lower bounds. More importantly, Jackson and Agunwamba [1977] and Woodhouse and Jackson [1977] managed to solve the problem of finding the greatest lower bound. We shall first compare the several lower bound indices available and then discuss the greatest lower bound.

Woodhouse and Jackson [1977, p. 586, Table 4.3] gave a partial ordering of the lower bounds λ_1 to λ_6 [Guttman, 1945] and the newly developed indices λ_7, ρ_3 and ρ_4 [Jackson & Agunwamba, 1977]. It appears that max (λ_4, λ_5, λ_6, λ_7) is always at least as high as any one of the nine lower bounds considered. In particular, λ_7 is always at least as high as λ_2 (μ_1). However, μ_2 may exceed each of the nine lower bounds. For instance, the first data set of Table 4.2 [Woodhouse & Jackson, 1977, p. 586] has $\lambda_7 = .711$ and $\mu_2 = .720$. For the second data set of the same Table we have $\mu_2 = .923$ hence μ_2 exceeds λ_4, λ_5 and λ_6. These results prove that the indices μ_2, μ_3, $\cdots$ are not dominated by any one of the nine lower bound indices considered by Woodhouse and Jackson. However, the greatest lower bound (g.l.b.) for which Woodhouse and Jackson offered a search procedure is by definition always at least as high as any explicit lower bound estimate. All lower bound estimates are inferior to the g.l.b. provided that the population variance-covariance matrix of the items be known or can be estimated accurately from the sample data. The present authors can see one reason for still using the μ_r lower bound estimates, apart from the case where one does not have a rapid computer program for searching the g.l.b. available. The g.l.b. treats the covariances between items asymmetrically, thus possibly capitalizing on sampling error in the covariances between the items. The same holds for the lower bound indices λ_4, λ_5, λ_6 and λ_7. Until the sampling properties of the g.l.b. are known one may resort to the μ_r-coefficients for small-sized samples. Woodhouse and Jackson [1977, p. 591] report that simulation studies on the sampling properties of the g.l.b. are in progress. Thus it seems that more decisive recommendations will be made shortly.

REFERENCES

Cronbach, L. J. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951, *16*, 297–334.

Guttman, L. A basis for analyzing test-retest reliability. *Psychometrika* 1945, *10*, 255–282.

Jackson, P. H. & Agunwamba, C. C. Lower bounds for the reliability of the total score on a test composed of nonhomogeneous items: I. Algebraic lower bounds. *Psychometrika*, 1977, *42*, 567–578.

Kuder, G. F. & Richardson, M. W. The theory of the estimation of test reliability. *Psychometrika*, 1937, *2*, 151–160.

Lord, F. M. & Novick, M. R. *Statistical theories of mental test scores;* Reading, Mass.: Addison-Wesley, 1974.

Woodhouse, B. & Jackson, P. H. Lower bounds for the reliability of the total score on a test composed of nonhomogeneous items: II. A search procedure to locate the greatest lower bound. *Psychometrika*, 1977, *42*, 579–591.

Manuscript received 11/18/77
Final version received 5/8/78