Shiken: JALT Testing & Evaluation SIG Newsletter. 7 (1) March 2003 (p. 14 - 16)

Statistics Corner: **Questions and answers about language testing statistics:**

The coefficient of determination

James Dean Brown
University of Hawai'i at Manoa

QUESTION: In a recent paper on cloze tests (Brown, Yamashiro, & Ogane, 2001, p. 143), you mentioned that the coefficients of determination for cloze tests vary from .19 to .83. Can you explain what coefficients of determination are? How are they calculated?

ANSWER: The central issue underlying your question is: how can we interpret correlation coefficients? To answer that question, I must deal with three sub-questions: (1) What does a correlation coefficient mean? (2) What does a coefficient of determination mean? And, (3) how does the coefficient of determination help put the interpretation of correlation coefficients into perspective?

(1) *What does a correlation coefficient mean?*

Table 1. *Sample data for tests A, B, and C*

Test A	Test B	Test C
9	8	1
8	7	2
7	6	3
6	5	4
5	4	5
4	3	6
3	2	7
2	1	8
1	0	9

Correlation was once defined for me by one of my professors at UCLA as the "go-togetherness" of two sets of numbers. That definition has always made sense to me. The degree to which two sets of numbers go together can be calculated statistically as a *correlation coefficient*[1]. Correlation coefficients (often symbolized by r, or r_{xy}) can turn out to be as high in a positive direction as +1.00 if the relationship between the two sets of numbers is perfect and in the same direction (as is the case in Table 1 below for Tests A and B). A correlation coefficient can also turn out to be as high in a negative direction as -1.00 if the relationship is perfect and in opposite directions (as is the case in Table 1 for Tests A and C, or B and C). A correlation coefficient can also turn out to be zero if no relationship at all exists between the two sets of numbers (as would be the case if two sets of numbers were, say, random numbers). Basically, this is what a correlation coefficient represents. However, this coverage of the topic has necessarily been brief because the focus here is on the *coefficient of determination*. [For much more on calculating and interpreting correlation coefficients see Brown 1996, 1999, or any other good applied statistics or testing book.]

One problem that arises in interpreting correlation coefficients is that their relative magnitudes are not proportional. That is to say, a correlation coefficient of .80 cannot be said to be accounting for twice as much "go-togetherness" as a coefficient of .40. Thus it is difficult to interpret and understand correlation coefficients, especially relative to each other.

(2) *What does a coefficient of determination mean?*

The *coefficient of determination* makes interpreting correlation coefficients easier. Notwithstanding its impressive name, calculating this coefficient is simple. The coefficient of determination is simply the squared value of the correlation coefficient. That is why the symbol for this statistic is r_{xy}. The resulting coefficient of determination provides an estimate of the *proportion* of overlapping variance between two sets of numbers (i.e., the degree to which the two sets of numbers vary together).

By simply moving the decimal point two places to the right, you can interpret a coefficient of determination as the *percentage* of variance shared by the two sets of numbers. So, a coefficient of determination of .81 can be interpreted as a proportion, or as 81%. For example, if you have two sets of scores on Tests X and Y, and they correlate at .90, you could square that value to get the coefficient of determination of .81 and interpret that result as meaning that 81% of the variance in Test X is shared with Test Y, or for that matter, that 81% of the variance on Test Y is shared with Test X. By extension, you should recognize that you don't know what the remaining 19% (100% - 81% = 19%) on each test is related to.

(3) *How does the coefficient of determination put the interpretation of correlation coefficients into perspective?*

Table 2. *Sample correlation coefficients and corresponding coefficients of determination*

Correlation Coefficient (r_{xy})	Coefficient of Determination (r_{xy}^2)
1.00	1.00
0.90	0.81
0.80	0.64
0.70	0.49
0.60	0.36
0.50	0.25
0.40	0.16
0.30	0.09
0.20	0.04
0.10	0.01

Why bother calculating the coefficient of determination? Well, as explained in the previous section, it is worthwhile because the proportions (or percentages) represented by the coefficient of determination are easier for most people to understand and because you can truly say that the ratios represented by various values of this coefficient have meaning. In other words, a coefficient of determination of .80 can be said to represent twice as much overlapping variance between two sets of numbers as a coefficient of .40. Table 2. *Sample correlation coefficients*

In addition, Table 2 illustrates how sharply the coefficients of determination decline in magnitude when compared with their respective correlation coefficients. For example (as shown in Table 2), a correlation of:

.90 squared equals .81 (i.e., 81%) or about four-fifths overlap,
.80 squared equals .64 (i.e., 64%) or about two-thirds overlap,
.70 squared equals .49 (i.e., 49%) or about one-half overlap,
.60 squared equals .36 (i.e., 36%) or about one-third overlap,
.50 squared equals .25 (i.e., 25%) or about one-quarter overlap,

.40 squared equals .16 (i.e., 16%) or about one-fifth overlap,
.30 squared equals .09 (i.e., 9%) or about one-tenth overlap,
.20 squared equals .04 (i.e., 4%) or about one-twenty-fifth overlap (almost nothing)
.10 squared equals .01, or less than one-hundredth overlap (definitely nothing)

Notice how much more rapidly the coefficients of determination drop as you scan down Table 2 than do the correlation coefficients. This should help you recognize that correlation coefficients can be misleading. For example, I have seen a correlation coefficient of .60 called a "moderate" correlation. After all, .60 on a scale from .00 to 1.00 looks like it represents about three-fifths overlap. But when you square that value to find the coefficient of determination, you quickly realize that the proportion of relationship is .36, or 36%, which is only about one-third overlap. How can that be said to represent a moderate relationship? Even a "moderate" correlation of .70 is only .49 when squared, and thus represents less than one-half overlap between whatever two sets of numbers are involved. The bottom line is that the coefficient of determination transforms a correlation coefficient into a statistic that you can more readily interpret and compare to other coefficients.

Conclusion

Your original question asked about our statement (Brown, Yamashiro, & Ogane, 2001, p. 143) that the coefficients of determination for cloze tests have varied from .19 to .83. You should now understand that these values of .19 to .83 would result from squaring correlation coefficients ranging from .44 (.442 = .1936 or about .19) to .91 (.912 = .8281 or about .83) and that each of these coefficients of determination represent the proportion of overlap between two sets of numbers (in this case, between cloze test scores and some other measure of overall English language proficiency, e.g., the TOEFL).

For more on coefficients of determination, see Brown 1996, 1999, or any other good applied statistics or testing book.

References

Brown, J. D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall.

Brown, J. D. (translated into Japanese by M. Wada). (1999). *Gengo tesuto no kisochishiki* (Literally: Basic knowledge of language testing). Tokyo: Taishukan Shoten.

Brown, J. D., Yamashiro, A. D., & Ogane, E. (2001). The emperor's new cloze: Strategies for revising cloze tests. In T. Hudson & J. D. Brown (Eds.), *A focus on language test development: Expanding the language proficiency construct across a variety of tests*. Honolulu, HI: University of Hawai'i Press.

[1] Throughout this explanation, I am referring to the most commonly reported correlation coefficient, the Pearson product-moment correlation coefficient, which is appropriate when the two sets of numbers are on continuous scales. The coefficient of determination is only appropriate for use with the Pearson coefficient.