

Chapter 25

Derivation of Classical Test Theory Equations and Coefficient α



Formalization and Derivation of CTT Eqs. (3.1)–(3.5) in Chap. 3

In many traditional texts, the subscript for the person is taken for granted. However, though quicker to write, it can contribute to confusion when it is not clear if the summation is over persons or over items, or both. Therefore, in this book, we continue to use subscripts n and i for a person and an item, respectively. According to Eq. (3.1),

$$y_n = t_n + e_n, \tag{25.1}$$

where y_n is the observed score on a test, that is, a person's total score on a set of items, and t_n, e_n are, respectively, the person's true and error scores. This is the fundamental simple equation of CTT and note again that it has no item parameters.

Thus, we see that the observed score on the variable is the sum of two other variables, the true and error scores. Both are unobserved and therefore latent scores. In addition, they are real numbers, whereas y_n is typically an integer.

If the error is not correlated with the actual true score, then from what we have just learned about the variance of the sum of two variables, it follows that the variance of the observed scores is the sum of the variances of the true scores and the error scores. That is

$$s_y^2 = s_t^2 + s_e^2 \tag{25.2}$$

This is the second most relevant equation in CTT.

The next important concept developed in CTT is the formalization of reliability. It begins with the idea of two tests being administered to measure the same construct. These tests are often termed parallel tests. In some situations, there really are two tests, but we do not need them to develop a theory and then see what other ways we can calculate the reliability.

Each test will have its error of measurement, but the true score for a person will be the same. Here we have to use double subscripts briefly.

Let the score of person n on test 1 be y_{n1} and on test 2 be y_{n2} .

That is,

$$y_{n1} = t_{n1} + e_{n1} \quad (25.3a)$$

and

$$y_{n2} = t_{n2} + e_{n2}. \quad (25.3b)$$

Because we are interested in how consistent the observed scores are from the two tests, we calculate the correlation between y_{n1} and y_{n2} . We would want the correlation to be high. We begin with the calculation of the covariance between y_1 and y_2 .

Derivation of Covariance

We have not derived these relationships in full elsewhere, and therefore for completeness include them here. We could use the random variable notation but for simplicity, we use the chosen scores in the derivations. In calculating the covariance, we immediately estimate the population value and therefore use $N - 1$ rather than N in the divisor, where N is the number of persons in the sample.

$$\begin{aligned} c_{12} &= \frac{\sum_{n=1}^N (y_{n1} - \bar{y}_1)(y_{n2} - \bar{y}_2)}{N - 1} \\ &= \frac{\sum_{n=1}^N [(t_n + e_{n1}) - (\bar{t} + \bar{e}_1)][(t_n + e_{n2}) - (\bar{t} + \bar{e}_2)]}{N - 1} \\ &= \frac{\sum_{n=1}^N [(t_n + e_{n1}) - \bar{t}][(t_n + e_{n2}) - \bar{t}]}{N - 1} \\ &= \frac{\sum_{n=1}^N [t_n^2 + t_n e_{n2} - t_n \bar{t} + e_{n1} t_n + e_{n1} e_{n2} - e_{n1} \bar{t} - \bar{t} e_{n2} + \bar{t} \bar{t}]}{N - 1} \\ &= \frac{\sum_{n=1}^N t_n^2 + \sum_{n=1}^N t_n e_{n2} - \sum_{n=1}^N t_n \bar{t} + \sum_{n=1}^N e_{n1} t_n + \sum_{n=1}^N e_{n1} e_{n2} - \sum_{n=1}^N e_{n1} \bar{t} - \sum_{n=1}^N \bar{t} e_{n2} + \sum_{n=1}^N \bar{t}^2}{N - 1} \end{aligned}$$

Now because the error is assumed to be not correlated with the true score, nor with itself across two different tests, the sum of the products of all terms which contain an error term will be 0.

Therefore, the last term above simplifies to

$$\begin{aligned} c_{12} &= \frac{\sum_{n=1}^N t_n^2 - \sum_{n=1}^N t_n \bar{t} - \sum_{n=1}^N \bar{t} t_n + \sum_{n=1}^N \bar{t}^2}{N - 1} \\ &= \frac{\sum_{n=1}^N t_n^2 - 2 \sum_{n=1}^N t_n \bar{t} + N \bar{t}^2}{N - 1} \end{aligned}$$

$$\begin{aligned}
 &= \frac{\sum_{n=1}^N t_n^2 - 2 \sum_{n=1}^N t_n \sum_{n=1}^N \frac{t_n}{N} + N \left(\frac{\sum_{n=1}^N t_n}{N} \right)^2}{N - 1} \\
 &= \frac{\sum_{n=1}^N t_n^2 - \frac{2}{N} \left(\sum_{n=1}^N t_n \right) \left(\sum_{n=1}^N t_n \right) + \frac{N}{N^2} \left(\sum_{n=1}^N t_n \right)^2}{N - 1} \\
 &= \frac{\sum_{n=1}^N t_n^2 - \frac{2}{N} \left(\sum_{n=1}^N t_n \right)^2 + \frac{1}{N} \left(\sum_{n=1}^N t_n \right)^2}{N - 1} \\
 &= \frac{\sum_{n=1}^N t_n^2 - \frac{1}{N} \left(\sum_{n=1}^N t_n \right)^2}{N - 1} \\
 &= \frac{\sum_{n=1}^N (t_n - \bar{t})^2}{N - 1} \\
 &= \frac{SS_t}{N - 1} = s_t^2
 \end{aligned}$$

The second last step is proved as follows:

To show that

$$\frac{\sum_{n=1}^N t_n^2 - \frac{1}{N} \left(\sum_{n=1}^N t_n \right)^2}{N - 1} = \frac{\sum_{n=1}^N (t_n - \bar{t})^2}{N - 1},$$

we work in reverse and just take the numerator for convenience, that is, we show that

$$\sum_{n=1}^N (t_n - \bar{t})^2 = \sum_{n=1}^N t_n^2 - \frac{1}{N} \left(\sum_{n=1}^N t_n \right)^2$$

Proof

$$\begin{aligned}
 \sum_{n=1}^N (t_n - \bar{t})^2 &= \sum_{n=1}^N (t_n^2 - 2\bar{t}t_n + (\bar{t})^2) \\
 &= \sum_{n=1}^N t_n^2 - 2\bar{t} \sum_{n=1}^N t_n + \sum_{n=1}^N \left(\frac{\sum_{n=1}^N t_n}{N} \right)^2 \\
 &= \sum_{n=1}^N t_n^2 - 2 \frac{\sum_{n=1}^N t_n}{N} \sum_{n=1}^N t_n + N \left(\frac{\sum_{n=1}^N t_n}{N} \right)^2 \\
 &= \sum_{n=1}^N t_n^2 - \frac{2}{N} \left(\sum_{n=1}^N t_n \right)^2 + \frac{N}{N^2} \left(\sum_{n=1}^N t_n \right)^2
 \end{aligned}$$

$$\begin{aligned}
&= \sum_{n=1}^N t_n^2 - \frac{2}{N} \left(\sum_{n=1}^N t_n \right)^2 + \frac{1}{N} \left(\sum_{n=1}^N t_n \right)^2 \\
&= \sum_{n=1}^N t_n^2 - \frac{1}{N} \left(\sum_{n=1}^N t_n \right)^2
\end{aligned}$$

In summary, the covariance between two parallel tests is simply the variance of the true scores.

$$c_{12} = s_t^2. \quad (25.4)$$

From this equation, we can derive another relevant relationship based on correlations.

In *Statistics Review 4*, it is shown that the covariance is standardized to a correlation by dividing the covariance by the standard deviations.

$$r_{12} = \frac{c_{12}}{s_1 s_2}. \quad (25.5)$$

However, $s_1^2 = s_t^2 + s_e^2$ and $s_2^2 = s_t^2 + s_e^2$.

Note that here we assume that the error variance on both tests is the same—of course the true scores must be the same. Therefore, the variance of the observed scores is the same. That is, if s_y^2 is the variance of the observed scores y of any two parallel tests 1 and 2, then these variances should be equal: $s_1^2 = s_2^2 = s_y^2$.

Therefore, Eq. (25.5) reduces to

$$r_{yy} = \frac{s_t^2}{s_y s_y} = \frac{s_t^2}{s_y^2}, \quad (25.6)$$

where the correlation of a test with itself is denoted by r_{yy} .

Equation (25.6) represents the proportion of the total variance that is true variance. Thus the reliability r_{yy} of the test is equivalent to the proportion of the total variance that is true score variance.

Equation (25.6) can be further rearranged to give

$$r_{yy} = \frac{s_t^2}{s_t^2 + s_e^2} \quad (25.7)$$

and

$$r_{yy} = \frac{s_y^2 - s_e^2}{s_y^2}. \quad (25.8)$$

It is evident that because (i) $s_e^2 \geq 0$, (ii) s_t^2 cannot be greater than s_y^2 and (iii) s_t^2 and s_y^2 are positive, that r_{yy} has the range 0 to 1, that is

$$0 \leq r_{yy} \leq 1. \tag{25.9}$$

Derivation of the Standard Error of Measurement

By rearranging Eq. (25.8) it is possible to write an equation of the error variance in terms of the reliability.

$$\begin{aligned} s_e^2 &= s_y^2 - r_{yy} s_y^2 \\ &= s_y^2(1 - r_{yy}). \end{aligned} \tag{25.10}$$

Now we need to appreciate what the error variance is; it is the variation of scores from test to test for persons with the same true score or from the same person on more than one occasion.

The *standard deviation* of these scores is given by

$$s_e = s_y \sqrt{(1 - r_{yy})}. \tag{25.11}$$

Equation (25.11) is known as the *standard error of measurement*.

Notice that if the reliability is 1, that is, the scores on the two parallel tests are perfect, then the standard error is 0; if the reliability is 0, then the standard error is the standard deviation of the original scores which means that all of the variations is error variance.

Derivation of the Equation for Predicting the True Score from the Observed Score

From the observed score y_n and the test's reliability, it is possible to estimate the true score t_n and the standard error of this estimate. This estimate is made using the equation for regression.

It will be recalled from *Statistics Review 4* that variable Y for person n can be predicted from variable X using the equation $\hat{Y}_n = b_0 + b_1 X_n$. Note that we now use the subscript n for the person rather than i to be consistent with the presentation here. In this equation, $b_1 = c_{xy}/s_x^2$ and $b_0 = \bar{Y} - b_1 \bar{X}$.

Now in the case where the score y corresponds to the true score t , and y remains as the observed score y , we obtain

$$t_n = b_0 + b_1 y_n, \tag{25.12}$$

where $b_1 = c_{yt}/s_y^2$ and $b_0 = \bar{t} - b_1 \bar{y}$.

In this special case, the covariance c_{yt} can be rearranged as follows. Again, you do not need to know this proof, but it is provided for completeness.

$$\begin{aligned}
 c_{yt} &= \frac{1}{N-1} \sum_{n=1}^N (y_n - \bar{y})(t_n - \bar{t}) \\
 &= \frac{1}{N-1} \sum_{n=1}^N (t_n + e_n - (\bar{t} + \bar{e}))(t_n - \bar{t}) \\
 &= \frac{1}{N-1} \sum_{n=1}^N (t_n + e_n - \bar{t} - \bar{e})(t_n - \bar{t}) \\
 &= \frac{1}{N-1} \sum_{n=1}^N (t_n^2 - t_n\bar{t} + e_n t_n - e_n\bar{t} - \bar{t}t_n + \bar{t}^2 - \bar{e}t_n + \bar{e}\bar{t}) \\
 &= \frac{1}{N-1} \left(\sum_{n=1}^N t_n^2 - \sum_{n=1}^N t_n\bar{t} + \sum_{n=1}^N e_n t_n - \sum_{n=1}^N e_n\bar{t} - \sum_{n=1}^N \bar{t}t_n + \sum_{n=1}^N \bar{t}^2 - \sum_{n=1}^N \bar{e}t_n + \sum_{n=1}^N \bar{e}\bar{t} \right) \\
 &= \frac{1}{N-1} \left(\sum_{n=1}^N t_n^2 - \sum_{n=1}^N t_n\bar{t} - \sum_{n=1}^N \bar{t}t_n + \sum_{n=1}^N \bar{t}^2 \right) \\
 &= \frac{1}{N-1} \left(\sum_{n=1}^N t_n^2 - 2 \sum_{n=1}^N t_n\bar{t} + N\bar{t}^2 \right) \\
 &= \frac{1}{N-1} \left(\sum_{n=1}^N t_n^2 - 2\bar{t} \sum_{n=1}^N t_n + N \left(\frac{\sum_{n=1}^N t_n}{N} \right)^2 \right) \\
 &= \frac{1}{N-1} \left(\sum_{n=1}^N t_n^2 - 2 \frac{\sum_{n=1}^N t_n}{N} \sum_{n=1}^N t_n + \frac{1}{N} \left(\sum_{n=1}^N t_n \right)^2 \right) \\
 &= \frac{1}{N-1} \left(\sum_{n=1}^N t_n^2 - \frac{2}{N} \left(\sum_{n=1}^N t_n \right)^2 + \frac{1}{N} \left(\sum_{n=1}^N t_n \right)^2 \right) \\
 &= \frac{1}{N-1} \left(\sum_{n=1}^N t_n^2 - \frac{1}{N} \left(\sum_{n=1}^N t_n \right)^2 \right) \\
 &= \frac{1}{N-1} \sum_{n=1}^N (t_n - \bar{t})^2 \\
 &= s_t^2
 \end{aligned}$$

Therefore

$$\begin{aligned}
 b_1 &= \frac{c_{yt}}{s_y^2} = \frac{s_t^2}{s_y^2} = r_{yy} \\
 b_0 &= \bar{t} - r_{yy}\bar{y}
 \end{aligned}$$

Therefore

$$\begin{aligned} \hat{t}_n &= \bar{t} - r_{yy}\bar{y} + r_{yy}y_n \\ &= \bar{t} + r_{yy}(y_n - \bar{y}) \end{aligned}$$

In this equation, we apparently do not know the value of \bar{t} .

However, in the population $E[\bar{y}] = \bar{t}$. Therefore, we substitute \bar{y} as an estimate of \bar{t} .

Therefore, finally we can write

$$\hat{t}_n = \bar{y} + r_{yy}(y_n - \bar{y}). \tag{25.13}$$

Thus using Eq. (25.13) we can predict the true score from the observed score and from Eq. (25.11) we can estimate the error in this prediction.

Derivation of Coefficient α

To derive the calculation of coefficient α , we need to involve the number of items. Therefore, we begin with the relationships of the variances at the item level. From Eq. (3.2) in Chap. 3,

$$s_i^2 = s_t^2 + s_e^2, \tag{25.14}$$

where s_i^2 is the variance of the observed scores of any item, s_t^2 is the variance of true scores and s_e^2 is the error variance relative to an item.

Then

$$\begin{aligned} \sum_{i=1}^I s_i^2 &= \sum_{i=1}^I (s_t^2 + s_e^2) \\ \sum_{i=1}^I s_i^2 &= Is_t^2 + Is_e^2 \end{aligned} \tag{25.15}$$

Therefore, subtracting Eq. (25.15) from Eq. (4.3) in Chap. 4 gives

$$\begin{aligned} s_y^2 - \sum_{i=1}^I s_i^2 &= I^2s_t^2 + Is_e^2 - (Is_t^2 + Is_e^2) \\ &= I^2s_t^2 - Is_t^2 \\ &= I(I - 1)s_t^2 \end{aligned} \tag{25.16}$$

Therefore, dividing Eq. (25.16) by Eq. (4.3) in Chap. 4 gives

$$\begin{aligned} \frac{s_y^2 - \sum_{i=1}^I s_i^2}{s_y^2} &= \frac{I(I-1)s_i^2}{I^2 s_i^2 + I s_e^2} \\ &= \frac{I(I-1)s_i^2}{I^2(s_i^2 + s_e^2/I)} \\ &= \frac{(I-1)s_i^2}{I(s_i^2 + s_e^2/I)} \end{aligned} \quad (25.17)$$

from which

$$\begin{aligned} \frac{I}{I-1} \left(\frac{s_y^2 - \sum_{i=1}^I s_i^2}{s_y^2} \right) &= \frac{I}{(I-1)} \cdot \frac{(I-1)s_i^2}{I(s_i^2 + s_e^2/I)} \\ &= \frac{s_i^2}{s_i^2 + s_e^2/I} \\ &= r_{yy} \end{aligned} \quad (25.18)$$

This is the expression for reliability in Eq. (4.5).

Thus to calculate an estimate of reliability according to coefficient α , we can write

$$\alpha = \frac{I}{I-1} \left(\frac{s_y^2 - \sum_{i=1}^I s_i^2}{s_y^2} \right). \quad (25.19)$$

The variances of the total scores and the items, s_y^2 and s_i^2 , are calculated simply as $s_y^2 = \frac{1}{N-1} \sum_{n=1}^N (y_n - \bar{y})^2$ and $s_i^2 = \frac{1}{N-1} \sum_{n=1}^N (x_n - \bar{x})^2$, where N is the number of persons involved.