LETTER TO THE EDITORS

Negative Values of the Intraclass Correlation Coefficient Are Not Theoretically Possible

In their methodological review of the indices of reproducibility, Tammemagi et al. [1] mentioned the intraclass correlation coefficient, which, they say, varies “from -1 for perfect disagreement, to 0 for random agreement, and to +1 for perfect agreement.” However, such an interpretation of negative values of this statistic, although previously asserted by Deyo et al. [12], is misleading. We argue that negative values are not theoretically possible although possibly observable as explained below.

In the framework of a two-way layout model, we define

\[ X_{ij} = \mu + s_i + r_j + e_{ij}, \quad i = 1 \ldots N, \quad j = 1 \ldots k \]

where \( X_{ij} \) is the random variable associated to the \( j \)th measure of the \( i \)th subject, \( \mu \) is a constant equal to the grand mean of all the observations in the population, the subject effects \( S_i \) are independent random variables normally distributed with mean 0 and variance \( \sigma_s^2 \), the rater effects \( R_j \) are independent random variables normally distributed with mean 0 and variance \( \sigma_r^2 \), and the residual errors \( E_{ij} \) are independent random variables normally distributed with mean 0 and variance \( \sigma_e^2 \).

\[
\text{cov}(S_i, E_{ij}) = 0, \quad \text{cov}(R_j, E_{ij}) = 0,
\]

and

\[
\text{cov}(S_i, R_j) = 0
\]

which implies that

\[
\sigma_s^2 = \sigma_r^2 + \sigma_e^2 + \sigma_{sr}^2.
\]

The intraclass correlation coefficient is defined as

\[
\rho = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_r^2 + \sigma_e^2}.
\]

Then, when

\[
\sigma_{sr}^2 = 0
\]

(1) which means perfect agreement (the whole variance is attributable to the variance between subjects).

The estimator of the intraclass correlation coefficient is then estimated, as indicated in the article by Tammemagi et al., by

\[
\hat{\rho} = \frac{N(MSS - MSE)}{(N(MSS) + k(MSR) + (Nk - N - k)MSE)}
\]

This estimator can indeed lead to negative values since if \( MSS = 0 \) and \( MSR = 0 \) the estimate is equal to

\[
-1/(k - 1 - k/N),
\]

which is the lowest value. In the case of two measures by subject, this lowest bound reduces to \( -N/(N - 2) \), which is not very far from -1. Negative values are then possibly observable but there is no other possible interpretation but poor agreement. In the same way, if \( MSS = 0 \) and \( MSR = 0 \) this leads to negative values for \( \hat{\rho}_s \) and \( \hat{\rho}_r \) and, nevertheless, no one would state that a variance could be negative. These negative values of \( \hat{\rho}_s \) and \( \hat{\rho}_r \) are no longer possible as both \( N \) and \( p \) tend to \( +\infty \), because of the consistency of the estimators. It is the same for \( \hat{\rho} \), which is known to be biased but consistent [3,4].

Finally, it should be noted that the authors only consider a two-way layout model (with random or fixed effects) while the one-way layout model (no rater effects) is more widely used in assessing agreement. In this usual framework, the lower bound of the estimate is known to be \(-1/(k - 1) \) [3], which then leads to the value -1 when only two measures by subject are considered. Nevertheless, here again, negative values of the estimate have no theoretical legitimacy.

Bruno Giraudou

Département de Biostatistique et Informatique Médicale
Hôpital Saint-Louis
75475 Paris, France

and Épidémiologie et Sciences de l’Information
INSERM U681
75571 Paris, France.

References

PII: S0895-4356(96)00053-4