

0.1 Intraclass Correlation as a Measure of Agreement

Intraclass correlation is the strength of a linear relationship between subjects belonging to the same class or the same subgroup or the same family. In the agreement setup, the two measurements obtained on the same subject by two observers or two methods is a subgroup. If they agree, the intraclass correlation will be high. This method of assessing an agreement was advocated by Lee et al. (1989).

In the usual correlation setup, the values of two different variables are obtained on a series of subjects. For example, you can have the weight and height of 20 girls aged 57 years. You can also have the weight of the father and mother of 30 low birthweight newborns. Both are weights and the productmoment correlation coefficient is a perfectly valid measure of the strength of the relationship in this case. Now consider the weight of 15 persons obtained on two machines.

Any person, say number 7, may be measured by machine 2 first and then by machine 1. Others may be measured by machine 1 then by machine 2. The order does not matter in this setup as the interest is in finding whether the values are in agreement or not.

Statistically, intraclass correlation is that part of the total variance that is accounted for by the differences in the paired measurements obtained by two methods.

0.1.1 Lin's Reproducibility Index

Lin proposes the use of a reproducibility index, called the Concordance Correlation Coefficient (CCC). While it is not strictly a measure of agreement as such, it can form part of an overall method comparison methodology. Lin proposes the use of a reproducibility index, called the Concordance Correlation Coefficient (CCC). While it is not strictly a measure of agreement as such, we have included it.

0.2 ICC, Reproducibility Index and Passing-Bablok

The ICC, which takes on values between 0 and 1, is based on analysis of variance techniques. It is close to 1 when the differences between paired measurements is very small compared to the differences between subjects. Of these three procedures—t test, correlation coefficient, intra-class correlation coefficient—the ICC is best because it can be large only if there is no bias and the paired measurements are in good agreement, but it suffers from the same faults ii and iii as ordinary correlation coefficients. The magnitude of the ICC can be manipulated by the choice of samples to split and says nothing about the magnitude of the paired differences.

0.3 ICC, Reproducibility Index and Passing-Bablok

0.3.1 Intraclass Correlation Coefficient

This measure of agreement is estimated using variance components from appropriate analysis of variance models. Measures of agreement are variance dependent, and so the ICC can be misleading. The ICC takes a value between 0 and 1, and is based on Analysis of Variance methodologies.

The ICC is a measure of reliability.

? considers the ICC as just another measure of agreement.

Bibliography

Bartko, J. (1994). Measures of agreement: A single procedure. *Statistics in Medicine* 13, 737–745.