Unbiased estimators of kappa coefficients for two raters

Martín Andrés, A.[1] *and Álvarez Hernández, M.*[2]

[1]amartina@ugr.es, Biostatistics, Faculty of Medicine, University of Granada, Spain.
[2]maria.alvarez@cud.uvigo.es, Defense University Center at the Spanish Naval Academy, Spain.

It is often necessary to assess the degree of concordance or agreement between two raters which independently classify n subjects within $K \geq 2$ nominal categories. As some of the observed agreements may be due to chance, it is most common to eliminate the effect of chance by defining a *kappa*-type coefficient. The agreement measure has the expression $\kappa = (I_o - I_e)/(1 - I_e)$ where I_o is the observed index of agreements (sum of the observed proportions of agreements) and I_e is the expected index of agreements (sum of the proportions of agreements that would happen if the two raters acted independently). When the categories are ordinal, or when they are nominal and weights are also assigned to the disagreements obtained, the above definition provides a weighted *kappa* coefficient. In any case, it is common to use the Cohen's *kappa*, Scott's *pi*, Gwet's AC1/2, and Krippendorf's *alpha* coefficients, which are obtained according to the definition adopted for I_e. However, all estimators of previous *kappa* coefficients are biased, since they estimate the product of two population proportions through the product of their sample estimators. The first objective of this study is to correct this bias by proposing unbiased estimators. We also provide the variances of these as a function of the variances of the biased estimators, except in the case of the Gwet estimators. The methodology is easy to apply to any other *kappa* coefficient studied but it may be unnecessary when the sample size n is sufficiently large (e.g. $n \geq 100$). In order to prove this, some simulations and data examples are shown. Finally, we demonstrate that the new unbiased estimator of the Cohen's *kappa* coefficient also coincide with the unbiased estimator of Lin's concordance correlation coefficient if the former are defined assuming quadratic weights.

Keywords: Agreement; *kappa*-type coefficient; Lin's concordance correlation coefficient.