Correct Formulation of the Kappa Coefficient of Agreement

William D. Hudson
Center for Remote Sensing and Department of Forestry, Michigan State University, East Lansing, MI 48824-1111

Carl W. Ramm
Department of Forestry, Michigan State University, East Lansing, MI 48824-1222

INTRODUCTION

SINCE ITS INTRODUCTION to the remote sensing community by Congalton et al. (1983), an increasing number of studies have utilized the Kappa coefficient of agreement as a measure of classification accuracy. It was recommended as a standard by Rosenfield and Fitzpatrick-Lins (1986). Their article is an excellent review of the Kappa coefficient, its variance, and its use for testing for significant differences. Unfortunately, a large number of erroneous formulas and incorrect numerical results have been published. This paper briefly reviews the correct formulation of the Kappa statistic.

Although the Kappa statistic was originally developed by Cohen (1960), most articles cite Bishop et al. (1975) as a source of formulation:

\[
\kappa = \frac{N \sum_{i=1}^{I} x_{ii} - \sum_{i=1}^{I} \sum_{j=1}^{J} x_{ij} x_{ji}}{N^2 - \sum_{i=1}^{I} \sum_{j=1}^{J} x_{ij} x_{ji}}
\]

where \( + \) represents summation over the index.

For computational purposes, the following form is often presented:

\[
\kappa = \frac{\theta_1 - \theta_2}{1 - \theta_2}
\]

where \( \theta_1 = \sum_{i=1}^{I} x_i/N \) and \( \theta_2 = \sum_{i=1}^{I} x_{ii}/N^2 \).

As indicated by Rosenfield and Fitzpatrick-Lins (1986), several earlier versions of the variance of Kappa are incorrect (Cohen, 1960; Spitzer et al., 1967; Cohen, 1968; Everitt, 1968). Although the correct formulation is given by Fleiss et al. (1969), most articles cite Bishop et al. (1975) as their source. The approximate large sample variance of Kappa is

\[
\hat{\sigma}^2 [\kappa] = \frac{1}{N} \left[ \theta_1 (1 - \theta_1) \right] + \frac{2 (1 - \theta_2) (2 \theta_1 - \theta_2)}{(1 - \theta_2)^3} + \frac{(1 - \theta_1)^2 (\theta_2 - 4 \theta_2^2)}{(1 - \theta_2)^4}
\]

where \( \theta_1 = \sum_{i=1}^{I} x_i/N \),

\( \theta_2 = \sum_{i=1}^{I} x_{ii}/N^2 \),

\( \theta_3 = \sum_{i=1}^{I} x_{ii}(x_{ii} + x_{ji})/N^2 \), and

\( \theta_4 = \sum_{i=1}^{I} x_{ii}(x_{ii} + x_{ji})^2/N^3 \).

Although the formulas as presented by Fleiss et al. (1969) and Bishop et al. (1975) appear substantially different, they are algebraically equivalent. Also note that the formulas for \( \theta_1 \), \( \theta_2 \), \( \theta_3 \), and \( \theta_4 \) used by Bishop et al. (1975) assume that proportions have been calculated for individual cells of the classification error matrix. Although the formulation of the Kappa statistic and its variance is correct, the numerical example provided by Bishop et al. (1975, p. 397) contains a numerical error. The correct value of \( \theta_1 \) is 0.49536501 and the correct estimated variance is 0.00823495 (R.G. Oderwald, personal communication).

The potential user of the Kappa coefficient of agreement is cautioned that a number of remote sensing articles contain errors in the formula for the Kappa statistic or its variance. Although the erratum (Photogrammetric Engineering and Remote Sensing, Vol. 50, No. 10, p. 1477) for an article by Congalton et al. (1983) does contain the proper formulas, its appearance ten months after the original article has not been generally referenced when authors cite this work.

A number of published research results (Congalton and Mead, 1983; Congalton et al., 1983; Benson and DeGloria, 1985) contain numerical errors in the reporting of the variance of the Kappa statistic. The errors appear to be caused by the improper computation of the \( \theta_2 \) term in a published computer program (Congalton et al., 1981, 1982). Line 53 of this FORTRAN program uses the \( \theta_4 \) row total plus the \( \theta_4 \) column total instead of the \( \theta_4 \) row total plus the \( \theta_4 \) column total as specified by the formula for \( \theta_2 \); i.e.,

\( \theta_4 = \sum_{i=1}^{I} x_{ii}(x_{ii} + x_{ji})^2/N^3 \)

(line 53 should be corrected to read \( \text{TH4} = \text{TH4} + X(I,J) * (SXR(J) + SXC(I))^2 \); this error is not present in the current version of the program (R.G. Oderwald, personal communication)). Although the error in computing \( \theta_2 \) is numerically small, and thus changes the variance term by only a very small amount, the continued use of the improper formula should be discouraged.

Although potentially very useful in remote sensing accuracy assessment, the user of the Kappa coefficient of agreement should be conscious of its correct formulation and numerical computation.

REFERENCES


I was very happy to see this brief written to clear up some of the confusion surrounding the use of the Kappa statistic. Since we introduced these techniques back in 1980 (Congalton et al., 1980), they have enjoyed increasing popularity. Unfortunately, due to a mistake on my part, incorrect versions of the equations were published in the literature. I applaud Hudson and Ramm for writing this paper and settling once and for all this confusion. The corrections to the KAPPA computer program as described by Hudson and Ramm have indeed been made. In addition, all three programs (KAPPA, MARGFIT, and CONTRABLE) have been converted to run on an IBM PC. If you would like copies of these programs or just an updated KAPPA program, send a request and a blank diskette to me and I will be happy to provide you with this software.

REFERENCES


—Russell G. Congalton
Assistant Professor
Dept. of Forestry and Resource Management,
University of California,
Berkeley, CA 94720

Comments on the Remote Sensing Brief entitled “Correct Formulation of the Kappa Coefficient of Agreement”


(Received 25 September 1986; accepted 11 November 1986; revised 5 December 1986)