

Yilmaz, Ayfer Ezgi; Saracbası, Tulay

Agreement and adjusted degree of distinguishability for square contingency tables. (English)

Zbl 1471.62383

Hacet. J. Math. Stat. 48, No. 2, 592-604 (2019).

Summary: In square contingency tables, analysis of agreement between the row and column classifications is of interest. In such tables, kappa or weighted kappa coefficients are used to summarize the degree of agreement between two raters. In addition to investigate the agreement between raters for square contingency tables, category distinguishability should be considered. Because the kappa coefficient is insufficient to measure the category distinguishability, the degree of distinguishability is suggested to use. In practice, some problems have occurred with regards to the use of the degree of distinguishability. The aim of this study is to assess the agreement coefficient and degree of distinguishability in square contingency tables together. In this study, the adjusted degree of distinguishability is suggested to solve the problem of calculating the degree of distinguishability falls outside the defined range. A simulation study is performed to compare the proposed adjusted degree of distinguishability and the classical degree of distinguishability. Furthermore, interpretation levels for the degree of distinguishability are determined based on a simulation study. The results are discussed over numerical examples and simulation.

MSC:

62H17 Contingency tables

Keywords:

agreement; degree of distinguishability; kappa coefficient; square contingency table

Software:

SPSS; SAS

Full Text: [Link](#)

References:

- [1] Agresti, A. Categorical data analysis (John Wiley and Sons, New York, 2002). · [Zbl 1018.62002](#)
- [2] Becker, M.P. and Agresti, A. Log-linear modelling of pairwise interobserver agreement on a categorical scale, *Statistics in Medicine* 11 (1), 101-114, 1992.
- [3] Cicchetti, D. and Allison, T. A new procedure for assessing reliability of scoring eeg sleep recordings, *American Journal EEG Technology* 11, 101-109, 1971.
- [4] Cohen, J. A coefficient of agreement for nominal scales, *Educational and Psychological Measurement* 20 (1), 37-46, 1960.
- [5] Cohen, J. Weighted Kappa: Nominal scale agreement with provision for scaled disagreement or partial credit, *Psychological Bulletin* 70 (4), 213-220, 1968.
- [6] Darroch, J.N. and McCloud, P.I. Category distinguishability and observer agreement, *Australian Journal of Statistics* 28 (3), 371-388, 1986. · [Zbl 0609.62140](#)
- [7] Fleiss, J.L. and Cohen, J. The equivalence of weighted kappa and the intraclass correlation coefficient as measure of reliability, *Educational and Psychological Measurement* 33, 613- 619, 1973.
- [8] Goktas, A., Isci, O. A comparison of the most commonly used measures of association for doubly ordered square contingency tables via simulation, *Metodoloski Zvezki* 8 (1), 17-37, 2011.
- [9] Holmquist, N.D., McMahon, C.A., and Williams, O.D. Variability in classification of carcinoma in situ of the uterine cervix, *Archives of pathology* 84, 334-345, 1967.
- [10] Landis, J.R. and Koch, G.G. The measurement of observed agreement for categorical data, *Biometrics* 33 (1), 159-174, 1977a. · [Zbl 0351.62039](#)
- [11] Landis, J.R. and Koch, G.G. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers, *Biometrics* 33 (2), 363-374, 1977b. · [Zbl 0357.62037](#)
- [12] Lawal, B. Categorical data analysis with SAS and SPSS applications (Lawrence Erlbaum Associates, Publishers, Inc., New Jersey, 2003).
- [13] Oh, M. Inference on measurements of agreement using marginal association, *Journal of the Korean Statistical Society* 38,

41-46, 2009. · [Zbl 1293.62061](#)

- [14] Perkins, S.M. and Becker, M.P. Assessing rater agreement using marginal association models, *Statistics in Medicine* 21, 1743-1760, 2002.
- [15] Saracbası, T. Agreement models for multiraters, *Turkish Journal of Medical Sciences* 41 (5), 939-944, 2011.
- [16] Shoukri, M.M. Measures of interrater agreement (Chapman & Hall/CRC Press LLC., Florida, 2004).
- [17] Terry, M.B., Neugut, A.I., Bostick, R.M., Potter, J.D., and Haile, R.W. Reliability in the classification of advanced colorectal adenomas, *Cancer Epidemiol Biomarkers & Prevention* 11, 660-663, 2002.
- [18] Tinsley, H.E.A. and Weiss, D.J. Interrater reliability and agreement, in: *Handbook of applied multivariate statistics and mathematical modeling* (Academic Press, New York, 2010).
- [19] Valet, F. and Mary, J.-Y. Power estimation of tests in log-linear nonuniform association models for ordinal agreement, *BMC Medical Research Methodology* 11 (1), 70-80, 2011.
- [20] Agresti, A. *Categorical data analysis* (John Wiley and Sons, New York, 2002). · [Zbl 1018.62002](#)
- [21] Becker, M.P. and Agresti, A. Log-linear modelling of pairwise interobserver agreement on a categorical scale, *Statistics in Medicine* 11 (1), 101-114, 1992.
- [22] Cicchetti, D. and Allison, T. A new procedure for assessing reliability of scoring eeg sleep recordings, *American Journal EEG Technology* 11, 101-109, 1971.
- [23] Cohen, J. A coefficient of agreement for nominal scales, *Educational and Psychological Measurement* 20 (1), 37-46, 1960.
- [24] Cohen, J. Weighted Kappa: Nominal scale agreement with provision for scaled disagreement or partial credit, *Psychological Bulletin* 70 (4), 213-220, 1968.
- [25] Darroch, J.N. and McCloud, P.I. Category distinguishability and observer agreement, *Australian Journal of Statistics* 28 (3), 371-388, 1986. · [Zbl 0609.62140](#)
- [26] Fleiss, J.L. and Cohen, J. The equivalence of weighted kappa and the intraclass correlation coefficient as measure of reliability, *Educational and Psychological Measurement* 33, 613- 619, 1973.
- [27] Goktas, A., Isci, O. A comparison of the most commonly used measures of association for doubly ordered square contingency tables via simulation, *Metodoloski Zvezki* 8 (1), 17-37, 2011.
- [28] Holmquist, N.D., McMahon, C.A., and Williams, O.D. Variability in classification of carcinoma in situ of the uterine cervix, *Archives of pathology* 84, 334-345, 1967.
- [29] Landis, J.R. and Koch, G.G. The measurement of observed agreement for categorical data, *Biometrics* 33 (1), 159-174, 1977a. · [Zbl 0351.62039](#)
- [30] Landis, J.R. and Koch, G.G. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers, *Biometrics* 33 (2), 363-374, 1977b. · [Zbl 0357.62037](#)
- [31] Lawal, B. *Categorical data analysis with SAS and SPSS applications* (Lawrence Erlbaum Associates, Publishers, Inc., New Jersey, 2003).
- [32] Oh, M. Inference on measurements of agreement using marginal association, *Journal of the Korean Statistical Society* 38, 41-46, 2009. · [Zbl 1293.62061](#)
- [33] Perkins, S.M. and Becker, M.P. Assessing rater agreement using marginal association models, *Statistics in Medicine* 21, 1743-1760, 2002.
- [34] Saracbası, T. Agreement models for multiraters, *Turkish Journal of Medical Sciences* 41 (5), 939-944, 2011.
- [35] Shoukri, M.M. Measures of interrater agreement (Chapman & Hall/CRC Press LLC., Florida, 2004).
- [36] Terry, M.B., Neugut, A.I., Bostick, R.M., Potter, J.D., and Haile, R.W. Reliability in the classification of advanced colorectal adenomas, *Cancer Epidemiol Biomarkers & Prevention* 11, 660-663, 2002.
- [37] Tinsley, H.E.A. and Weiss, D.J. Interrater reliability and agreement, in: *Handbook of applied multivariate statistics and mathematical modeling* (Academic Press, New York, 2010).
- [38] Valet, F. and Mary, J.-Y. Power estimation of tests in log-linear nonuniform association models for ordinal agreement, *BMC Medical Research Methodology* 11 (1), 70-80, 2011.

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.