# This Week's Citation Classic

**Goodman L A & Kruskal W H.** Measures of association for cross classification.
*J. Amer. Statist. Ass.* **49**:732-64, 1954.

**In examining a cross classification, how might one usefully measure its degree of association? Our 1954 paper proposed a number of then non-traditional methods, all characterized by operational interpretability in probabilistic terms. Subsequent papers examined distribution theory and statistical inference. [The *Science Citation Index®* (*SCI®*) and the *Social Sciences Citation Index™* (*SSCI™*) indicate that this paper has been cited over 330 times since 1961.]**

Leo A. Goodman & William H. Kruskal
Department of Statistics
University of Chicago
Chicago, IL 60637

March 17, 1978

"In the early 1950s, as young faculty members at the University of Chicago, we had separate conversations with senior colleagues there about statistical treatment of data that were naturally arranged as cross classifications of counts. One of us talked to Bernard Berelson (then Dean of the Graduate Library School and later the President of the Population Council), who was at that time dealing with extensive cross classifications related to voting behavior. For example, he might have a number of cross classifications of intended vote against educational level for different sections of a city.

"The other conversations were with the late Louis Thurstone (a major figure in the field of psychometrics and in particular in the development of factor analysis), who also was dealing with multiple cross classifications in the context of the relationship between various personal characteristics (e.g., leadership ability) and results from various psychological tests.

"In both cases the investigator had substantial numbers of cross classifications and needed a sensible way to reduce the data to try to make it coherent One promising approach was felt to be replacement of each cross classification by a single number that measured in a reasonable way the degree of association between the characteristics corresponding to the rows and columns of the tabulated cross classification. "Thus, the two of us were independently thinking about the same question. We discovered our mutual interest during a conversation at a party –we think that it was a New Year's Eve party at the Quadrangle (Faculty) Club –and the paper grew out of that interaction

"We knew something of the existing literature on measures of association for cross classifications, and as we studied it further we recognized that most suggested measures of association were formal and arbitrary, without relevant interpretations –or. without interpretations at all. Our contribution was to suggest a number of association measures that have interesting interpretations and to provide a simple taxonomy for cross classifications. As an example of the latter, we emphasized the importance of knowing whether or not the categories of a classification have or have not a natural ordering.

"Since cross classifications occur throughout science, since our emphasis on interpretation was perhaps novel, and since our work was quickly incorporated into textbook expositions, citations to the paper became numerous We continued work on the topic, digging more deeply into its history and fields of application, and treating at length the relevant approximate sampling in an effort to contribute some new approaches and to effect some changes in statistical thinking and practice.[1]

"One of us[2] also developed an interest in ordinal measures of association beyond cross classifications as such. The other[3] was led to extensive research in the analysis of multi-way cross classifications, leading to what have come to be known as log-linear model theory and methodology. Another outgrowth, we dare to hope, of our paper has been fresh general concern with descriptive statistics from the viewpoint of finding usefully interpretable characteristics of populations and samples."

1. **Goodman L A & Kruskal W H.** Measures of association of cross classifications. IV: Simplification of asymptotic variances. *J. Amer. Statist. Ass.* **67**:415-21, 1972.
2. **Kruskal W H.** Ordinal measures of association. *J Amer. Statist. Ass.* **53**:814-61, 1958.
3. **Goodman L A.** The muiltivariate analysis of qualitative data: interactions among multiple classifications. *J. Amer. Statist. Ass.* **65**:226-56, 1970.