Comment on Kannemann’s “The Exact Evaluation of 2-way Cross-classifications”

Pieter M. Kroonenberg and Albert Verbeek

University of Leiden and Netherlands Central Bureau of Statistics

Abstract

Kannemann (1982a, b) is criticized on several points, primarily using results available in the literature prior to his papers.

In two papers Kannemann (1982a, b) presents algorithms for the exact distribution of the probability and test statistics in 2-way contingency tables with fixed margins. In his first paper he states that he will not review the literature on the subject as he purports it to be of a purely theoretical nature and not relevant for applied statisticians, researchers, and experimenters.

This disregard for the literature is, however, unfortunate as an essentially identical algorithm to Kannemann (1982b) has been proposed by Hancock (1975). Howell & Gordon (1976) and Cantor (1979) made a number of small improvements on Hancock’s algorithm. More efficient algorithms were, however, already available at that time, viz. Boulton (1974), Agresti & Wackerly (1977), and Baker (1977). The latter has also provisions for handling one or two fixed margins.

At present the state of the art has even been further developed by Pagano & Taylor-Halvorson (1981), Mehta & Patel (1983), and Verbeek, Kroonenberg, and Kroonenberg (1983). A virtually complete survey of the field is given by Verbeek & Kroonenberg (1985).

With respect to closed form formulae for the number of possible tables given the row and column margins, they still do not exist, but good approximation formulae have been developed by Boulton & Wallace (1973), Good (1976), and especially Gail & Mantel (1977).

As to the use of the two-parametric gamma distribution as a better approximation than the one parametric chi-square, it should be noted that Yarnold (1970) presents extensive proof against the superiority of the gamma distribution.

Kannemann’s claim that he can analyse 50×50-tables and three-dimensional...
tables by complete enumeration seems untested and unlikely. The smallest family of 50×50-tables with fixed margins is the one with 50 observations, each marginal frequency being 1. It contains 50!>10^{64} tables, and with current processing speeds this defies enumeration. For three-dimensional tables the smallest families are larger than the third power of families of two-dimensional tables. Only for very small tables and small numbers of observations, this procedure is practical.

References


GAIL, M., and N. MANTEL, 1977: Counting the Number of r×c Contingency Tables. J. Amer. Statist. Assoc. 72, 859—862.


Authors' address:

Dr. P. M. KROONENBERG
Dr. A. VERBEEK
University of Leiden
Department of Education
Stationsplein 12
P. O. Box 9507
2300 RA Leiden
The Netherlands

Manuscript received: June, 20, 1984