Beyond the Chi-Squared: Evaluating Cell-Wise and Sub-Table Methods for Detecting Association in Contingency Tables

Prof. Dr. Atila Göktaş
Professor, Department of Statistics, Mugla Sitki Kocman University, Muğla, Turkey

Angelo Mazza
Department of Economics and Business, University of Catania, Catania, Italy

Antonio Punzo
Department of Economics and Business, University of Catania, Catania, Italy

Abstract

The chi-square test of independence remains the default approach for analyzing contingency tables, however questions persist about its reliability with sparse data and the appropriateness of alternative exact methods. This study evaluates two classes of procedures, cell-wise residuals and Fisher's exact tests applied to individual cells, as well as sub-table scanning using Fisher's exact test on all possible 2×2 partitions across an extensive simulation design encompassing 23 table dimensions (3×3 to 6×12), eight sample sizes (25 to 1000), and four association levels (ρ = 0 to 0.75).

Under the null hypothesis, the chi-square test maintains rejection rates around the nominal 0.05 level across most conditions, including severely sparse tables where 100% of expected cell counts fall below five, conditions where asymptotic approximations are often cautioned against. Cell-wise Fisher exact tests with Bonferroni correction are also conservative but begin to detect associations at moderate sample sizes. In contrast, Fisher's exact test applied to all possible 2×2 sub-tables with Bonferroni adjustment proves extraordinarily conservative, with empirical rejection rates near zero for sample sizes below 200 even under substantial association (ρ = 0.50). Residual-based tests using Bonferroni-adjusted minimum p-values exhibit inflated Type I error, particularly for larger tables. Power comparisons reveal that chi-square detects association at considerably smaller sample sizes than either exact approach, though cell-wise Fisher tests become competitive as sample sizes increase, while the sub-table Fisher lags substantially.

Disagreement between methods is most pronounced for moderate associations (ρ = 0.25-0.50) and larger tables, where rates exceed 0.20 for sample sizes of 200 and reach as high as 0.66 for the largest tables at the smallest sample sizes. These disagreements are driven primarily by the chi-square test rejecting while Fisher 2×2 tests do not, reflecting the substantial conservatism of the sub-table approach or by residual-based tests rejecting while Fisher tests do not, reflecting Type I error inflation in the residual methods. These results suggest that the asymptotic chi-square test is remarkably robust to sparse conditions that would seemingly violate conventional guidelines, that cell-wise Fisher tests with Bonferroni correction offer a viable exact alternative for moderate to large samples, but that the sub-table Fisher approach may be impractically conservative except with very large samples and strong associations. Residual-based tests require more stringent multiplicity adjustments than are commonly recognized. The high disagreement rates between methods underscore the importance of understanding each test's operating characteristics and suggest that reliance on a single method, particularly overly conservative exact tests or inadequately corrected residual approaches may lead to incorrect conclusions about the presence or absence of association.

Keywords

Contingency table, chi-square test, Fisher's exact test, cell-wise tests, sub-table analysis, Bonferroni correction, sparse data.