Contents

conveniently using median differences. These medians are then combined, in a way related to usual analyses of paired-comparisons data, to yield row effects and column effects. This approach may be preferable to median polish, particularly when the two-way table contains holes and has at most 20% bad data values.

3. RESISTANT NONADDITIVE FITS FOR TWO-WAY TABLES

Two-way tables of data may exhibit structure that goes beyond that describable in a simple additive fit. One approach to such nonadditive structure (discussed by Mandel and others) adds a general multiplicative term to produce an additive-plus-multiplicative model. A generalization of median polish yields a resistant fit of this model. New resistant strategies supplement existing ones in assessing various fits and their residuals.

4. THREE-WAY ANALYSES

Tables with a measured response and three or more factors extend the ideas of the two-way table. To gain protection against the adverse effects of isolated unusual data values, median polish is extended to higher-way tables. The analyses parallel the more traditional analysis based on means, to which they are compared. Generalizations of the

two-way diagnostic plot aid in detecting systematic patterns of nonadditivity and in learning whether a power transformation would help promote additivity.

An archaeological example illustrates three methods of identifying outliers in the analysis of large contingency tables. The simulation method generates random entries for tables with independent rows and columns and given margins to get the distribution of the largest entries, using a new standardization. The second method uses fixed margins and applies an exploratory approach to locate outliers among the standardized residuals. The third method adjusts the margins to reduce the impact of anomalous cell counts and again applies an exploratory approach to the residuals to locate outliers. The new standardization improves the normal approximation of the far right-hand tail probability for such right-skewed distributions as binomial, Poisson, and hypergeometric.

These methods, including those based on M-estimators and W-estimators as well as bounded-influence regression, protect against distortion by anomalous data and have good efficiency over a wide range of possibilities for the error structure.

New graphical techniques help to assess how closely an observed discrete frequency distribution follows a member of one of the common families of discrete distributions, such as the binomial or the Poisson. Used carefully, these techniques prevent unusual counts from distorting the overall assessment. The slope of a straight-line pattern in the plot identifies the main parameter of the family of distributions. Other plots help to show whether a cell is discrepant, as well as to assess whether different parts of the data indicate different values of the main parameter.

Instead of the classical measures based on the third and fourth moments, numerical and graphical techniques based on quantiles

serve to describe skewness and elongation. Variants of quantile–quantile plots aid in comparing distributions.