We commend Professor Agresti for pooling together the tremendous recent developments in the area of exact inference for contingence tables and suggesting additional research for the next decade. We would also like to congratulate Dr. Cyrus Mehta on his successful development of the extremely useful computer software, StatXact, an important contribution to our profession.

ACKNOWLEDGMENTS

This research was supported by grants from the National Institutes of Health.

Comment: An Interdisciplinary Approach to Exact Inference for Contingency Tables

Cyrus R. Mehta

I congratulate Professor Agresti for a masterful survey of the blossoming field of exact inference. This paper would not have been as exciting had it been written a decade earlier. Few algorithms were then available for generating permutational distributions or their tail probabilities. Statisticians urgently needing exact tests relied either on brute force exhaustive enumeration of the reference set, or on Monte Carlo sampling. The personal computer industry was in its infancy, and one had to factor the cost of expensive CPU time on a mainframe computer into the decision to compute an exact p-value. But, today, one can buy a 33 mHz 80486 IBM-PC clone for the same price as I paid for my first IBM-XT, $3,300. Yet the 80486 is a hundred times faster. The trend toward increased computing power, more random access memory and more disk storage space, at reduced prices, continues with no end in sight. Are all these computing resources being fully utilized? Let me draw an analogy from the automobile industry. Manufacturers of sports cars are always on the look out for skilled racing drivers able to push a car to its limit by fully utilizing all the available horse power. Similarly, computer manufacturers eagerly solicit software developers whose products can take full advantage of the phenomenal power inside their new machines. Permutational inference is one of the few fields that can satisfy the appetite of an 80486-based PC, a SUN SPARC 2 or a DECstation 5000, eager to devour hard computational problems. Professor Agresti is to be commended for opening up the field and pointing out so many new research directions, guaranteed to keep us occupied for the next decade.

Exact permutational inference is interesting because of its interdisciplinary nature. It draws on ideas from four disciplines: statistics, discrete mathematics, computer science and operations research. I will illustrate this through an exact treatment of the $2 \times k$ contingency table.

1. STATISTICS: THE LINEAR RANK TESTS

Let x denote a generic $2 \times k$ contingency table of the form:

	Col 1	Col 2	...	Col k	Total
Row 1	x_1	x_2	...	x_k	m_k
Row 2	x'_1	x'_2	...	x'_k	m'_k
Total	n_1	n_2	...	n_k	N

Define the reference set of all such $2 \times k$ contingency tables with fixed row and column margins by

$$\Gamma = \left\{ \mathrm{x}: \sum_{i=1}^{k} x_i = m_k,\ x_i + x'_i = n_i,\ i = 1, 2, \ldots k \right\}.$$

For a rich class of statistical problems, the linear rank tests [see, e.g., Chapter 4 of the StatXact, (1991) manual], one needs the permutational distribution of

$$T = \sum_{i=1}^{k} w_i X_i, \tag{1.1}$$

where the w_i's are arbitrary scores, and for any $\mathrm{x} \in \Gamma$,

$$\Pr(X = x) = \prod_{i=1}^{k} \binom{n_i}{x_i} \Big/ \binom{N}{m_k}. \tag{1.2}$$

Cyrus R. Mehta is Associate Professor of Biostatistics, Harvard School of Public Health and President of Cytel Software Corporation, Cambridge, Massachusetts. His mailing address is 137 Erie Street, Cambridge, Massachusetts 02139.