Comment

D. Y. Lin and L. J. Wei

Professor Agresti has provided an excellent review of exact inference for contingency tables. We appreciate the opportunity to discuss this thoughtful and well-written paper.

Like most statistical problems, inferences for contingency tables require elimination of nuisance parameters. In comparing two success probabilities, π_1 and π_2, in a 2×2 table, Professor Agresti emphasizes exact conditional inference that eliminates the nuisance parameter by conditioning on both the row and column margins. An alternative is the unconditional approach, which eliminates the nuisance parameter, say, π_2, by maximizing type I error rate over all plausible values of π_2. The third approach is to integrate out the nuisance parameter by specifying a prior distribution for π_2. A fourth technique is to replace the nuisance quantity by a "reasonable" estimate. The last two approaches are not "exact," in the sense that type I error may not be preserved.

The exact conditional test can be conservative if one insists on a given type I error rate. Although the severity of the problem can be alleviated by reporting the p-value rather then declaring "significant" versus "nonsignificant" at a preset level, the conservativeness does create difficulties in certain situations, such as the sample size calculation. In our opinion, the real problem of using the conditional approach is that it is impossible to construct exact intervals for association measures that are not functions of the odds ratio, a quantity that does not have an intuitive physical interpretation to most practitioners. By contrast, the exact unconditional approach can be used to derive confidence intervals for any function of π_1 and π_2 (Santner and Snell, 1980), for example, the difference $\pi_1 - \pi_2$ and the relative risk π_1 / π_2. A disadvantage of the unconditional inference is that the resulting interval estimate can be quite conservative in certain applications. Moreover, unconditional inference is not feasible for high-dimensional contingency tables at the present stage. So, before we can make "exact" methods more flexible, we still need those "nonpurists" around in our profession. Recent work by Storer and Kim (1990) indicate that some nonexact methods are quite reliable for small samples.

We have found exact methods very useful in settings not considered by Professor Agresti. In the recent paper by Lin, Wei, and DeMets (1991), we studied exact statistical inference for 2×2 tables under group sequential designs. In group sequential trials, the data are reviewed periodically to detect early evidence for the treatment difference. Since there are typically very few subjects in early interim analyses, the large-sample approximations do not work well. The conservativeness of the exact conditional test does become a problem here, because it is necessary to prespecify a sequence of exit probabilities in the group sequential study. We found the unconditional approach quite useful in this case.

When investigating therapies of potentially great benefit, it is more ethical to utilize an adaptive design in which the treatment allocation ratio changes as a function of the observed outcomes. For example, two such designs, the randomized play-the-winner rule (Wei and Durham, 1978) and a two-stage model (Ware, 1990), were instrumental in establishing the effectiveness of extracorporeal membrane oxygenation for treating persistent pulmonary hypertension of the newborn. In adaptive designs, the observed counts tend to be small and unbalanced. As a result, large-sample approximations are usually inadequate. In fact, asymptotic methods may not even exist for some complicated designs (such as Ware's two-stage design). We have developed efficient network algorithms for conducting exact conditional and unconditional inferences for 2×2 tables under arbitrary designs. These algorithms have been used successfully to analyze data from the randomized play-the-winner design (Wei, Smythe, Lin and Parks, 1990) and data from the two-stage design (Lin and Wei, 1990). It should be pointed out that, with an adaptive design, substantial loss of information about the relative merit of the two treatments may result from conditioning on the numbers of subjects assigned to the two groups (Wei, Smythe, Lind and Parks, 1990). Furthermore, because conditioning on both margins drastically reduces the sample space, the conditional distribution for the parameter of interest may be degenerate, as occurred in Ware's design (Lin and Wei, 1990). By contrast, the unconditional approach has been met with greater success.

D. Y. Lin is an Assistant Professor, Department of Biostatistics, University of Washington, Seattle, Washington 98195. L. J. Wei is a Professor, Department of Biostatistics, Harvard University, 677 Huntington Avenue, Boston, Massachusetts 02115.

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.

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.

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

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

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 $\mathrm{x} \in \Gamma$,

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