The Annals of Statistics
1989, Vol. 17, No. 4, 1693–1708

INTERSECTION UNION TESTS FOR STRICT COLLAPSIBILITY IN THREE-DIMENSIONAL CONTINGENCY TABLES

BY LINDA JUNE DAVIS

The structure of the strict collapsibility model in a three-dimensional contingency table is investigated with the goal of developing more powerful tests. After showing that strict collapsibility can be written as a finite union of distinct models, an intersection union test is proposed which is at least as powerful asymptotically as the likelihood ratio test. Intersection union tests for strong collapsibility and marginal independence are also suggested.

1. Introduction. The advantages of collapsing a high-dimensional contingency table include increasing cell frequencies, decreasing number of parameters, easing data interpretation and simplifying graphical and tabular data presentation. Unfortunately, the disadvantages of collapsing include information loss concerning interactions of interest and distortion of these interactions, a phenomenon known as Simpson's paradox [Simpson (1951)]. Whittemore (1978), however, characterizes tables called strictly collapsible for which certain interactions remain unchanged when the contingency table is collapsed. For such tables, the benefits of collapsing can be reaped without paying the price of distorted interactions.

Several tests of strict collapsibility have been proposed. Whittemore (1978), for a $2 \times 2 \times 8$ table, uses the likelihood ratio test (LRT) which cannot be computed using standard log linear model fitting packages. To avoid this computational difficulty, she also suggests using the weighted linear regression technique of Grizzle, Starmer and Koch (1969). Several papers have exploited this last suggestion. Specifically, Cohen, Gatsonis and Marden (1983b) develop such a test for a $2 \times 2 \times 2$ table using the asymptotic normal distribution of the maximum likelihood estimators of two- and three-factor interactions. Ducharme and Lepage (1986) propose a test for a $2 \times 2 \times K$ table using the asymptotic normal distribution of the differences of observed log odds ratios. They also explain how their test extends to $I \times J \times K$ tables.

This paper explores yet another testing method suggested by the model structure of strict collapsibility. Specifically, the strict collapsibility model equals a finite union of distinct models (i.e., no one model entirely contained in another). For such a model, Berger (1982), Berger and Sinclair (1984) and Gleser (1973) suggest the so-called intersection union test (IUT); in fact, Berger and Sinclair (1984) show that the IUT is more powerful than the LRT. This paper considers the IUT specifically applied to testing strict collapsibility and compares it to the LRT.

Received March 1988; revised October 1988.
AMS 1980 *subject classifications*. Primary 62F03; secondary 62F05.
Key words and phrases. Likelihood ratio test, union hypotheses, strong collapsibility, marginal independence.

This paper is organized as follows: In Section 2, strict collapsibility is shown to equal a union of distinct models. In Section 3, the IUT and LRT for union null hypotheses are defined and examined. In Section 4, intersection union tests are suggested for strong collapsibility, as defined by Ducharme and Lepage (1986), and marginal independence via their relationship to strict collapsibility. Finally, in Section 5, the LRT and the IUT are compared via several examples.

2. Strict collapsibility. Suppose objects in a population are classified by three factors, say A, B and C. Factor A has I levels, $i = 1, 2, \dots, I$, factor B has J levels, $j = 1, 2, \dots, J$ and factor C has K levels, $k = 1, 2, \dots, K$, n_{ijk} equals the number of objects in a sample of size N with $A = i$, $B = j$ and $C = k$ and n the entire table of sample frequencies. Also, $p_{ijk} > 0$ is the probability that a selected object has $A = i$, $B = j$ and $C = k$ and p the table of these probabilities. For simplicity, n is assumed to follow a multinomial distribution with parameters $N = \sum_{i,j,k} n_{ijk}$ and p so $\sum_{i,j,k} p_{ijk} = 1$.

As defined by Whittemore (1978), collapsibility over C with respect to the AB two-factor interaction means that this two-factor interaction in the three-dimensional table is identical to that in the AB-marginal table. Strict collapsibility entails, in addition, that the ABC three-factor interaction is 0. Thus, testing strict collapsibility involves testing both equality of a two-factor interaction and the absence of a three-factor interaction.

In a three-dimensional table, Whittemore shows strict collapsibility is equivalent to

$$p_{ijk} = p_{ij+} h_{ik} g_{jk},$$

for some positive constants $\{h_{ik}\}$ and $\{g_{jk}\}$ where $p_{ij+} = \sum_k p_{ijk}$. Hidden in this parametrization are the constraints that $\sum_{i,j} p_{ij+} = 1$ and $\sum_k h_{ik} g_{jk} = 1$ since $p_{ijk}/p_{ij+} = h_{ik} g_{jk}$ and $\sum_k p_{ijk}/p_{ij+} = 1$. Thus, one parametrization of strict collapsibility is

$$\left\{ p > 0 | p_{ijk} = p_{ij+} h_{ik} g_{jk}, \sum_{i,j} p_{ij+} = 1, \sum_k h_{ik} g_{jk} = 1 \, \forall \, i, j \right\}. \tag{1}$$

(Implicitly, $p_{ij+} > 0$, $h_{ik} > 0$ and $g_{jk} > 0$ for all i, j, k.)

As an example of this parametrization, consider a class of tables investigated by Darroch (1962) called perfect contingency tables. A three-dimensional perfect table satisfies

$$\sum_i p_{ij+} p_{i+k} / p_{i++} = p_{+j+} p_{++k} \quad \text{for all } j, k,$$

$$\sum_j p_{ij+} p_{+jk} / p_{+j+} = p_{i++} p_{++k} \quad \text{for all } i, k,$$

$$\sum_k p_{i+k} p_{+jk} / p_{++k} = p_{i++} p_{+j+} \quad \text{for all } i, j.$$

As Darroch shows, if the three-factor interaction is 0, then these constraints imply

$$p_{ijk} = \frac{p_{ij+} p_{i+k} p_{+jk}}{p_{i++} p_{+j+} p_{++k}}.$$

Setting

$$h_{ik} = \frac{p_{i+k}}{p_{i++}p_{++k}} \quad \text{and} \quad g_{jk} = \frac{p_{+jk}}{p_{+j+}},$$

p_{ijk} has the form developed by Whittemore for strictly collapsible tables. The constraint that $\sum_k h_{ik}g_{jk} = 1$ is then just $\sum_k p_{i+k}p_{+jk}/p_{++k} = p_{i++}p_{+j+}$.

This parametrization in (1), however, has two shortcomings. First, $\{h_{ik}\}$ and $\{g_{jk}\}$ are not identifiable since for any nonzero constant f_k, $h_{ik}g_{jk} = (h_{ik}/f_k)(g_{jk}f_k)$. For example, for a perfect contingency table, one could also take

$$h_{ik} = \frac{p_{i+k}}{p_{i++}} \quad \text{and} \quad g_{jk} = \frac{p_{+jk}}{p_{+j+}p_{++k}}.$$

Second, the number of nonredundant constraints among the $I \times J$ constraints $\sum_k h_{ik}g_{jk} = 1$ is difficult to determine due to its dependence on the number of levels I, J and K. This dependency stems from the nonlinearity of these constraints caused by constraining sums of products of individual parameters. In practice, this dependency underlies such properties of strictly collapsible tables as:

> An $I \times J \times 2$ table is strictly collapsible if and only if A and C are conditionally independent given B, or B and C are conditionally independent given A. An $I \times J \times K$ table with $K > 2$, however, can be strictly collapsible without any two factors being conditionally independent given the third.

Also, this dependency makes it impossible to use standard iterative packages for computing maximum likelihood estimates and likelihood ratio tests. As noted by a referee, a practitioner cannot rely on his or her software to determine the number of constraints on $\{h_{ik}\}$ and $\{g_{jk}\}$. Thus, the constraints must be reformulated analytically before computation can proceed using this model.

Reparametrizing (1) to

$$(2) \quad \begin{aligned} &\Big\{p > 0 \,|\, p_{ijk} = p_{ij+}f_k h_{ik} g_{jk}, \sum_{i,j} p_{ij+} = 1, \prod_i h_{ik} = 1\,\forall\, k, \\ &\qquad \prod_j g_{jk} = 1\,\forall\, k, \sum_k f_k h_{ik} g_{jk} = 1\,\forall\, i,j\Big\} \end{aligned}$$

resolves the first problem.

The proposed resolution to the second problem involves relating the constraints $\sum_k f_k h_{ik} g_{jk} = 1$ to constraints on inner products of vectors. This reformulation is suggested by the "sums of products of parameters" form of the constraints. Define

$$\begin{aligned} f &\equiv (f_1, f_2, \ldots, f_K), \\ h_i &\equiv (h_{i1}, h_{i2}, \ldots, h_{iK}) \quad \text{for } i = 1,2,\ldots, I, \\ g_j &\equiv (g_{j1}, g_{j2}, \ldots, g_{jK}) \quad \text{for } j = 1,2,\ldots, J. \end{aligned}$$

Also, let V_A denote the vector space spanned by $\{h_i - h_I, i = 1,2,\ldots, I-1\}$ and V_B the vector space spanned by $\{g_j - g_J, j = 1,2,\ldots, J-1\}$. Furthermore, define an inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^K$ by

$$\langle x, y \rangle = \sum_k f_k x_k y_k.$$

For any two subspaces V_1 and V_2 of $\mathbb{R}^K$, $V_1 \perp V_2$ means V_1 and V_2 are orthogonal with respect to $\langle \cdot, \cdot \rangle$ and $V_1 \perp x$ that $\langle x, y \rangle = 0$ for all $y \in V_1$.

Using this notation, the constraints $\sum_k f_k h_{ik} g_{jk} = 1$ for all i, j, i.e., $\langle h_i, g_j \rangle = 1$ for all i, j, are equivalent to:

1. $\langle h_I, g_J \rangle = 1$.
2. $V_A \perp g_J$.
3. $V_B \perp h_I$.
4. $V_A \perp V_B$.

Of course, this reformulation does not remove the redundancy in the constraints; however, it does relate the redundancy in the constraints to the dimensions of V_A and V_B [denoted $\dim(V_A)$ and $\dim(V_B)$] which as shown below prove easier to handle both analytically and computationally.

Specifically, the next two propositions relate

$$Q \equiv \left\{ (f, h_i, g_j) | f_k, h_{ik}, g_{jk} > 0, \langle h_i, g_j \rangle = 1 \, \forall \, i, j \right\}$$

to a union of

$$S_{a,b} \equiv \left\{ (f, h_i, g_j) | f_k, h_{ik}, g_{jk} > 0, \langle h_I, g_J \rangle = 1, \right.$$
$$\left. V_A \perp g_J, V_B \perp h_I, V_A \perp V_B, \dim(V_A) \le a, \dim(V_B) \le b \right\}$$

over certain a and b. Having done so solves the second problem with the parametrization given in (1) since the number of nonredundant constraints in $S_{a,b}$ is tractable.

PROPOSITION 1. $Q = \bigcup_{a=0}^{K-1} S_{a, K-1-a}$.

PROOF. $Q = \bigcup_{a=0}^{K-1} \bigcup_{b=0}^{K-1} S_{a,b}$ upon noting $\dim(V_A) \le K-1$ since $V_A \perp g_J$ and $g_J \ne 0$ and $\dim(V_B) \le K-1$ since $V_B \perp h_I$ and $h_I \ne 0$. Thus, it suffices to prove $S_{a^*, b^*} \subseteq \bigcup_{a=0}^{K-1} S_{a, K-1-a}$ for fixed $0 \le a^*, b^* \le K-1$. Let $(f, h_i, g_j) \in S_{a^*, b^*}$. Since $\langle h_I, g_J \rangle = 1$ and $V_A \perp g_J$, $h_I \notin V_A$. Therefore, $V_B \perp h_I$ and $V_A \perp V_B$ imply $\dim(V_B) \le K - 1 - \dim(V_A)$. Thus,

$$(f, h_i, g_j) \in S_{\dim(V_A), K-1-\dim(V_A)} \subseteq \bigcup_{a=0}^{K-1} S_{a, K-1-a}. \qquad \square$$

PROPOSITION 2. *If* $I \le K$, $S_{a, K-1-a} \subseteq S_{I-1, K-I}$ *for* $I - 1 \le a \le K - 1$; *if* $J \le K$, $S_{a, K-1-a} \subseteq S_{K-J, J-1}$ *for* $0 \le a \le K - J$.

PROOF. Fix $I - 1 \le a \le K - 1$ and let $(f, h_i, g_j) \in S_{a, K-1-a}$. Then, $\dim(V_A) \le a$ and $\dim(V_B) \le K - 1 - a$. But, since V_A is spanned by $I - 1$ vectors, $\dim(V_A) \le I - 1$. Furthermore, since $a \ge I - 1$, $K - 1 - a \le K - I$.

Thus, $\dim(V_A) \le I - 1$ and $\dim(V_B) \le K - I$ so $(f, h_i, g_j) \in S_{I-1, K-I}$. The second part is proved similarly. □

Combining these two propositions:

THEOREM 3. *If* $K < I + J - 1$, *then*

$$Q = \bigcup_{a=a_1}^{a_2} S_{a, K-1-a},$$

where $a_1 = \max(0, K - J)$ *and* $a_2 = \min(I - 1, K - 1)$. *If* $K \ge I + J - 1$, *then*

$$Q = S_{I-1, J-1}.$$

PROOF. Suppose $K < I + J - 1$. Then, the values of a_1 and a_2 are given in the following table:

	$I \le K, J \le K$	$I \le K, J > K$	$I > K, J \le K$	$I > K, J > K$
a_1	$K - J$	0	$K - J$	0
a_2	$I - 1$	$I - 1$	$K - 1$	$K - 1$

Thus,

$$Q = \bigcup_{a=0}^{K-1} S_{a, K-1-a} \quad \text{(Proposition 1)}$$

$$= \bigcup_{a=a_1}^{a_2} S_{a, K-1-a} \quad \text{(Proposition 2)}.$$

Next, suppose $K \ge I + J - 1$. Then, $0 \le I - 1$ and $J - 1 \le K - 1$ so $S_{I-1, J-1} \subseteq Q$. To prove $Q \subseteq S_{I-1, J-1}$, let $(f, h_i, g_j) \in Q$. Then, since $K \ge I$, $\dim(V_A) \le I - 1$ and since $K \ge J$, $\dim(V_B) \le J - 1$ so $(f, h_i, g_j) \in S_{I-1, J-1}$. Thus, $Q \subseteq S_{I-1, J-1}$. □

To implement the IUT, the union hypothesis must involve nonempty and distinct models. (Here, distinct does not mean disjoint but rather, each model containing at least one point not contained in any other model.) Thus, for later use in developing an IUT:

PROPOSITION 4. *If* $K < I + J - 1$, *then for each* $a_1 \le a^* \le a_2$, *there exists a* $(f, h_i, g_j) \in S_{a^*, K-1-a^*}$ *such that* $\dim(V_A) = a^*$ *and* $\dim(V_B) = K - 1 - a^*$ *and also,* $S_{a^*, K-1-a^*} \not\subseteq \bigcup_{a \ne a^*} S_{a, K-1-a}$. *If* $K \ge I + J - 1$, *there exists a* $(f, h_i, g_j) \in S_{I-1, J-1}$ *such that* $\dim(V_A) = I - 1$ *and* $\dim(V_B) = J - 1$.

PROOF. Suppose $K < I + J - 1$ and fix $a_1 \le a^* \le a_2$. Proving the existence of a $(f, h_i, g_j) \in S_{a^*, K-1-a^*}$ such that $\dim(V_A) = a^*$ and $\dim(V_B) = K - 1 - a^*$ requires only checking subspace dimensions. Now, if this $(f, h_i, g_j) \in S_{a, K-1-a}$ for $a < a^*$, then $\dim(V_A) = a^* \le a$ which is a contradiction. Similarly, if $(f, h_i, g_j) \in S_{a, K-1-a}$ for $a > a^*$, then $\dim(V_B) = K - 1 - a^* \le K - 1 - a$, i.e., $a \le a^*$ which is a contradiction. Thus, $S_{a^*, K-1-a^*} \not\subseteq \bigcup_{a \ne a^*} S_{a, K-1-a}$. When $K \ge I + J - 1$, the existence of a $(f, h_i, g_j) \in S_{I-1, J-1}$ such that $\dim(V_A) = I - 1$ and $\dim(V_B) = J - 1$ also only requires checking subspace dimensions. □

Combining the results of Theorem 3 with the parametrization of strict collapsibility in (2), the strict collapsibility model equals

$$\bigcup_{a=a_1}^{a_2} S^*_{a, K-1-a} \quad \text{if } K < I + J - 1, \tag{3}$$
$$S^*_{I-1, J-1} \quad \text{if } K \ge I + J - 1,$$

where

$$S^*_{a,b} = \Big\{ p | p_{ijk} = p_{ij+} f_k h_{ik} g_{jk};\ p_{ij+}, f_k, h_{ik}, g_{jk} > 0 \,\forall\, i, j, k; \sum_{i,j} p_{ij+} = 1;$$
$$\prod_i h_{ik} = 1, \prod_j g_{jk} = 1 \,\forall\, k; \langle h_I, g_J \rangle = 1, V_A \perp g_J, V_B \perp h_I, \tag{4}$$
$$V_A \perp V_B; \dim(V_A) \le a, \dim(V_B) \le b \Big\}.$$

Furthermore, for $K < I + J - 1, \{S^*_{a, K-1-a}, a_1 \le a \le a_2\}$ are nonempty and distinct models and for $K \ge I + J - 1$, $S^*_{I-1, J-1}$ is nonempty.

3. IUT and LRT for union hypotheses. Let X_m be a random vector whose likelihood function P^m_θ belongs to a class $\{P^m_\theta, \theta \in \Theta^m\}$. Suppose $\Theta^m_r \subseteq \Theta^m$ for $r = 1, 2, \dots, R_m \le R$, where R is a finite, fixed integer not depending on m, and for each r, there exist a test $\phi^m_r(\cdot)$ and a test statistic $T^m_r(\cdot)$ with

$$\phi^m_r(x) = \begin{cases} 1 & \text{if } T^m_r(x) \ge c^m_r, \\ 0 & \text{if } T^m_r(x) < c^m_r, \end{cases} \tag{5}$$

for which ϕ^m_r is asymptotic size α for testing $\theta \in \Theta^m_r$ against $\theta \in \Theta^m - \Theta^m_r$, i.e., $\lim_{m \to \infty} \sup_{\theta \in \Theta^m_r} P_\theta\{\phi^m_r(X_m) = 1\} = \alpha$.

Using T^m_r, $r = 1, 2, \dots, R_m$, define the following two tests for testing $\theta \in \bigcup_{r=1}^{R_m} \Theta^m_r$ against $\theta \in \Theta^m - \bigcup_{r=1}^{R_m} \Theta^m_r$:

$$\phi^*_m(x) = \begin{cases} 1 & \text{if } T^m_r(x) \ge c^m_r \text{ for all } r = 1, 2, \dots, R_m, \\ 0 & \text{if } T^m_r(x) < c^m_r \text{ for some } r = 1, 2, \dots, R_m \end{cases} \tag{6}$$

and

$$\phi^{**}_m(x) = \begin{cases} 1 & \text{if } \min_r T^m_r(x) \ge \max_r c^m_r, \\ 0 & \text{if } \min_r T^m_r(x) < \max_r c^m_r. \end{cases} \tag{7}$$

ϕ^*_m is called an intersection union test (IUT) since it tests each model separately and rejects the union model if and only if each individual model is rejected. ϕ^{**}_m is the likelihood ratio test (LRT) when T^m_r is the likelihood ratio statistic for testing $\theta \in \Theta_r^m$ since

$$-2\log\left(\sup_{\theta \in \bigcup_{r=1}^{R_m} \Theta_r^m} P_\theta^m \Big/ \sup_{\theta \in \Theta^m} P_\theta^m\right) = \min_r \left\{-2\log\left(\sup_{\theta \in \Theta_r^m} P_\theta^m \Big/ \sup_{\theta \in \Theta^m} P_\theta^m\right)\right\}.$$

The next theorem explores properties of ϕ_m^* and ϕ_m^{**}. Specifically, using results in Berger (1982), it shows both ϕ_m^* and ϕ_m^{**} are asymptotic size at most α. Also, it gives sufficient conditions, implicitly in Berger and Sinclair (1984), for ϕ_m^* and ϕ_m^{**} to be asymptotically size α. Berger (1982) also gives sufficient conditions for ϕ_m^* to be exact size α, but these apply only to one-sided individual hypotheses and thus are not applicable to testing strict collapsibility.

THEOREM 5. *The asymptotic size of ϕ_m^* and ϕ_m^{**} is at most α. Furthermore, if there exists a sequence $\theta_m \in \bigcup_{r=1}^{R_m} \Theta_r^m$ such that*

$$\lim_{m\to\infty} \sum_{r=1}^{R_m} P_{\theta_m}\{T_r^m(X_m) < c_r^m\} \leq 1 - \alpha, \tag{8}$$

then the asymptotic size of ϕ_m^* is α; if there exists a sequence $\theta_m \in \bigcup_{r=1}^{R_m} \Theta_r^m$ such that

$$\lim_{m\to\infty} \sum_{r=1}^{R_m} P_{\theta_m}\left\{T_r^m(X_m) < \max_r c_r^m\right\} \leq 1 - \alpha, \tag{9}$$

then the asymptotic size of ϕ_m^{**} is α.

PROOF. For any r^* and $\theta^* \in \Theta_{r^*}^m$,

$$\begin{aligned} P_{\theta^*}\{\phi_m^*(X_m) = 1\} &= P_{\theta^*}\{T_r^m(X_m) \geq c_r^m, \forall\, r\} \\ &\leq P_{\theta^*}\{T_{r^*}^m(X_m) \geq c_{r^*}^m\}. \end{aligned}$$

Therefore, since $R_m \leq R < \infty$ and the individual tests are asymptotic size α for testing $\theta \in \Theta_r^m$,

$$\begin{aligned} \lim_{m\to\infty} \sup_{\theta \in \bigcup \Theta_r^m} P_\theta\{\phi_m^*(X_m) = 1\} &= \lim_{m\to\infty} \max_r \sup_{\theta \in \Theta_r^m} P_\theta\{\phi_m^*(X_m) = 1\} \\ &\leq \max_r \lim_{m\to\infty} \sup_{\theta \in \Theta_r^m} P_\theta\{\phi_m^*(X_m) = 1\} \\ &= \alpha. \end{aligned} \tag{10}$$

Next, suppose there exists a sequence $\{\theta_m\}$ satisfying (8). Since, for any θ,

$$\begin{aligned} P_\theta\{\phi_m^*(X_m) = 1\} &= P_\theta\{T_r^m(X_m) \geq c_r^m, \forall\, r\} \\ &= 1 - P_\theta\{T_r^m(X_m) < c_r^m, \text{for some } r\} \\ &\geq 1 - \sum_{r=1}^{R_m} P_\theta\{T_r^m(X_m) < c_r^m\}, \end{aligned}$$

(8) implies

$$\lim_{m\to\infty} P_{\theta_m}\{\phi_m^*(X_m) = 1\} \geq 1 - \lim_{m\to\infty} \sum_{r=1}^{R_m} P_{\theta_m}\{T_r^m(X_m) < c_r^m\} \geq \alpha.$$

Thus,

$$\lim_{m\to\infty} \sup_{\theta\in\cup\Theta_r^m} P_\theta\{\phi_m^*(X_m) = 1\} \geq \alpha,$$

which, in conjunction with (10), implies the asymptotic size of ϕ_m^* is α. The proof for ϕ_m^{**} is similar. □

These sufficient conditions given in (8) and (9) for asymptotic size α of ϕ_m^* and ϕ_m^{**}, respectively, are mild. They are satisfied, for example, if the individual tests are consistent and no one individual model is contained in the union of the others. For, under these conditions, a sequence θ_m can be chosen such that all summands but one in (8) and (9) go to 0 while the remaining one approaches $1 - \alpha$ since tests of individual models are asymptotic size α.

Berger (1982) and Berger and Sinclair (1984) address other properties of the IUT. First, the IUT performs more like a simultaneous inference procedure than an omnibus test since accepting indicates which hypothesis is most likely true. But, unlike simultaneous inference where inference about individual parameters must be done at error rates less than α to achieve an overall rate of α, each individual test is performed at asymptotic size α to achieve overall asymptotic size at most α. Second, under mild conditions on the individual test statistics and their distributions, when all the individual tests are one-sided, then among all level α tests with monotone rejection regions based on these test statistics, the IUT is uniformly most powerful.

As shown by Berger and Sinclair (1984), however, the main motivation for using an IUT is that the rejection region of ϕ_m^{**} is contained in the rejection region of ϕ_m^*. Thus:

THEOREM 6. *When ϕ_m^* and ϕ_m^{**} have the same asymptotic size, ϕ_m^* is asymptotically uniformly more powerful than ϕ_m^{**}.*

The power advantage of ϕ_m^* over ϕ_m^{**} stems from a difference in the "effective" sizes of the tests of the individual models under ϕ_m^* and ϕ_m^{**}. The critical region of ϕ_m^* consists of the random vectors rejected by all tests of the individual models each at asymptotic size α. The critical region of ϕ_m^{**}, on the other hand, consists of random vectors rejected by all tests of the individual models some of which may be of asymptotic size less than α. (The size may be less than α since the individual tests are essentially compared to the maximum of the critical levels for all individual tests instead of the critical level specific to that test.) Thus, the power of ϕ_m^* is greater than that of ϕ_m^{**} because with ϕ_m^{**}, the asymptotic size of the tests of individual models can be less than α.

Returning to the problem of testing strict collapsibility over C with respect to the AB two-factor interaction, let T_r in (5), (6) and (7) be the LRT for testing $\theta \in \Theta_r$. Results for other tests with similar asymptotic properties, such as Pearson's chi-squared, are immediate.

Now, as shown in Section 2, the strict collapsibility model's structure depends upon whether or not $K \geq I + J - 1$. When $K \geq I + J - 1$, the model equals

TABLE 1

Constraint	Number of restrictions when $K \geq I + J - 1$	Number of restrictions when $K < I + J - 1$
$\Sigma_{i,j} p_{ij+} = 1$	1	1
$\Pi_i h_{ik} = 1 \forall k$	K	K
$\Pi_j g_{jk} = 1 \forall k$	K	K
$\langle h_I, g_J \rangle = 1$	1	1
$V_A \perp g_J$	$I - 1$	a
$V_B \perp h_I$	$J - 1$	$K - 1 - a$
$V_A \perp V_B$	$(I - 1)(J - 1)$	$a(K - 1 - a)$
$\dim(V_A) \leq a$	0	$(I - 1 - a)(K - a)$
$\dim(V_B) \leq b$	0	$(J - K + a)(a + 1)$
	$IJ + 2K + 1$	$a(J + 1 - I - K + a)$ $+(J + 1 + K + KI)$

$S^*_{I-1, J-1}$ as defined in (4). Since in this case the strict collapsibility model does not equal a union of distinct models, one cannot define an IUT. Thus, when $K \geq I + J - 1$, the power of the LRT cannot be improved upon by an IUT.

Using standard asymptotic theory for the LRT [see, e.g., Diamond (1963) and Serfling (1980), Section 4.4], the asymptotic distribution of the LRT when $K \geq I + J - 1$ is chi-squared (χ^2_d) with degrees of freedom (d.f.)

$$d \equiv (\text{number of parameters under alternative model}) - (\text{number of parameters under null model}). \tag{11}$$

The number of parameters under the alternative model is $IJK - 1$ since the only restriction placed on p is $\Sigma_{i,j,k} p_{ijk} = 1$. The number of parameters under the null model equals the number of parameters in $p_{ij+} f_k h_{ik} g_{jk}$ which is $IJ + K + IK + JK$ minus the number of restrictions on p_{ij+}, f_k, h_{ik} and g_{jk} which as shown in Table 1 is $IJ + 2K + 1$. Thus,

$$d = IJK - 1 - (IJ + K + IK + JK - IJ - 2K - 1)$$
$$= (I - 1)(J - 1)K.$$

Therefore, one test for strict collapsibility when $K \geq I + J - 1$ is the LRT having an asymptotic χ^2_d distribution with $d = (I - 1)(J - 1)K$.

Next, suppose $K < I + J - 1$. Then, the strict collapsibility model equals $\bigcup_{a=a_1}^{a_2} S^*_{a, K-1-a}$ with $S^*_{a, K-1-a}$ defined in (4). Since in this case the strict collapsibility model equals a union of distinct models, one can define an IUT.

To this end, for $S^*_{a, K-1-a}$, unlike for $S^*_{I-1, J-1}$, the constraints $\dim(V_A) \leq a$ and $\dim(V_B) \leq K - 1 - a$ place restrictions on the parameters. In fact, $S^*_{a, K-1-a}$ must also be expressed as a union to apply asymptotic theory since $S^*_{a, K-1-a}$ does not specify which vectors in $\{h_i - h_I, i = 1, 2, \ldots, I - 1\}$ and which in $\{g_j - g_J, j = 1, 2, \ldots, J - 1\}$ are linearly independent. In particular, $S^*_{a, K-1-a} = \bigcup_{l_A, l_B} Z_{l_A, l_B}$, where l_A ranges over all subsets of a indices from $\{1, 2, \ldots, I - 1\}$, l_B ranges over all subsets of $K - 1 - a$ indices from

$\{1, 2, \ldots, J-1\}$ and $Z_{l_A l_B}$ is identical to $S^*_{a, K-1-a}$ with the added constraints that $h_i - h_I$ for $i \notin l_A$ is in the span of $\{h_i - h_I, i \in l_A\}$ and $g_j - g_J$ for $j \notin l_B$ is in the span of $\{g_j - g_J, j \in l_B\}$. Since the LRT for testing $Z_{l_A l_B}$ has the same d.f. regardless of the values of l_A and l_B, the IUT and the LRT for testing $S^*_{a, K-1-a}$ are identical. Furthermore, the assumptions in Theorem 5 hold for the LRT of $Z_{l_A l_B}$. Thus, the LRT of $S^*_{a, K-1-a}$ has an asymptotic $\chi^2_{d(a)}$ distribution where $d(a)$ equals the d.f. associated with $Z_{l_A l_B}$.

Now, $d(a)$ associated with $Z_{l_A l_B}$ is still given by (11) so using Table 1,

$$
\begin{aligned}
d(a) &= IJK - 1 - \{IJ + K + IK + JK - a(J + 1 - I - K + a) \\
&\qquad\qquad - (J + 1 + K + KI)\} \qquad (12)\\
&= IJ(K-1) - a(K - 1 - a) - aI - (K - 1 - a)J.
\end{aligned}
$$

In particular, note that $d(0) = (I-1)J(K-1)$ and $d(K-1) = I(J-1) \times (K-1)$.

Thus, to test for strict collapsibility when $K < I + J - 1$, one can use either the IUT or the LRT based on the individual LRT for testing $S^*_{a, K-1-a}$ which has an asymptotic $\chi^2_{d(a)}$ distribution with $d(a)$ given in (12). Proposition 4 and Theorem 5 imply both tests are asymptotically size α. If $d(a)$ does not depend upon a, then the two tests are identical. If $d(a)$ depends on a, however, Theorem 6 implies the IUT is asymptotically uniformly more powerful than the LRT.

The amount of improvement in power offered by the IUT over the LRT hinges on the variation in the degrees of freedom associated with individual models. For example, in a $4 \times 8 \times 4$ table, $d(a)$ varies from 72 to 84. Thus, for the LRT, all individual likelihood ratio statistics are compared to a quantile of χ^2_{84} while for an IUT, the individual likelihood ratio statistics are each compared to the associated quantile with degrees of freedom ranging from 72 to 84. Since, for example, the 0.95 quantile of χ^2_{84} is 106.39 while that for χ^2_{72} is only 92.81, the IUT is much more powerful in this case. On the other hand, for a $8 \times 4 \times 2$ table, $d(a)$ varies only from 24 to 28 for all individual models so the IUT and LRT are more comparable in power.

In general, computing the LRT for $S^*_{I-1, J-1}$ or $S^*_{a, K-1-a}$ cannot be done using standard statistical packages for contingency table analysis since the model $S^*_{a, b}$ is not log linear or linear. The statistic, however, can be computed, using any package which minimizes a smooth function subject to constraints including simple bounds, linear constraints and smooth nonlinear constraints on the variables.

When $a = 0$ or $a = K - 1$, however, the LRT of $S^*_{a, K-1-a}$ can be computed directly. When $a = 0$, $S^*_{a, K-1-a}$ is the hierarchical log linear model that A and C are conditionally independent given B (denoted $A \otimes C|B$) while when $a = K - 1$, it is the model that B and C are conditionally independent given A (denoted $B \otimes C|A$).

In particular, when $a_1 = 0$ and/or $a_2 = K - 1$, the LRT of $A \otimes C|B$ and $B \otimes C|A$ supply a lower bound on the p-value of the IUT and LRT. For example, if $K \le I$ and $K \le J$, then provided the maximum p-value of these two tests is large, the strict collapsibility model would be accepted via the IUT without computing the additional LRT. A similar trick works for $K \geq I + J - 1$. Hence, since $A \otimes C|B$ and $B \otimes C|A$ imply strict collapsibility, the LRT of $S^*_{I-1,\,J-1}$ is smaller than the minimum of the LRT of these two conditional independence models. Thus, the p-value for $S^*_{I-1,\,J-1}$ is bounded below by the p-value of the minimum of the two LRT's.

As an aside, since $S^*_{a,\,K-1-a}$ is a hierarchical log linear model if and only if $a = 0$ or $K - 1$, the strict collapsibility model reduces to a union of hierarchical log linear models if and only if $K = 2$ (in which case it reduces to the union of $A \otimes C|B$ and $B \otimes C|A$). The sufficiency of $K = 2$ is proved by Whittemore (1978).

Other individual test statistics besides the LRT can be used in ϕ^* and ϕ^{**}. For example, Cohen, Gatsonis and Marden (1983b) propose a test for strict collapsibility in a $2 \times 2 \times 2$ table which rejects if

$$\min\left\{(\hat{u}_{123}, \hat{u}_{13})\hat{\Sigma}_{13}^{-1}(\hat{u}_{123}, \hat{u}_{13})', (\hat{u}_{123}, \hat{u}_{23})\hat{\Sigma}_{23}^{-1}(\hat{u}_{123}, \hat{u}_{23})'\right\} > z_2(\alpha),$$

where $\hat{u}_{123}$, $\hat{u}_{13}$ and $\hat{u}_{23}$ are the maximum likelihood estimators under a saturated model of u_{123} the ABC three-factor interaction, u_{13} the AC two-factor interaction and u_{23} the BC two-factor interaction, $\hat{\Sigma}_{13}$ and $\hat{\Sigma}_{23}$ estimators of the corresponding asymptotic covariance matrices and $z_2(\alpha)$ the $1 - \alpha$ quantile of a χ^2_2 distribution. This is an IUT and also a ϕ^{**} test since the individual tests for $\{u_{123} = u_{13} = 0\}$ and $\{u_{123} = u_{23} = 0\}$ have the same d.f. Although originally proposed and studied only for testing strict collapsibility in a $2 \times 2 \times 2$ table, this test works for any $I \times J \times 2$ table via treating $\hat{u}_{123}$, $\hat{u}_{13}$ and $\hat{u}_{23}$ as vectors; ϕ^* and ϕ^{**}, however, are now distinct tests provided $I \neq J$.

4. Strong collapsibility and marginal independence. Durcharme and Lepage (1986) define an $I \times J \times K$ table to be strongly collapsible over C with respect to the AB two-factor interaction if all $I \times J \times K'$ tables with $K' \leq K$ obtained from the original table by combining levels of C are strictly collapsible over C with respect to the AB interaction. They also prove an $I \times J \times K$ table is strongly collapsible with respect to the AB interaction if and only if $A \otimes C|B$ or $B \otimes C|A$.

Since strong collapsibility equals a union of distinct models, the results of Section 3 suggest using an IUT. Therefore, one test of strong collapsibility is to test $A \otimes C|B$ and $B \otimes C|A$ separately by LRT and reject if and only if both models are rejected. This IUT is asymptotically size α and provided $I \neq J$, asymptotically more powerful than the LRT (which is also asymptotically size α).

Cohen (1981) and Cohen, Gatsonis and Marden (1983a, b) consider testing marginal independence of A and B (denoted $A \otimes B$) given $A \otimes B|C$. Given $A \otimes B|C$, $A \otimes B$ if and only if p is strictly collapsible over C with respect to the AB interaction. Hence, in this special case, testing marginal independence is equivalent to testing the null hypothesis that $A \otimes B|C$ and p is strictly collapsible against the alternative that $A \otimes B|C$. Using (1), the set of p such that p is strictly collapsible and $A \otimes B|C$ equals

$$\left\{p > 0 | p_{ijk} = p_{i++}p_{+j+}h_{ik}g_{jk}, \sum_i p_{i++} = 1, \sum_j p_{+j+} = 1, \sum_k h_{ik}g_{jk} = 1 \forall\, i, j\right\}.$$

Thus, arguing as in Section 2, the null model equals

$$(13) \qquad \begin{aligned} &\bigcup_{a=a_1}^{a_2} S^{**}_{a, K-1-a} \quad \text{if } K < I + J - 1, \\ &S^{**}_{I-1, J-1} \quad \text{if } K \geq I + J - 1, \end{aligned}$$

where

$$S^{**}_{a,b} = \Big\{ p | p_{ijk} = p_{i++}p_{+j+}f_k h_{ik} g_{jk};\ p_{i++}, p_{+j+}, f_k, h_{ik}, g_{jk} > 0 \forall\, i, j, k;$$
$$\sum_i p_{i++} = 1, \sum_j p_{+j+} = 1; \prod_i h_{ik} = 1, \prod_j g_{jk} = 1 \forall\, k; \langle h_I, g_J \rangle = 1,$$
$$V_A \perp g_J, V_B \perp h_I, V_A \perp V_B; \dim(V_A) \leq a, \dim(V_B) \leq b \Big\}.$$

Using (13), when $K \geq I + J - 1$, one can test $A \otimes B$ against the alternative $A \otimes B|C$ using the LRT of $S^{**}_{I-1, J-1}$ against $A \otimes B|C$ which has an asymptotic χ^2_d distribution with $d = (I - 1)(J - 1)$. When $K < I + J - 1$, one can test $A \otimes B$ against the alternative $A \otimes B|C$ using the LRT or the IUT based on the individual LRT for testing $S^{**}_{a, K-1-a}$ against $A \otimes B|C$ which has an asymptotic $\chi^2_{d(a)}$ distribution with

$$d(a) = (I + J - 1)(K - 1) - a(K - 1 - a) - aI - (K - 1 - a)J.$$

The IUT and LRT are asymptotically size α but the IUT is asymptotically more powerful.

As before, when $a = 0$ or $a = K - 1$, $S^{**}_{a, K-1-a}$ reduces to a hierarchical log linear model—to $A \otimes (B, C)$ when $a = 0$ and to $B \otimes (A, C)$ when $a = K - 1$. Furthermore, the LRT of $A \otimes (B, C)$ against $A \otimes B|C$ equals the LRT of $A \otimes C$ against the saturated model in the $I \times K$ AC-marginal table and similarly, the LRT of $B \otimes (A, C)$ against $A \otimes B|C$ equals the LRT of $B \otimes C$ against the saturated model in the $J \times K$ BC-marginal table. Thus, the LRT of $S^{**}_{a, K-1-a}$ when $a = 0$ or $a = K - 1$ is easily computed. In particular, if $K = 2$, the IUT of $A \otimes B$ given $A \otimes B|C$ rejects if and only if the LRT of $A \otimes C$ and the LRT of $B \otimes C$ reject.

Other tests of $A \otimes B$ given $A \otimes B|C$ have been proposed. For example, Cohen, Gatsonis and Marden (1983a) propose several tests for $2 \times 2 \times 2$ tables. One is an exact IUT using two-sided Fisher's exact tests for independence in the AC-marginal table and in the BC-marginal table. Cohen, Gatsonis and Marden (1983a) prove this IUT is not size α when the individual tests are.

They also propose a large sample IUT based on

$$T_1 = (\hat{\tau}_1 - \hat{\tau}_3)/\hat{\sigma}_{13},$$
$$T_2 = (\hat{\tau}_5 - \hat{\tau}_6)/\hat{\sigma}_{56},$$

where

$$\tau_1 = P\{B = 2|A = 1, C = 1\},$$
$$\tau_3 = P\{B = 2|A = 1, C = 2\},$$
$$\tau_5 = P\{C = 1|A = 1\},$$
$$\tau_6 = P\{C = 1|A = 2\}$$

and $\hat{\tau}_1$, $\hat{\tau}_3$, $\hat{\tau}_5$, $\hat{\tau}_6$ are maximum likelihood estimators under the conditional independence model with $\hat{\sigma}_{13}$ and $\hat{\sigma}_{56}$ the associated estimated asymptotic standard errors. If instead of estimating σ_{13} (σ_{56}) by replacing τ_1 (τ_5) by $\hat{\tau}_1$ ($\hat{\tau}_5$) and τ_3 (τ_6) by $\hat{\tau}_3$ ($\hat{\tau}_6$) [as done by Cohen, Gatsonis and Marden (1983a)], one estimates σ_{13} (σ_{56}) by replacing τ_1 and τ_3 (τ_5 and τ_6) by the pooled estimate under the null hypothesis, then the above IUT is based on Pearson's chi-squared statistics for independence in the AC-marginal and BC-marginal tables.

5. Examples. Fienberg (1980), pages 11–13, considers a $2 \times 2 \times 2$ table involving the habitat of sagrei adult male Anolis lizards of Bimini as compared to distichus adult and subadult Anolis lizards of Bimini. For each lizard, perch height and perch diameter were recorded and subsequently dichotomized resulting in the following $2 \times 2 \times 2$ table:

		Sagrei lizards		Distichus lizards	
		Perch diameter (inches)		Perch diameter (inches)	
		≤ 4.0	> 4.0	≤ 4.0	> 4.0
Perch height (feet)	> 4.75	32	11	61	41
	≤ 4.75	86	35	73	70

Let A denote perch height, B perch diameter and C species.

Now, for a $2 \times 2 \times 2$ table, the strict collapsibility model is the union of $S^*_{0,1}$ and $S^*_{1,0}$ with the LRT for testing $S^*_{0,1}$ and $S^*_{1,0}$, each having asymptotic χ^2_2 distributions. Since both individual LRTs have the same asymptotic distribution, the IUT and the LRT for strict collapsibility over C with respect to the AB two-factor interaction are identical. Also, since $K = 2$, the IUT and LRT for strict collapsibility are also tests of strong collapsibility. The LRT for testing $S^*_{0,1}$ which equals the log linear model $A \otimes C|B$ is 11.82 on 2 d.f. with p-value = 0.003 and the LRT for testing $S^*_{1,0}$ which equals the model $B \otimes C|A$ is 14.02 on 2 d.f. with p-value = 0.001 making the p-value of the IUT (or LRT) for strict (or strong) collapsibility 0.003. Thus, the model of strict (or strong) collapsibility is rejected so collapsing the two tables into one by summing over species would change the relationship between perch height and perch diameter.

Other proposed tests for strict collapsibility lead to the same conclusion. For example, the test statistic proposed by Ducharme and Lepage (1986) equals 7.19 on 2 d.f. with p-value = 0.027. For the IUT proposed by Cohen, Gatsonis and Marden (1983a), $T_1 = 11.42$ and $T_2 = 13.53$ both on 2 d.f. resulting in an overall p-value of $\max(0.003, 0.001) = 0.003$.

Turning to a more complex table, consider a $3 \times 2 \times 4$ table from Aickin (1983), page 158. Prior to voting, voters expressed their political preference (Democrat, Independent, Republican) and then after voting, their actual vote (Democrat, Republican) was determined. The voters were also classified according to income level (in thousands of dollars per year).

	< 10		10–15		15–20		> 20	
	Dem	**Rep**	**Dem**	**Rep**	**Dem**	**Rep**	**Dem**	**Rep**
Dem	112	7	83	13	76	8	86	11
Ind	67	37	75	45	67	57	67	63
Rep	5	35	3	28	9	35	14	71

Let A denote political preference, B actual vote and C income level.

First, consider testing collapsibility over C with respect to the AB interaction. In this case, $K = I + J - 1$. Thus, testing strict collapsibility involves computing the LRT of the model $S_{2,1}^*$ having an asymptotic χ_8^2 distribution. Since the minimum of the LRT of $A \otimes C|B$ and $B \otimes C|A$ is an upper bound on the LRT of $S_{2,1}^*$, the trick mentioned in Section 3 can be used to alleviate this computational task. The LRT of $A \otimes C|B$ is 25.20 (12 d.f., p-value = 0.014) and of $B \otimes C|A$ is 11.81 (9 d.f., p-value = 0.224) so the p-value for strict collapsibility is at least $P\{\chi_8^2 \geq 11.81\} = 0.160$. Thus, the null hypothesis of strict collapsibility would not be rejected so the relationship between political preference and actual vote could be studied in the marginal table. For testing strong collapsibility, the values of the necessary LRT are given above. The p-value of the IUT is $\max(0.014, 0.224) = 0.224$ so the strong collapsibility model is not rejected. The p-value for the LRT is $P\{\chi_{12}^2 > 11.81\} = 0.461$ so the model is again not rejected.

Next, consider testing strict collapsibility over A with respect to the BC interaction. Here, $K < I + J - 1$. Now, $a_1 = \max(0, K - J) = 0$ and $a_2 = \min(I - 1, K - 1) = 1$ so the strict collapsibility model equals $S_{0,2}^* \cup S_{1,1}^*$. As pointed out before, $S_{0,2}^*$ equals $B \otimes A|C$; the LRT for this model is 343.61 on 8 d.f. with p-value ≈ 0. Using NPSOL, a FORTRAN package for nonlinear programming developed by Gill, Murray, Saunders and Wright (1984) of the Systems Optimization Laboratory in the Department of Operations Research at Stanford University, the LRT for $S_{1,1}^*$ is 16.52 on 9 d.f. with p-value = 0.057 implying the p-value of the IUT equals 0.057 and the p-value of the LRT equals 0.057. Therefore, at the 0.05 level, the model of strict collapsibility over A is not rejected using either the IUT or the LRT.

For testing strong collapsibility, the LRT of $B \otimes A|C$ is 343.61 on 8 d.f. with p-value ≈ 0 and the LRT of $C \otimes A|B$ is 25.20 on 12 d.f. with p-value = 0.014 implying the p-value for the IUT and LRT is 0.014. Thus, at the 0.05 level, the strong collapsibility model is rejected even though the strict collapsibility model is not rejected.

Since the LRT of $B \otimes C|A$ is 11.81 on 9 d.f. with p-value = 0.224, one might test $B \otimes C$ given $B \otimes C|A$. The LRT associated with $S_{0,2}^{**}$ is 349.76 on 2 d.f. with p-value ≈ 0 and with $S_{1,1}^{**}$ is 22.67 on 3 d.f. with p-value ≈ 0. Thus, both the IUT and the LRT p-value for marginal independence is approximately 0.

The LRT for independence of B and C computed using the BC-marginal table is 17.96 on 3 d.f. with p-value ≈ 0 agreeing with the test of $B \otimes C$ given $B \otimes C|A$.

6. Concluding remarks. The results of this paper revolve around the parametrization in (3) and (4) of the strict collapsibility model. This is certainly not the only way to reparametrize the model; for example, as pointed out by Cohen (1985), the $\{f_k\}$ could be eliminated by removing the restriction on $\prod g_{jk}$. Of several different parametrizations investigated, this one was selected for several reasons. First, it is symmetric in its treatment of h and g which simplifies the discussion. Also, the sets making up the union are open so that the LRT can be applied to each set separately while still appealing to regular asymptotic theory.

No claim is made as to this being the best parametrization in the sense of creating the most powerful test. The intent is only to identify a test at least as powerful as the LRT and to expose the complexity of the strict collapsibility model as rooted in its union nature.

This parametrization of the strict collapsibility model, however, reveals that testing strict collapsibility hinges on the size of the table. When $K \geq I + J - 1$, one test of strict collapsibility is the LRT having an asymptotic χ_d^2 distribution with $d = (I-1)(J-1)K$. When $K < I + J - 1$, however, one can test strict collapsibility via an IUT, based on LRT of individual models, which is asymptotically at least as powerful as the LRT for strict collapsibility.

Acknowledgments. I wish to thank Professor Arthur Cohen for his comments and suggestions which helped clarify my presentation of this topic. Also, thanks to several referees for their comments on the paper and their suggestions of some additional topics to investigate.

REFERENCES

AICKIN, M. (1983). *Linear Statistical Analysis of Discrete Data*. Wiley, New York.

BERGER, R. L. (1982). Multiparameter hypothesis testing and acceptance sampling. *Technometrics* **24** 295–300.

BERGER, R. L. and SINCLAIR, D. F. (1984). Testing hypotheses concerning unions of linear subspaces. *J. Amer. Statist. Assoc.* **79** 158–163.

COHEN, A. (1981). Inference for marginal means in contingency tables with conditional independence. *J. Amer. Statist. Assoc.* **76** 895–902.

COHEN, A. (1985). Personal communication.

COHEN, A., GATSONIS, C. and MARDEN, J. I. (1983a). Hypothesis testing for marginal probabilities in a $2 \times 2 \times 2$ contingency table with conditional independence. *J. Amer. Statist. Assoc.* **78** 920–929.

COHEN, A., GATSONIS, C. and MARDEN, J. I. (1983b). Hypothesis tests and optimality properties in discrete multivariate analysis. In *Studies in Econometrics, Time Series, and Multivariate Statistics* (S. Karlin, T. Amemiya and L. A. Goodman, eds.) 379–405. Academic, New York.

DARROCH, J. N. (1962). Interactions in multi-factor contingency tables. *J. Roy. Statist. Soc. Ser. B* **24** 251–263.

DIAMOND, E. L. (1963). The limiting power of categorical data chi-square tests analogous to normal analysis of variance. *Ann. Math. Statist.* **34** 1432–1441.

DUCHARME, G. R. and LEPAGE, Y. (1986). Testing collapsibility in contingency tables. *J. Roy. Statist. Soc. Ser. B* **48** 197–205.

FIENBERG, S. E. (1980). *The Analysis of Cross-Classified Categorical Data*. MIT Press, Cambridge, Mass.

GILL, P. E., MURRAY, W., SAUNDERS, M. A. and WRIGHT, M. H. (1984). User's guide to NPSOL (Version 2.1): A FORTRAN package for nonlinear programming. Technical Report SOL 84-7, Dept. Operations Research, Stanford Univ.

GLESER, L. J. (1973). On a theory of intersection–union tests. *Inst. Math. Statist. Bull.* **2** 233.

GRIZZLE, J. E., STARMER, C. F. and KOCH, G. C. (1969). Analysis of categorical data by linear models. *Biometrics* **25** 489–504.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

SIMPSON, C. H. (1951). The interpretation of interaction in contingency tables. *J. Roy. Statist. Soc. Ser. B* **13** 238–341.

WHITTEMORE, A. S. (1978). Collapsibility of multidimensional contingency tables. *J. Roy. Statist. Soc. Ser. B* **40** 328–340.

11539 MAMIE LANE
FAIRFAX STATION, VIRGINIA 22039