INTERACTIONS IN MULTIDIMENSIONAL CONTINGENCY TABLES[1]

By Leo A. Goodman

University of Chicago

1. Introduction and summary. In the present article, we shall propose a definition of the rth order interactions in a m-dimensional $d_1 \times d_2 \times \cdots \times d_m$ contingency table ($r = 0, 1, 2, \cdots, m - 1$), and we shall present methods for testing the hypothesis that any specified subset of these interactions is equal to zero. In addition, we shall present simple methods for obtaining simultaneous confidence intervals for these interactions or for any specified subset of them.

In the special case where the m-dimensional contingency table is a $2 \times 2 \times \cdots \times 2$ table (i.e., where $d_i = 2$ for $i = 1, 2, \cdots, m$), the rth order interactions defined herein are the same as Good's interactions [9], but the tests proposed by Good are different even in this case from those presented herein. When $d_i > 2$ for some values of i, Good's interactions are complex valued, whereas the interactions presented here are real valued. We shall show herein that the hypothesis H_r that all rth order and higher-order complex interactions (defined by Good) are equal to zero is equivalent to the hypothesis H_r^* that all rth order and higher-order real interactions (defined herein) are equal to zero, and that the test of H_r within H_s ($r < s$) presented by Good is asymptotically equivalent (under H_r) to the test of H_r^* within H_s^* presented herein. The tests presented herein are, in some cases, easier to apply than Good's tests. In addition, the methods presented herein are applicable to a wider range of problems in the sense that Good's methods can be used to test the null hypothesis that all rth order and higher-order interactions are equal to zero, whereas the methods presented herein can be used to test the more general null hypothesis that any specified subset of these interactions is equal to zero.

The tests presented herein are generalizations of methods proposed earlier by Plackett [13] and Goodman [10] for testing the null hypothesis H_2^* in a three-dimensional table. The test proposed by Good [9] is a generalization of the methods proposed earlier by Bartlett [5], Roy and Kastenbaum [14], and Darroch [8] for testing H_2^* in the three-dimensional table. All of these earlier papers were concerned mainly with the testing of null hypotheses. In the present article, in addition to our treatment of hypothesis testing, we shall also present two different methods for obtaining confidence intervals for the rth order real interactions in the m-dimensional contingency table ($r = 0, 1, 2, \cdots, m - 1$).

2. The 2^m contingency table. We number the two classes in the kth dimension of the m-dimensional $2 \times 2 \times \cdots \times 2$ contingency table $i_k = 0$ and 1 (for

---

Received 29 November 1962; revised 6 January 1964.

[1] This research was supported in part by Research Grant No. NSF-G10368 from the Division of Mathematical, Physical and Engineering Sciences of the National Science Foundation. For very helpful comments, the author is indebted to I. J. Good.

$k = 1, 2, \cdots, m$). Let p_i denote the multinomial probability associated with cell $i = \{i_1, i_2, \cdots, i_m\}$ of the contingency table. We have $\sum_i p_i = 1$, and we assume that $p_i > 0$ for all i. Let b_i denote the natural logarithm of p_i; i.e., $b_i = \log p_i$. Let $L_m = 2^m$, and let b denote the $L_m \times 1$ column vector

$$b = \{b_i, \text{ with } i \text{ taking on all } L_m \text{ possible values}\}.$$

Let $|i|$ denote the number of non-zero i_k in i. (Thus, for the 2^m contingency table $|i| = \sum_{k=1}^m i_k$.) Let $j = \{j_1, j_2, \cdots, j_m\}$ where $j_k = 0$ or 1 for $k = 1, 2, \cdots, m$. Let $w(i, j) = (-1)^{i \cdot j}$, where $i \cdot j = \sum_{k=1}^m i_k j_k$. Let $w(\cdot, j)$ denote the $L_m \times 1$ column vector $w(\cdot, j) = \{w(i, j)$, with i taking on all L_m possible values$\}$, and let $I(j) = b' w(\cdot, j)$. Good [9] defined the rth order interactions to be the $I(j)$ for all j such that $|j| = r + 1$. The number of such rth order interactions is C_{r+1}^m.

We shall now consider the problem of testing the null hypothesis H_{m-1} within the alternate hypothesis H_m. (Following Good [9], for formal convenience we write H_m for the hypothesis that states nothing at all.) The null hypothesis specifies that $I_{m-1} = \sum_i (-1)^i b_i = 0$, where $(-1)^i$ denotes $(-1)^{|i|}$. If a sample of size n is drawn from the 2^m contingency table, the maximum likelihood estimator of I_{m-1} is $\hat{I}_{m-1} = \sum_i (-1)^i \hat{b}_i$, where $\hat{b}_i = \log (n_i/n)$, and n_i is the observed frequency in cell i. Since it is assumed that $p_i > 0$ for all i, the probability that the n_i are all positive will approach one as $n \to \infty$. However, if any observed n_i is actually zero, following Berkson [6] and Plackett [13] we might replace the n_i of zero by one-half when calculating the corresponding $\hat{b}_i$. Note that the factor n appearing in the formula for $\hat{b}_i$ can be ignored when the $\hat{b}_i$ are used to calculate $\hat{I}_{m-1}$. The asymptotic variance of $\hat{b}_i$ is $V(\hat{b}_i) = (p_i^{-1} - 1)n^{-1}$, and the asymptotic covariance between $\hat{b}_i$ and $\hat{b}_j$ $(i \neq j)$ is $C(\hat{b}_i, \hat{b}_j) = -n^{-1}$ (see, for example, Plackett [13]). Thus, the asymptotic variance of $\hat{I}_{m-1}$ is

$$(1) \quad V(\hat{I}_{m-1}) = \sum_i (p_i^{-1} - 1)n^{-1} - \sum_i (-1)^{2i}(-n^{-1}) = \sum_i (np_i)^{-1}.$$

Note that $\hat{I}_{m-1}$ is a contrast of the $\hat{b}_i$, and $V(\hat{I}_{m-1})$ can be calculated by regarding the $\hat{b}_i$ as uncorrelated and having a variance of $(np_i)^{-1}$. To estimate $V(\hat{I}_{m-1})$, we take

$$(2) \qquad \hat{V}_1(\hat{I}_{m-1}) = \sum_i n_i^{-1},$$

or

$$(3) \qquad \hat{V}_2(\hat{I}_{m-1}) = (n + 1)n^{-1} \sum_i (n_i + 1)^{-1}.$$

The estimators $\hat{V}_1(\hat{I}_{m-1})$ and $\hat{V}_2(\hat{I}_{m-1})$ are asymptotically equivalent. If any observed n_i is actually zero, the adjustment mentioned above (viz., the replacement of zero by one-half) might be applied in the calculation of $\hat{V}_1(\hat{I}_{m-1})$. For the calculation of $\hat{V}_2(\hat{I}_{m-1})$, adjustments of this kind will usually not be necessary; the bias of $\hat{V}_2(\hat{I}_{m-1})$ is $- n^{-1} \sum_i [(1 - p_i)^{n+1}/p_i]$, which will be negligible for large enough samples.

When H_{m-1} is true, the asymptotic distribution of $\hat{I}_{m-1}$ is normal with mean zero and variance $V(\hat{I}_{m-1})$ given above. For both $j = 1$ and 2, we find that $\hat{V}_j(\hat{I}_{m-1})/V(\hat{I}_{m-1})$ converges in probability to one as $n \to \infty$. Thus the asymptotic distribution of the statistic

$$X^2 = \hat{I}_{m-1}^2/\hat{V}_j(\hat{I}_{m-1}) \qquad (\text{for } j = 1 \text{ or } 2) \tag{4}$$

is the chi-square distribution with one degree of freedom when H_{m-1} is true. The statistic X^2 can be used to provide us with a simple test of the null hypothesis H_{m-1}.

The null hypothesis H_{m-1} which we have considered here is a natural generalization of the null hypothesis considered earlier by Bartlett [5], Roy and Kastenbaum [14], Darroch [8], Plackett [13], and Goodman [10] for the special case where $m = 3$. In this special case, the null hypothesis is that there is no second-order interaction, and the statistic presented herein [viz. (4) for $j = 1$ or 2] for testing this hypothesis is the following (for $j = 1$):

$$\begin{aligned} X^2 = \hat{I}_2^2/\hat{V}_1(\hat{I}_2) &= [\textstyle\sum_i (-1)^i \log n_i]^2/\textstyle\sum_i n_i^{-1} \\ &= \frac{[\log (n_{000}\, n_{011}\, n_{101}\, n_{110}/n_{001}\, n_{010}\, n_{100}\, n_{111})]^2}{[n_{000}^{-1} + n_{011}^{-1} + n_{101}^{-1} + n_{110}^{-1} + n_{001}^{-1} + n_{010}^{-1} + n_{100}^{-1} + n_{111}^{-1}]}. \end{aligned} \tag{5}$$

To test this hypothesis using the method originally proposed by Bartlett [5] and later by Roy and Kastenbaum [14], the user is required to solve a certain cubic equation in order to calculate the maximum likelihood estimator of p_i under H_2. To calculate the maximum likelihood estimator of p_i under H_2 using the iterative solution presented by Darroch [8], the number of (non-linear) equations to be solved iteratively is 12. Since the system of equations presented by Good [9] for determining the maximum likelihood estimator of p_i under H_2 corresponds in this special case to the system of equations presented by Bartlett [5], Roy and Kastenbaum [14], and Darroch [8], to apply Good's system of equations in order to test the null hypothesis H_2 the user will presumably be required to use either Bartlett's method or the iterative solution given by Darroch. Good [9] has noted that the required solution to Bartlett's cubic equation is known to lie in a certain range in advance, and is unique in that range, so that calculation of the solution can be simplified somewhat. In contrast with these earlier methods, we note that the user is not required to solve any equations at all to calculate the X^2 statistic (5) presented herein. With the aid of a table of logarithms and reciprocals, the user can readily apply the methods presented herein, and he will find them easier to apply in this case than the other methods referred to above.

For $m = 3$, the X^2 statistic presented here was given earlier by Woolf [19], Plackett [13], and Goodman [10]. The results presented herein for $m \geqq 3$ provide a generalization of this earlier work.

The system of equations derived by Good [9] for determining the maximum likelihood estimator of p_i under H_{m-1} is a generalization of the system of equations given by the earlier writers for the case where $m = 3$. In the more general case, to determine the maximum likelihood estimator of p_i under H_{m-1} by a generalization of the Bartlett method the user will be required to solve an equation of degree $2^{m-1} - 1$. (The calculation of the required solution of this equation is easier than the general solution of an equation of this degree, since the solution in this case is known to lie in a certain range and it is unique in this range (see [9]).) To calculate the maximum likelihood estimator of p_i under H_{m-1} by a generalization of the Darroch method, the number of equations to be solved iteratively will be $m2^{m-1}$. To apply Good's system of equations to test the null hypothesis H_{m-1}, the user will presumably be required to use either the generalization of Bartlett's method or the generalization of Darroch's method referred to above. We noted earlier herein that for $m = 3$ the test presented herein was easier to apply than Good's test. Similarly, for $m > 3$ the test of H_{m-1} based on the statistic X^2 presented herein will also be easier to apply than Good's test. (The user applying the X^2 statistic (4) is not required to solve any equations at all.) To determine which test is more accurately approximated by the tabular chi-square probabilities when the sample size n is small or moderate, further research is needed.

The null hypothesis H_{m-1} is also a natural generalization of the hypothesis H_1 of independence in a two-way contingency table ($m = 2$), and the hypothesis H_0 that $p_i = \frac{1}{2}$ in a one-way table ($m = 1$). For $m = 1$, the statistic presented herein for testing the null hypothesis H_0 is

$$X^2 = [\log (n_0/n_1)]^2 n_0 n_1/n, \tag{6}$$

and for $m = 2$ the statistic presented herein for testing the null hypothesis H_1 is

$$X^2 = [\log (n_{00}n_{11}/n_{01}n_{10})]^2/[n_{00}^{-1} + n_{11}^{-1} + n_{01}^{-1} + n_{10}^{-1}]. \tag{7}$$

By the usual large sample approximation for the logarithm of a maximum likelihood estimator, we find that for $m = 1$ the X^2 statistic (6) is asymptotically equivalent under H_0 to

$$Y^2 = (n_0 n_1^{-1} - 1)^2 n_0 n_1/n = [2n_0 - n]^2 n_0/nn_1, \tag{8}$$

which in turn is asymptotically equivalent under H_0 to

$$Z^2 = [n_0 - \tfrac{1}{2}n]^2 n/(n_0 n_1), \tag{9}$$

and to

$$W^2 = [n_0 - \tfrac{1}{2}n]^2 4/n. \tag{10}$$

For $m = 2$, the statistic (7) is asymptotically equivalent under H_1 to

$$\begin{aligned} Y^2 &= [n_{00}n_{11}/n_{01}n_{10} - 1]^2/[n_{00}^{-1} + n_{11}^{-1} + n_{01}^{-1} + n_{10}^{-1}] \\ &= [n_{00}n_{11} - n_{01}n_{10}]^2/\{[n_{00}^{-1} + n_{11}^{-1} + n_{01}^{-1} + n_{10}^{-1}](n_{01}n_{10})^2\}, \end{aligned} \tag{11}$$

which in turn is asymptotically equivalent under H_1 to

$$(12) \qquad Z^2 = [n_{00}n_{11} - n_{01}n_{10}]^2/[n_{01}n_{10}n_{00}n_{11}(n_{00}^{-1} + n_{11}^{-1} + n_{01}^{-1} + n_{10}^{-1})]$$

and to

$$(13) \qquad W^2 = (n_{00}n_{11} - n_{10}n_{01})^2 n/n_{0\cdot}n_{1\cdot}n_{\cdot 0}n_{\cdot 1},$$

where $n_{\cdot k} = n_{0k} + n_{1k}$ and $n_{k\cdot} = n_{k0} + n_{k1}$. For $m = 1$ the statistics (9) and (10) are the usual large-sample test statistics for testing the null hypothesis H_0; for $m = 2$ the statistic (13) is the usual large-sample test statistic for H_1. Thus, we have shown that the X^2 statistics presented herein are asymptotically equivalent under H_{m-1} (for $m = 1$ and 2) to the usual large-sample test statistics for these hypotheses.

For $m = 1$ and 2, the W^2 statistics (10) and (13) are the Lagrange-multiplier test statistics (see Aitchison and Silvey [3]). This is also the case for the goodness-of-fit statistic obtained by Bartlett [5], Roy and Kastenbaum [14], and Darroch [8] for $m = 3$, and more generally by Good [9] for any $m > 0$. In other words, the methods presented by the earlier writers, and the generalization given by Good, all involve restricted maximum-likelihood estimation of p_i under H_{m-1} with its associated Lagrange-multiplier test. (For some general results relating to the Lagrange-multiplier test, see Aitchison [1], Aitchison and Silvey [2], [3], and Silvey [16].) The method presented by Woolf [19], Plackett [13], and Goodman [10], and the generalization presented herein, all involve unrestricted maximum likelihood estimation with its associated Wald test. (See Wald [17], Aitchison and Silvey [2], [3], and Silvey [16].) The test presented herein involved the unrestricted maximum likelihood estimation of I_{m-1} and the Wald test was used to test the null hypothesis that $I_{m-1} = 0$. The proof of the asymptotic equivalence of the Lagrange-multiplier test and the corresponding Wald test (see literature cited above) can be applied to show that the test presented herein is asymptotically equivalent under H_{m-1} to the test given by Good [9].

We now consider the null hypothesis that a specified interaction, say $I(j)$, of rth order is equal to zero ($r < m$). The null hypothesis states that $I(j) = \sum_i (-1)^{i \cdot j} b_i = 0$. The maximum likelihood estimator of $I(j)$ is $\hat{I}(j) = \sum_i (-1)^{i \cdot j} \hat{b}_i$. Since $\hat{I}(j)$ is a contrast of the $\hat{b}_i$, by applying the asymptotic variance and covariance formulas given earlier herein we find that the asymptotic variance of $\hat{I}(j)$ is

$$(14) \qquad V[\hat{I}(j)] = \sum_i (np_i)^{-1}.$$

Note that $V[\hat{I}(j)]$ is equal to $V[\hat{I}_{m-1}]$ given earlier herein. Thus, the asymptotic variance is the same for all interactions. The estimators $\hat{V}_1(\hat{I}_{m-1})$ or $\hat{V}_2(\hat{I}_{m-1})$ given earlier herein can therefore be used to estimate $V[\hat{I}(j)]$. We let $\hat{V}$ denote either of these estimators. As in our earlier study of $\hat{I}_{m-1}$, we find that the asymptotic distribution of the statistic

$$(15) \qquad X^2 = [\hat{I}(j)]^2/\hat{V},$$

is the chi-square distribution with one degree of freedom. The statistic (15) can be used to provide us with a simple test of the null hypothesis that $I(j) = 0$. This test is a generalization of the test presented earlier in this section where only interaction of order $m - 1$ was considered.

We now consider the null hypothesis that two specified interactions, say $I(j)$ and $I(j^\dagger)$, of order r and $r^\dagger$, respectively, are equal to zero ($r < m$, $r^\dagger < m$, $j \neq j^\dagger$). The maximum likelihood estimator of $I(j)$ and $I(j^\dagger)$ is $\hat{I}(j) = \sum_i (-1)^{i \cdot j} \hat{b}_i$ and $\hat{I}(j^\dagger) = \sum_i (-1)^{i \cdot j^\dagger} \hat{b}_i$, respectively, the asymptotic variances of $\hat{I}(j)$ and $\hat{I}(j^\dagger)$ are $V[\hat{I}(j)] = V[\hat{I}(j^\dagger)] = \sum_i (np_i)^{-1}$, and the asymptotic covariance between $\hat{I}(j)$ and $\hat{I}(j^\dagger)$ is

$$C[\hat{I}(j), \hat{I}(j^\dagger)] = \sum_i (-1)^{i \cdot (j+j^\dagger)} \{(p_i^{-1} - 1)n^{-1} - (-n^{-1})\}$$
$$= \sum_i (-1)^{i \cdot (j+j^\dagger)} (np_i)^{-1}.$$

To estimate the variances, we take $\hat{V}$ as defined earlier herein; to estimate the covariance we take

$$\hat{C}_1[\hat{I}(j), \hat{I}(j^\dagger)] = \sum_i (-1)^{i \cdot (j+j^\dagger)} n_i^{-1}, \tag{17}$$

or

$$\hat{C}_2[\hat{I}(j), \hat{I}(j^\dagger)] = [(n+1)/n] \sum_i (-1)^{i \cdot (j+j^\dagger)} (n_i + 1)^{-1}. \tag{18}$$

The estimators $\hat{C}_1[\hat{I}(j), \hat{I}(j^\dagger)]$ and $\hat{C}_2[\hat{I}(j), \hat{I}(j^\dagger)]$ are asymptotically equivalent, and we let $\hat{C}[\hat{I}(j), \hat{I}(j^\dagger)]$ denote either of these estimators. Applying the usual large sample multivariate normal theory (see, for example, [4]), we test the null hypothesis that $I(j) = I(j^\dagger) = 0$ by computing the statistic

$$X^2 = \{[\hat{I}(j)]^2 + [\hat{I}(j^\dagger)]^2 - 2\rho \hat{I}(j)\hat{I}(j^\dagger)\}^2 / \{\hat{V}(1 - \rho^2)\}, \tag{19}$$

where $\rho = \hat{C}[\hat{I}(j), \hat{I}(j^\dagger)]/\hat{V}$. The statistic (19) provides us with a simple test of the null hypothesis, since its asymptotic distribution under the null hypothesis is the chi-square distribution with two degrees of freedom.

In the preceding paragraph, we gave simple formulas for the asymptotic variances and covariance of $\hat{I}(j)$ and $\hat{I}(j^\dagger)$, where $|j| = r + 1$ and $|j^\dagger| = r^\dagger + 1$. We noted that the asymptotic variance of $\hat{I}(j)$ is identical for all j regardless of the order r of $I(j)$, and that the estimated variance and covariance are easily computed. Having obtained the estimated variance and covariance, it is possible to apply the usual large sample multivariate normal theory (see, for example [4]) to test the null hypothesis that a specified set of K different interactions, say the interactions $I_{(1)}, I_{(2)}, \cdots, I_{(K)}$, are all equal to zero. Under the null hypothesis, the asymptotic distribution of the X^2 statistic obtained is the chi-square distribution with K degrees of freedom. More generally, if we let $H^\dagger$ denote the alternate hypothesis that a specified set of $K^\dagger$ different interactions are all equal to zero ($K^\dagger \geqq 0$), and if we let H denote the null hypothesis that a specified set of $K + K^\dagger$ interactions, including the $K^\dagger$ interactions in $H^\dagger$, are all equal to zero; then by a direct application of the usual multivariate theory we can test the null hypothesis H within $H^\dagger$. The X^2 statistic obtained will have an asymptotic chi-square distribution with K degrees of freedom under H. The hypotheses H and $H^\dagger$ are generalizations of the hypotheses H_r and H_s $(r < s)$ considered by Good [9], and the X^2 statistic given here provides us with a test of the generalized hypothesis. In the special case where $H = H_r$ and $H^\dagger = H_s$, the number of degrees of freedom will be $\sum_{k=r}^{s-1} C_{k+1}^m$. The test obtained here is quite different from the test proposed by Good [9], though the two tests are asymptotically equivalent under H_r. We noted earlier herein that for $r = m - 1$ and $s = m$, the test given here is preferable for computational purposes. It is easy to see that when $r = 0$ and $s = m$ the test given by Good [9] will be preferable.

The preceding discussion was concerned solely with hypothesis testing. We shall now present a simple method for obtaining confidence intervals for the interactions $I(j)$. Since $\hat{I}(j)$ is asymptotically normal with mean $I(j)$ and variance V given by (14), we have the following approximate two-sided confidence interval for $I(j)$, at the level of probability $1 - \alpha$:

$$\hat{I}(j) \pm \phi(1 - \alpha)\hat{V}^{\frac{1}{2}}, \tag{20}$$

where $\phi(1 - \alpha)$ is the $[(1 - \frac{1}{2}\alpha) \times 100]$th percentile of the unit normal variate. For any specified set of K interactions, say $I_{(1)}, I_{(2)}, \cdots, I_{(K)}$, we then have the following approximate two-sided simultaneous confidence intervals, at a probability level which is at least $1 - \alpha$:

$$\hat{I}_{(k)} \pm \phi_K(1 - \alpha)\hat{V}^{\frac{1}{2}} \qquad (k = 1, 2, \cdots, K), \tag{21}$$

where $\phi_K(1 - \alpha)$ is the $[(1 - \alpha/2K) \times 100]$th percentile of the unit normal variate. (For related results, see, for example, Wilks [18], p. 291.) The following approximate two-sided simultaneous confidence intervals are also at a probability level of at least $1 - \alpha$:

$$\hat{I}_{(k)} \pm \chi_K(1 - \alpha)\hat{V}^{\frac{1}{2}} \qquad (k = 1, 2, \cdots, K), \tag{22}$$

where $\chi_K^2(1 - \alpha)$ is the $[(1 - \alpha) \times 100]$th percentile of the chi-square distribution with K degrees of freedom. The confidence intervals (22) can be derived by a method similar to that used by Scheffé [15] to derive his simultaneous confidence intervals. For $K = 1$, the confidence intervals (21) and (22) are identical; but for $K > 1$, the confidence intervals (21) are shorter than the corresponding (22), for the usual probability levels ($\alpha = .05$ or $.01$).

If any of the K confidence intervals (22) exclude the value zero, then it can be shown that the chi-square statistic X^2 (with K degrees of freedom), presented earlier herein for testing the null hypothesis that the K interactions $I_{(k)}$ $(k = 1, 2, \cdots, K)$ are equal to zero, will be larger than $\chi_K^2(1 - \alpha)$, and thus the test based upon this statistic will lead to rejection of the null hypothesis at the level of significance α. The statistic X^2 will be less than or equal to $\chi_K^2(1 - \alpha)$ if and only if all the confidence intervals (22) include zero, and the corresponding confidence intervals for all linear functions of the $I_{(k)}$ $(k = 1, 2, \cdots, K)$ also include zero. (For somewhat related results, see Scheffé [15].) The simultaneous confidence intervals referred to above, for all linear functions of the $I_{(k)}$ $(k = 1, 2, \cdots, K)$, are computed as in (22), using the constant $\chi_K(1 - \alpha)$. Since linear functions of the $I_{(k)}$ are also contrasts of the $\hat{b}_i$, to determine the estimated variances to be used in the calculation of the simultaneous confidence intervals, we can regard the $\hat{b}_i$ as uncorrelated with variance $(np_i)^{-1}$, as was the case earlier herein. If a specified set of K^* linear functions of the $I_{(k)}$ are of interest, in order to obtain simultaneous confidence intervals for these functions, the square roots of the estimated variances are multiplied by the constant $\chi_{K^*}(1 - \alpha)$ or by $\chi_K(1 - \alpha)$ as in (22), or by ϕ_{K^*} $(1 - \alpha)$ as in (21), whichever is smaller.

3. Interactions in the m-dimensional $d_1 \times d_2 \times \cdots \times d_m$ contingency table. We number the d_k classes in the kth dimension of the m-dimensional $d_1 \times d_2 \times \cdots \times d_m$ contingency table from $i_k = 0$ to $d_k - 1$ $(k = 1, 2, \cdots, m)$. As in the preceding section, we let p_i denote the multinomial probability associated with cell $i = \{i_1, i_2, \cdots, i_m\}$ of the contingency table, and we assume that $p_i > 0$ for all i. Let $\log p_i = b_i$ and $L_m = \prod_{k=1}^m d_k$, and let b denote the $L_m \times 1$ column vector $b = \{b_i$, with i taking on all L_m possible values$\}$. Let $j = \{j_1, j_2, \cdots, j_m\}$, where the values of j_k run from $j_k = 0$ to $d_k - 1$ for $k = 1, 2, \cdots, m$. Let $w(i, j) = \prod_{k=1}^m w_k^{i_k j_k}$, where w_k is the primitive root of unity, $w_k = \exp(2\pi\sqrt{-1}/d_k)$. Let $w(\cdot, j)$ denote the $L_m \times 1$ column vector $w(\cdot, j) = \{w(i, j)$, with i taking on all L_m possible values$\}$, and let $I(j) = b'w(\cdot, j)$. Good [9] defined the rth order complex interactions to be the $I(j)$ for all j such that $|j| = r + 1$. The number K_r of these rth order complex interactions is the sum of the products of $d_1 - 1, d_2 - 1, \cdots, d_m - 1$ taken $(r + 1)$ at a time. We shall now define a different set of K_r real interactions of rth order.

Let $i = \{i_1, i_2, \cdots, i_m\}$ and $j = \{j_1, j_2, \cdots, j_m\}$. We define $f(i, j) = 0$ unless, for each k, either $i_k = j_k$ or $i_k = 0$ or $j_k = 0$, and then $f(i, j) = (-1)^\nu$, where ν is the number of k for which $i_k j_k \neq 0$. Let $f(\cdot, j)$ denote the $L_m \times 1$ column vector $f(\cdot, j) = \{f(i, j)$, with i taking on all L_m possible values$\}$, and let $I^*(j) = b'f(\cdot, j)$. We define the rth order real interactions to be the $I^*(j)$ for all j such that $|j| = r + 1$. There are K_r real interactions of rth order.

In the special case where $d_k = 2$ for $k = 1, 2, \cdots, m$, we have $I(j) = I^*(j)$; but the real interactions given here are quite different from the complex interactions defined by Good [9] when $d_k > 2$ for some k. In the case where $d_k > 2$ for some k, the real interaction $I^*(j)$ of order $r < m$ is a sum of interactions defined for the $2 \times 2 \times \cdots \times 2$ subtables of size 2^{r+1} of the m-dimensional $d_1 \times d_2 \times \cdots \times d_m$ contingency table. For example, for $r = m - 1$ let $S(j)$ denote the subtable of size 2^m formed by restricting the indices of the m-dimensional $d_1 \times d_2 \times \cdots \times d_m$ table to $(0, j_1; 0, j_2; \cdots; 0, j_m)$, where $|j| = m$. Let I_{m-1} denote the interaction of order $m - 1$ in a table of size 2^m, and let $I_{m-1}(j)$ denote the interaction I_{m-1} calculated for the subtable $S(j)$. We can write $I_{m-1}(j)$ in the following terms: Define $g(i, j) = 0$ unless, for each k either $i_k = j_k$ or $i_k = 0$, and then $g(i, j) = (-1)^\nu$, where ν is the number of k for which $i_k j_k \neq 0$. Let $g(\cdot, j)$ denote the $L_m \times 1$ column vector $g(\cdot, j) = \{g(i, j)$, with i taking on all L_m possible values of $i\}$. Then $I_{m-1}(j) = b'g(\cdot, j) = I^*(j)$. Thus, for $|j| = m$ the interaction $I^*(j)$ is the interaction of order $m - 1$ calculated for $S(j)$. For $m > 1$, note that $I_{m-1}(j) = I_{m-2}(j^\dagger \mid 0) - I_{m-2}(j^\dagger \mid j_m)$, where $j^\dagger = \{j_1, j_2, \cdots, j_{m-1}\}$, and $I_{m-2}(j^\dagger \mid k)$ is the interaction of order $m - 2$ calculated for the subtable $S(j^\dagger \mid k)$ of size 2^{m-1} formed by restricting the indices of the m-dimensional table to $(0, j_1; 0, j_2; \cdots; 0, j_{m-1}; k)$. Thus the real interactions of order $m - 1$ can be obtained from the interactions of order $m - 2$ calculated for the subtables $S(j^\dagger \mid 0)$ and $S(j^\dagger \mid j_m)$ of size 2^{m-1}, when $m > 1$.

The results presented above can now be extended to cover the case where the interactions are of order $r < m - 1$. For the sake of simplicity, let us first consider $j = \{j_1, j_2, \cdots, j_m\}$ where $j_k > 0$ for $k = 1, 2, \cdots, r + 1$, and $j_k = 0$ otherwise. Let $j' = \{j'_{r+2}, j'_{r+3}, \cdots, j'_m\}$ and let $S(j \mid j')$ denote the subtable of size 2^{r+1} formed by restricting the indices of the m-dimensional table to $(0, j_1; 0, j_2; \cdots; 0, j_{r+1}; j'_{r+2}; j'_{r+3}; \cdots; j'_m)$. Let I_r denote the interaction of order r in a table of size 2^{r+1}, and let $I_r(j \mid j')$ denote the real interaction I_r calculated for the subtable $S(j \mid j')$ of size 2^{r+1}. We can write $I_r(j \mid j')$ in the following terms: Define $g(i, j \mid j') = 0$ unless, for $k = 1, 2, \cdots, r + 1$ we have either $i_k = j_k$ or $i_k = 0$, and for $k = r + 2, r + 3, \cdots, m$ we have $i_k = j'_k$, and then $g(i, j \mid j') = (-1)^\nu$ where ν is the number of k for which $i_k j_k \neq 0$. Let $g(\cdot, j \mid j')$ denote the $L_m \times 1$ column vector $g(\cdot, j \mid j') = \{g(i, j \mid j')$, with i taking on all L_m possible values$\}$. Then $I_r(j \mid j') = b'g(\cdot, j \mid j')$. The real interaction $I^*(j)$ is easily seen to be $\sum_{j'} I_r(j \mid j')$, where the summation is over all $d_{r+2} \times d_{r+3} \times \cdots \times d_m$ possible values of j'. For $r > 1$ note that $I_r(j \mid j') = I_{r-1}(j^\dagger \mid 0, j') - I_{r-1}(j^\dagger \mid j_{r+1}, j')$, where $j^\dagger = \{j_1, j_2, \cdots, j_r\}$, and $I_{r-1}(j^\dagger \mid k, j')$ is the interaction of order $r - 1$ calculated for the subtable $S(j^\dagger \mid k, j')$ of size 2^r formed by restricting the indices of the m-dimensional table to $(0, j_1; 0, j_2; \cdots; 0, j_r; k; j'_{r+2}; j'_{r+3}; \cdots, j'_m)$. Thus, $I^*(j) = \sum_{j'} I_{r-1}(j^\dagger \mid 0, j') - \sum_{j'} I_{r-1}(j^\dagger \mid j_{r+1}, j')$, which indicates that the real interaction of order r can be obtained from the interactions of order $r - 1$ calculated for the subtables $S(j^\dagger \mid 0, j')$ and $S(j^\dagger \mid j_{r+1}, j')$ of size 2^r, when $r > 1$. Similar results can be obtained for any j such that $|j| = r + 1$.

Again for simplicity let us take $j = \{j_1, j_2, \cdots, j_m\}$ where $j_k > 0$ for $k = 1, 2, \cdots, r + 1$, and $j_k = 0$ otherwise. Let $\prod_{k=1}^{r+1} (d_k - 1) = D_r$. There are D_r values of j. For each of the D_r values of j, we can calculate $I^*(j)$, and all linear functions of the $I^*(j)$ can be written in the form

$$I^*(\alpha) = \sum_i \alpha_{i_1 i_2 \cdots i_{r+1}} b_i, \tag{23}$$

where the summation is over all L_m possible values of $i = \{i_1, i_2, \cdots, i_m\}$, and where for each k ($k = 1, 2, \cdots, r + 1$) the following conditions will be satisfied:

$$\sum_{i_k} \alpha_{i_1 i_2 \cdots i_{r+1}} = 0, \tag{24}$$

this summation being over the d_k possible values of i_k. Conversely, all functions of the form $I^*(\alpha)$, where the $\alpha_{i_1 i_2 \cdots i_{r+1}}$ satisfy (24), will be linear functions of the $I^*(j)$. In particular, if we take $\alpha_{i_1 i_2 \cdots i_{r+1}} = f(i, j)$, we see in this case that $I^*(\alpha) = I^*(j)$. More generally, if for each possible value of i the constant α_i is of the form $\alpha_i = \alpha_{i_1 i_2 \cdots i_{r+1}}$ where the α_i satisfy (24), then $\alpha = \{\alpha_i$, with i taking on all L_m possible values of $i\}$ represents a point in a D_r dimensional vector space spanned by the D_r vectors $f(\cdot, j) = \{f(i, j)$, with i taking on all L_m possible values of $i\}$, and $I^*(\alpha) = \sum_i \alpha_i b_i$ is a linear function of the $I^*(j)$.

Returning now to Good's definition of interaction given earlier herein, note that for each k $(k = 1, 2, \cdots, r + 1)$ the following conditions will be satisfied:

$$\sum_{i_k} w(i, j) = 0, \tag{25}$$

this summation being over the d_k possible values of i_k. (Recall that we are considering the case where $j_k > 0$ for $k = 1, 2, \cdots, r + 1$, and $j_k = 0$ otherwise.) Thus, the $w(i, j)$ satisfy (24) for both the real parts and for the complex parts of $w(i, j)$. Considering the real and complex parts of $w(\cdot, j)$ as two real vectors, we find therefore that these vectors represent points in the vector space spanned by the D_r vectors $f(\cdot, j)$. Furthermore, it is possible to show that this space is also spanned by the D_r vectors $w(\cdot, j)$, considering the real and complex parts of $w(\cdot, j)$ as two real vectors. (The D_r complex vectors $w(\cdot, j)$ yield $2D_r$ real vectors, but only D_r independent real vectors.) Thus, the condition that $I(j) = b'w(\cdot, j) = 0$ for the D_r values of j is equivalent to the condition that $I^*(j) = b'f(\cdot, j) = 0$ for the D_r values of j.

The results presented above were concerned with the case where the first $r + 1$ entries in the vector j were positive and the rest were zero. These results can be generalized in a straightforward fashion to cover the case where a specified set of $r + 1$ entries in the vector j are positive and the remaining $m - r - 1$ entries are zero. Thus, suppose that the $r + 1$ entries $j_{t_1}, j_{t_2}, \cdots, j_{t_{r+1}}$ are positive, and the remaining $m - r - 1$ entries in j are zero. Let $\prod_{k=1}^{r+1} (d_{t_k} - 1) = D(t)$, where $t = \{t_1, t_2, \cdots, t_{r+1}\}$. There are $D(t)$ possible values of j. From the results presented above we see that the condition that $I(j) = 0$ for all $D(t)$ values of j is equivalent to the condition that $I^*(j) = 0$ for all $D(t)$ values of j. Thus, the hypothesis H_r considered by Good [9] is equivalent to the hypothesis H_r^* that all rth order and higher-order real interactions are equal to zero; i.e., that $I^*(j) = 0$ for all j such that $|j| \geqq r + 1$. To test the hypothesis H_r we shall test instead the equivalent hypothesis H_r^*.

4. Statistical methods for the m-dimensional table. We first consider the problem of testing the null hypothesis that a specified real interaction, say $I^*(j)$, of rth order is equal to zero $(r < m)$. The maximum likelihood estimator of $I^*(j)$ is $\hat{I}^*(j) = \hat{b}'f(\cdot, j)$, where $\hat{b}$ is the maximum likelihood estimator of b. Since $\hat{I}^*(j)$ is a contrast of the $\hat{b}_i$, by applying the asymptotic variance and covariance formulas given earlier herein we find that the asymptotic variance of $\hat{I}^*(j)$ is

$$V[\hat{I}^*(j)] = \sum_i c(i, j)(np_i)^{-1}, \tag{26}$$

where $c(i, j) = 0$ when $f(i, j) = 0$, and $c(i, j) = 1$ when $f(i, j) = \pm 1$. This variance can be estimated by replacing $(np_i)^{-1}$ in the above formula by n_i^{-1} or by $(n + 1)/[n(n_i + 1)]$, as earlier herein. We denote the estimated variance by $\hat{V}[\hat{I}^*(j)]$. The statistic

$$X^2 = [\hat{I}^*(j)]^2/\hat{V}[\hat{I}^*(j)] \tag{27}$$

will have an asymptotic chi-square distribution with one degree of freedom when the null hypothesis is true, and it can be used to test the null hypothesis.

The covariance between two estimated interactions, say $\hat{I}^*(j)$ and $\hat{I}^*(j^\dagger)$, of order r and $r^\dagger$, respectively, is

$$C[\hat{I}^*(j), \hat{I}^*(j^\dagger)] = \sum_i f(i, j)f(i, j^\dagger)(np_i)^{-1}, \tag{28}$$

which can be estimated by $\hat{C}[\hat{I}^*(j), \hat{I}^*(j^\dagger)]$, which is obtained by replacing $(np_i)^{-1}$ in the above formula by n_i^{-1} or by $(n + 1)/[n(n_i + 1)]$, as earlier herein. Having obtained the estimated variances and covariances, it is possible to apply, as in the preceding section, the usual large sample multivariate theory to test the null hypothesis that a specified set of K different interactions are all equal to zero, or more generally to test the null hypothesis H that a specified set of $K + K^\dagger$ different interactions ($K^\dagger \geqq 0$) are all equal to zero within the alternate hypothesis $H^\dagger$ that a specified subset consisting of $K^\dagger$ of these interactions is zero. The X^2 statistic used to test H will have an asymptotic chi-square distribution with K degrees of freedom when H is true. Since the hypothesis H and $H^\dagger$ are generalizations of the hypotheses H_r^* and H_s^*, respectively, the X^2 given here can also be used to test these hypotheses, which are equivalent to the hypotheses considered by Good [9]. As earlier, the test obtained by the method suggested here is quite different from Good's test. To illustrate this difference, we shall now consider in somewhat more detail the tests for the null hypothesis H_{m-1} within the alternate hypothesis H_m for the m-dimensional $d_1 \times d_2 \times \cdots \times d_m$ table. (Note that $K^\dagger = 0$ when the alternate hypothesis is H_m.)

There are $D_{m-1} = \prod_{k=1}^{m} (d_k - 1)$ real interactions $I^*(j)$ of order $m - 1$. For $m > 1$ we noted earlier herein that the real interaction $I^*(j)$ of order $m - 1$ can be obtained from the interactions of order $m - 2$ calculated for the subtables $S(j^\dagger \mid 0)$ and $S(j^\dagger \mid j_m)$ of size 2^{m-1}, where $j^\dagger = \{j_1, j_2, \cdots, j_{m-1}\}$ and $j = \{j^\dagger, j_m\}$. In particular, $I^*(j) = I_{m-2}(j^\dagger \mid 0) - I_{m-2}(j^\dagger \mid j_m)$. Therefore, the null hypothesis that $I^*(j) = 0$ is equivalent to the hypothesis that $I_{m-2}(j^\dagger \mid 0) = I_{m-2}(j^\dagger \mid j_m)$. Writing $j = \{j^\dagger, j_m\}$, we see more generally that the null hypothesis that the $d_m - 1$ interactions $I^*(j^\dagger, 1), I^*(j^\dagger, 2), \cdots, I^*(j^\dagger, d_m - 1)$ of order $m - 1$ are all equal to zero (for a given vector $j^\dagger$) is equivalent to the hypothesis that the d_m interactions $I_{m-2}(j^\dagger \mid 0), I_{m-2}(j^\dagger \mid 1), \cdots, I_{m-2}(j^\dagger \mid d_m - 1)$ of order

$m - 2$ are all equal to each other. Denoting the maximum likelihood estimator of $I_{m-2}(j^\dagger \mid k)$ by $\hat{I}_{m-2}(j^\dagger \mid k)$, the d_m statistics $\hat{I}_{m-2}(j^\dagger \mid k)$ (for $k = 0, 1, \cdots, d_m - 1$) are contrasts of mutually exclusive sets of $\hat{b}_i$. By applying the asymptotic covariance formulas given earlier herein, we find that the $\hat{I}_{m-2}(j^\dagger \mid k)$ and $\hat{I}_{m-2}(j^\dagger \mid k')$ are asymptotically uncorrelated for $k \neq k'$. (Similarly, $\hat{I}_{m-2}(j^\dagger \mid k)$ and $\hat{I}_{m-2}(j^* \mid k')$ are asymptotically uncorrelated for $k \neq k'$, when $j^* = \{j_1^*, j_2^*, \cdots, j_{m-1}^*\}$ and $|j^*| = m - 1$.) Thus, to test the hypothesis that the $d_m - 1$ interactions $I^*(j^\dagger, k)$ of order $m - 1$ (for $k = 1, 2, \cdots, d_m - 1$) are equal to zero (with $j^\dagger$ given), we can apply the usual chi-square test to determine whether the d_m statistics $\hat{I}_{m-2}(j^\dagger \mid k)$ differ significantly from each other ($k = 0, 1, \cdots, d_m - 1$). To compute the usual chi-square statistic (with $d_m - 1$ degrees of freedom), we note that the estimated variance of $\hat{I}_{m-2}(j^\dagger \mid k)$ is calculated readily by applying the results presented in the preceding section since $\hat{I}_{m-2}(j^\dagger \mid k)$ is the estimated interaction of order $m - 2$ in the subtable $S(j^\dagger \mid k)$ of size 2^{m-1}.

Let $D_{m-2} = \prod_{k=1}^{m-1} (d_k - 1)$. There are D_{m-2} possible values of $j^\dagger = \{j_1, j_2, \cdots, j_{m-1}\}$ such that $|j^\dagger| = m - 1$. To test the null hypothesis that the $D_{m-2}(d_m - 1)$ interactions $I^*(j^\dagger, 1), I^*(j^\dagger, 2), \cdots, I^*(j^\dagger, d_m - 1)$ are all equal to zero (for all D_{m-2} possible values of $j^\dagger$), we can apply the usual multivariate test to determine whether the d_m vectors $\hat{I}_{m-2}(\cdot, 0)$, $\hat{I}_{m-2}(\cdot, 1)$, $\cdots$, $\hat{I}_{m-2}(\cdot, d_m - 1)$ differ significantly from each other, where $\hat{I}_{m-2}(\cdot, k) = \{\hat{I}_{m-2}(j^\dagger, k)$, with $j^\dagger$ taking on all D_{m-2} possible values$\}$. The chi-square test thus obtained will have $D_{m-1} = \prod_{k=1}^{m} (d_k - 1)$ degrees of freedom. To calculate this chi-square test, estimates of the variances and covariances of the $\hat{I}_{m-2}(j^\dagger, k)$ are required. Noting that $\hat{I}_{m-2}(j^\dagger, k)$ is a contrast between the $\hat{b}_i$, we find by a straightforward generalization of the results presented earlier herein that the variances and covariances of the $\hat{I}_{m-2}(j^\dagger, k)$ can be calculated by regarding the $\hat{b}_i$ as uncorrelated and having a variance of $(np_i)^{-1}$ (see Plackett [13], and Goodman [10]). Estimates of the variances and covariances of the $\hat{I}_{m-2}(j^\dagger, k)$ can then be calculated by replacing the $(np_i)^{-1}$ in the formulae by one of their consistent estimators, as earlier herein.

From the results presented above, we see that to test H_{m-1} we can apply the usual large sample multivariate theory to determine whether all D_{m-1} interactions $I^*(j)$ with $|j| = m$ are equal to zero, which would require that the user invert one estimated variance-covariance matrix of side D_{m-1}; or equivalently we can apply the usual large sample theory to determine whether all d_m vectors $\hat{I}(\cdot, k)$ for $k = 0, 1, \cdots, d_m - 1$ differ significantly from each other, which would require that the user invert $d_m + 1$ matrices each of side D_{m-2}. Thus, in the special case of an m-dimensional contingency table with $d_k = 2$ for $k = 1, 2, \cdots, m - 1$, and $d_m \geqq 2$, the method given above for testing the null hypothesis H_{m-1} is as simple as the usual large sample test of whether d_m independent sample means differ significantly from each other. (In this special case, the user need not invert any matrices at all.) This generalizes the earlier results by Plackett [13] and Goodman [10] for the case where $m = 3$.

The case where $m = 3$ and $d_1 = d_2 = 2$ has also been studied by Norton [12] and Kastenbaum and Lamphiear [11]. To apply their methods for testing H_2, the user is required to solve a set of $d_3 - 1$ simultaneous third-degree equations in as many unknowns. To apply the iterative solution given by Darroch [8] in this case, the user would be required to solve iteratively $4(d_3 + 1)$ non-linear equations. Clearly the method of analysis presented herein is simpler to apply than these earlier methods.

For the general three-way table, the user applying the method suggested herein to test the null hypothesis H_2 will be required to invert $d_3 + 1$ matrices each of side $(d_1 - 1)(d_2 - 1)$ [or if he prefers he can invert only one matrix of side $(d_1 - 1)(d_2 - 1)(d_3 - 1)$]. It is also possible to make still further simplifications so that the user would be required to invert only one matrix of side $(d_1 - 1)(d_2 - 1)$ and d_3 matrices each of side $c - 1$, where $c = \min [d_1, d_2]$ (see Goodman [10]). To apply the methods of Roy and Kastenbaum [14] for testing H_2, the user is required to solve a set of $(d_1 - 1)(d_2 - 1)(d_3 - 1)$ simultaneous fourth-degree equations in as many unknowns. When $\min [d_1, d_2, d_3] = 2$, these equations reduce to third-degree equations. To apply the iterative solution given by Darroch [8], the user is required to solve iteratively $d_1d_2d_3 \sum_{k=1}^{3} d_k^{-1}$ nonlinear equations. (See Good's related remark on iterative scaling methods, [9] p. 927.) Again we find that the method presented herein is simpler to apply than the earlier methods presented in the literature.

To test H_{m-1} using the system of equations given by Good [9], for the general m-dimensional table $(m > 2)$, applying a generalization of the Darroch method, the number of equations to be solved iteratively is $\prod_{k=1}^{m} d_k \sum_{h=1}^{m} d_h^{-1}$. Applying a generalization of the Roy-Kastenbaum method, the user would be required to solve a set of $\prod_{k=1}^{m} (d_k - 1)$ simultaneous equations of degree 2^{m-1}. (When $\min [d_1, d_2, \cdots, d_m] = 2$, the degree of these equations is reduced by one.) To test H_{m-1} using the method presented herein, there will be $d_m + 1$ matrices each of side $\prod_{k=1}^{m-1} (d_k - 1)$ to invert, or alternatively one matrix of side $\prod_{k=1}^{m} (d_k - 1)$ to invert. Thus, for the more general m-dimensional table $(m > 2)$, we also find that the test of H_{m-1} presented herein is simpler to apply than Good's test of this hypothesis.

The results presented above concerning interactions of order $m - 1$ can be extended to the case where interactions of order $r \leqq m - 1$ are of interest. Let us now consider the interactions $I^*(j)$ of order r where $j_k > 0$ for $k = 1, 2, \cdots, r + 1$, and $j_k = 0$ otherwise. Let $S^*(\cdot \mid j')$ denote the $(r + 1)$-dimensional complete subtable formed by not restricting the first $r + 1$ indices $\{i_1, i_2, \cdots, i_{r+1}\}$ at all, and then restricting the remaining $m - r - 1$ indices $\{i_{r+2}, i_{r+3}, \cdots, i_m\}$ to be equal to $j' = \{j'_{r+2}, j'_{r+3}, \cdots, j'_m\}$. From the results presented earlier herein, we see that the interaction $I^*(j)$ is a sum of $d_{r+2} \times d_{r+3} \times \cdots \times d_m$ interactions $I_r(j \mid j')$, where $I_r(j \mid j')$ is an interaction of order r calculated for the $(r + 1)$-dimensional subtable $S^*(\cdot \mid j')$. For a given value of j', there are $D_r = \prod_{k=1}^{r+1} (d_k - 1)$ interactions $I_r(j \mid j')$ of order r, and the results presented herein for testing H_{m-1} in an m-dimensional table can be applied directly to test the null hypothesis $H_r(j')$ that all D_r interactions $I_r(j \mid j')$ in the $(r + 1)$-

dimensional table $S^*(\cdot \mid j')$ are equal to zero. For different values of j', the $S^*(\cdot \mid j')$ are mutually exclusive subtables. Denoting the estimator of $I_r(j \mid j')$ by $\hat{I}_r(j \mid j')$, for different values of j' the $\hat{I}_r(j \mid j')$ are asymptotically independent. Since $I^*(j) = \sum_{j'} I_r(j \mid j')$, the test of $H_r(j')$ can now be modified in a straightforward fashion to provide a test of the hypothesis that all D_r interactions $I^*(j)$ are equal to zero.

The hypothesis that all D_r interactions $I^*(j)$ are zero is the hypothesis that the rth order interactions between the first $r + 1$ dimensions $[1, 2, \cdots, r + 1]$, summed over the remaining $m - r - 1$ dimensions, are equal to zero. For any set of $r + 1$ dimensions, say $[\delta_1, \delta_2, \cdots, \delta_{r+1}]$, we can in a fashion similar to that described above test the hypothesis $H[\delta]$ that all rth order interactions between the dimensions $\{\delta_1, \delta_2, \cdots, \delta_{r+1}\} = \delta$, summed over the remaining $m - r - 1$ dimensions, are equal to zero.

The results presented earlier in the present section were concerned solely with hypothesis testing. As in Section 2 herein, it is also possible to obtain simultaneous confidence intervals for the interactions $I^*(j)$ in a multidimensional table. The simultaneous confidence intervals for the 2^m table given in Section 2 can be generalized in a straightforward fashion to obtain confidence intervals for the $d_1 \times d_2 \times \cdots \times d_m$ contingency table. To save space we shall not present the details here.

In closing, we take note of the fact that the results presented herein pertain to interactions defined in terms of certain linear functions of the b_i. Similar results could have been obtained if, for example, instead of the definition of the D_r interactions $I^*(j)$ of order r given herein (where for simplicity we took $j_k > 0$ for $k = 1, 2, \cdots, r + 1$, and $j_k = 0$ otherwise), we had taken as our definition any given non-singular transformation of these interactions, or if we had taken these interactions as a basis for a more general (and more symmetric) definition of the interactions given by the linear functions (23) satisfying the condition (24). For interactions $I^*(j)$ of order r, where the $r + 1$ entries $j_{t_1}, j_{t_2}, \cdots, j_{t_{r+1}}$ are positive, and the remaining $m - r - 1$ entries in j are zero, similar kinds of modifications can be made.

A somewhat different formulation of the concept of interaction, obtained by writing the b_i in terms of a general mean plus "interactions" of order r ($r = 0, 1, \cdots, m - 1$), in a manner similar to the usual analysis of variance model (see, for example, Birch [7]), would probably have seemed more familiar to analysis of variance users. The hypothesis H_r considered by Good [9] and the hypothesis H_r^* considered herein are equivalent to the hypothesis that all rth order and higher-order "interactions" in the analysis of variance model are equal to zero, but there are a number of reasons why the usual analysis of variance techniques can not be applied directly to test this hypothesis: (i) The assumption of constant variance which is made in the usual analysis of variance cannot be made in the present context where the variance of the $\hat{b}_i$ depends upon the value of p_i. (ii) Since $\sum_i \exp[b_i] = 1$ in the present context, the general mean is subject here to a restraint which does not appear in the usual analysis of variance model. (iii) The assumption that the observations are independent, which is made in the analysis of variance, cannot be made here since the $\hat{b}_i$ are correlated. (However, when dealing with contrasts of the b_i, we can regard the $\hat{b}_i$ as asymptotically independent, as earlier herein.) It is possible to modify some of the usual analysis of variance techniques so that they can be applied in the present context, but we shall not go into these details here.

REFERENCES

[1] Aitchison, John (1962). Large sample restricted parametric tests. *J. Roy. Statist. Soc. Ser. B* **24** 234–250.

[2] Aitchison, J. and Silvey, S. D. (1958). Maximum likelihood estimation of parameters subject to restraints. *Ann. Math. Statist.* **29** 813–828.

[3] Aitchison, J. and Silvey, S. D. (1960). Maximum-likelihood estimation procedures and associated tests of significance. *J. Roy. Statist. Soc. Ser. B* **23** 154–171.

[4] Anderson, T. W. (1958). *Introduction to Multivariate Statistical Analysis*. Wiley, New York.

[5] Bartlett, M. S. (1935). Contingency table interactions. *J. Roy. Statist. Soc. Suppl.* **2** 248–252.

[6] Berkson, J. (1955). Maximum likelihood and minimum chi-square estimates of the logistic function. *J. Amer. Statist. Assoc.* **50** 130–162.

[7] Birch, M. W. (1963). Maximum likelihood in three-way contingency tables. *J. Roy. Statist. Soc. Ser. B* **25** 220–233.

[8] Darroch, J. N. (1962). Interactions in multifactor contingency tables. *J. Roy. Statist. Soc. Ser. B* **24** 251–263.

[9] Good, I. J. (1963). Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables. *Ann. Math. Statist.* **34** 911–934.

[10] Goodman, L. A. (1963). On Plackett's test for contingency table interactions. *J. Roy. Statist. Soc. Ser. B* **25** 179–188.

[11] Kastenbaum, M. A. and Lamphiear, D. E. (1959). Calculation of chi-square to test the no three-factor interaction hypothesis. *Biometrics* **15** 107–115.

[12] Norton, H. W. (1945). Calculation of chi-square from complex contingency tables. *J. Amer. Statist. Assoc.* **40** 251–259.

[13] Plackett, R. L. (1962). A note on interactions in contingency tables. *J. Roy. Statist. Soc. Ser. B* **24** 162–166.

[14] Roy, S. N. and Kastenbaum, M. A. (1956). On the hypothesis of no "interaction" in a multi-way contingency table. *Ann. Math. Statist.* **27** 749–757.

[15] Scheffé, H. (1959). *The Analysis of Variance*. Wiley, New York.

[16] Silvey, S. D. (1959). The Lagrange-multiplier test. *Ann. Math. Statist.* **30** 389–407.

[17] Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans. Amer. Math. Soc.* **54** 426–482.

[18] Wilks, S. S. (1962). *Mathematical Statistics*. Wiley, New York.

[19] Woolf, B. (1955). On estimating the relation between blood group and disease. *Ann Human Genetics* **19** 251–253.