The Annals of Mathematical Statistics
1971, Vol. 42, No. 2, 594–606

MARGINAL HOMOGENEITY OF MULTIDIMENSIONAL CONTINGENCY TABLES[1]

BY S. KULLBACK

The George Washington University

0. Summary. Tests of marginal homogeneity in a two-way contingency table given by [1], [3], and [13] do not seem to lend themselves easily to extension to the problem of m-way marginal homogeneity in an N-way $r \times r \times \cdots \times r$ contingency table, $m < N$. The principle of minimum discrimination information estimation and the associated minimum discrimination information statistic applied in [5] to the problem of marginal homogeneity in an $r \times r$ contingency table can be easily extended to the case of a multidimensional contingency table. Estimates of the cell entries under the hypotheses of m-way marginal homogeneity are given. Relationships among the tests of homogeneity for m-way, $m = 1, 2, \cdots, N-1$, marginals are given by an analysis of information. Numerical results are given for two sample $3 \times 3 \times 3$ tables, and two 5×5 tables.

1. Introduction. In the study of the association between the characteristics in an $r \times r$ contingency table a particular hypothesis of interest is that of symmetry of the cell frequencies about the main diagonal. Bowker [2] who gave a large sample chi-square type test for the null hypothesis of symmetry noted that "the weaker hypothesis of equality of marginal distributions would also be of interest, especially in the absence of symmetry; this problem appears to be somewhat more difficult to handle by straightforward methods." Stuart [13] considered the problem of marginal homogeneity. Since the two marginals are not independent the usual test of homogeneity of independent samples is not applicable. It was noted by Stuart [13] that the likelihood-ratio principle yields an intractable result in this case. Stuart defined a test statistic which is a quadratic in the differences of the corresponding marginal values, with matrix the inverse of a consistent estimate of the covariance matrix of the differences under the null hypothesis. Stuart's statistic is asymptotically distributed as χ^2 with $r-1$ degrees of freedom under the null hypothesis of marginal homogeneity.

Bhapkar [1] considered the problem of marginal homogeneity using his result that Wald's statistic [14] is algebraically equivalent to the χ_1^2 statistic of Neyman [12] for testing linear hypotheses in categorical data. Bhapkar [1] proposed a test statistic which is also a quadratic in the differences of the corresponding marginal entries with matrix the inverse of a consistent estimate of the covariance matrix of the differences even if the null hypothesis does not hold. Under the null hypo-

Received August 21, 1969.

[1] Supported in part by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under Grant AFOSR-68-1513.

thesis of marginal homogeneity Bhapkar's statistic is asymptotically distributed as χ^2 with $r-1$ degrees of freedom.

Caussinus [3] defined the hypothesis of quasi-symmetry and showed that quasi-symmetry and marginal homogeneity together imply and are implied by symmetry. Caussinus gave maximum likelihood estimates for the cell frequencies of an observed contingency table under the hypothesis of quasi-symmetry, and a chi-square type statistic for the test of this hypothesis. He proposed the difference between the chi-square type statistics for symmetry and quasi-symmetry as the test for marginal homogeneity.

In [5] the principle of minimum discrimination information estimation [6], [7], was applied to obtain RBAN estimates of the cell frequencies of an $r \times r$ contingency table under the hypothesis of marginal homogeneity. The associated minimum discrimination information statistic which is distributed asymptotically as χ^2 under the null hypothesis provides a test of significance.

The tests of marginal homogeneity in a two-way table given by [1], [3], and [13] do not yield estimates of the cell frequencies under the hypothesis and do not seem to lend themselves easily to extension to the problem of marginal homogeneity in an N-way $r \times r \cdots \times r$ contingency table as does the method based on the principle of minimum discrimination information estimation. The m-way, $m = 1, 2, \cdots, N-1$, marginals are examined for homogeneity, and relationships among the associated minimum discrimination information statistics given by an analysis of information.

In [5] the solution to the problem of minimum discrimination information estimation of the cell frequencies in an $r \times r$ contingency table under an hypothesis of marginal homogeneity was obtained by a direct formulation using Lagrange undetermined multipliers. Although such a direct approach should also be applicable to the case of multidimensional contingency tables, it turns out that a two-step approach using certain results from [6] is more convenient algebraically.

The results of [6] that we shall need are summarized in the following two theorems. For convenience we give the results in terms of a four-way table, but the results generally are readily apparent.

THEOREM 1.1. *Given a contingency table* $\pi(ijkl)$, $i = 1, \cdots, r$, $j = 1, \cdots, s$, $k = 1, \cdots, t$, $l = 1, \cdots, u$, $\pi(ijkl) > 0$, $\sum\sum\sum\sum \pi(ijkl) = 1$. *Consider all contingency tables* $p(ijkl)$ *of the same dimensions. The minimum value of the discrimination information*

$$I(p:\pi) = \sum\sum\sum\sum p(ijkl) \ln (p(ijkl)/\pi(ijkl)) \tag{1.1}$$

(a) *when the one-way marginals*

$$p(i\cdots), \qquad p(\cdot j \cdot \cdot), \qquad p(\cdot \cdot k \cdot), \qquad p(\cdots l) \tag{1.2}$$

are given is attained for

$$p_1^*: p_1^*(ijkl) = a(i)b(j)c(k)d(l)\pi(ijkl), \tag{1.3}$$

$$\sum\sum\sum\sum p_1^*(ijkl) = 1,$$

where the parameters $a(i), \cdots, d(l)$ *are determined so that* $p_1{}^*$ *satisfies the marginal restraints and*

$$(1.4)\quad I(p_1{}^*:\pi) = \textstyle\sum_i p(i\cdots)\ln a(i) + \sum_j p(\cdot j\cdot\cdot)\ln b(j) + \sum_k p(\cdot\cdot k\cdot)\ln c(k) + \sum_l p(\cdots l)\ln d(l);$$

(b) *when the two-way marginals*

$$(1.5)\qquad p(ij\cdot\cdot),\ p(i\cdot k\cdot), \cdots, p(\cdot\cdot kl)$$

are given is attained for

$$(1.6)\qquad p_2{}^*: p_2{}^*(ijkl) = a(ij)b(ik)\cdots f(kl)\pi(ijkl),\qquad \textstyle\sum\sum\sum\sum p_2{}^*(ijkl) = 1,$$

where the parameters $a(ij), \cdots, f(kl)$ *are determined so that* $p_2{}^*$ *satisfies the marginal restraints and*

$$(1.7)\qquad I(p_2{}^*:\pi) = \textstyle\sum_i\sum_j p(ij\cdot\cdot)\ln a(ij) + \cdots + \sum_k\sum_l p(\cdot\cdot kl)\ln f(kl);$$

(*c*) *when the three-way marginals*

$$(1.8)\qquad p(ijk\cdot),\qquad p(i\cdot kl),\qquad p(ij\cdot l),\qquad p(\cdot jkl)$$

are given is attained for

$$(1.9)\qquad p_3{}^*: p_3{}^*(ijkl) = a(ijk)b(ikl)\cdots d(jkl)\pi(ijkl),\qquad \textstyle\sum\sum\sum\sum p_3{}^*(ijkl) = 1,$$

where the parameters $a(ijk), \cdots, d(jkl)$ *are determined so that* $p_3{}^*$ *satisfies the marginal restraints and*

$$(1.10)\qquad I(p_3{}^*:\pi) = \textstyle\sum_i\sum_j\sum_k p(ijk\cdot)\ln a(ijk) + \cdots + \sum_j\sum_k\sum_l p(\cdot jkl)\ln d(jkl).$$

THEOREM 1.2. *If* $\pi(ijkl)$ *in Theorem* 1.1 *is* $\hat{p}(ijkl) = x(ijkl)/n$, *where* $x(ijkl)$ *is the observed number in the* $ijkl$th *cell of a contingency table with* $\sum\sum\sum\sum x(ijkl) = n$, *the minimizing set* $\hat{p}^*(ijkl)$ *is a* RBAN *estimator and the minimum discrimination information statistic*

$$(1.11)\qquad 2nI(\hat{p}^*:\hat{p}) = 2n\textstyle\sum\sum\sum\sum \hat{p}^*(ijkl)\ln(\hat{p}^*(ijkl)/\hat{p}(ijkl)) = 2\sum\sum\sum\sum x^*(ijkl)\ln(x^*(ijkl)/x(ijkl)) = 2I(x^*:x)$$

where $x^*(ijkl) = n\hat{p}^*(ijkl)$, *is asymptotically distributed as* χ^2 *with appropriate degrees of freedom under the null hypothesis.*

2. Marginal homogeneity. (a) *One-way marginals.* Let us consider the specifications of Theorem 1.1 with $r = s = t = u$. Suppose there is now imposed the marginal homogeneity requirement that

$$(2.1)\qquad p(i\cdots) = p(\cdot i\cdot\cdot) = p(\cdot\cdot i\cdot) = p(\cdots i),\qquad \textstyle\sum_i p(i\cdots) = 1,\qquad i = 1, 2, \cdots, r,$$

then the minimum value of (1.1) over all $p(ijkl)$ satisfying (2.1) may be obtained by minimizing (1.4) subject to (2.1), that is,

$$\frac{\partial I(p_1{}^*:\pi)}{\partial p(i\cdots)} = \ln a(i)b(i)c(i)d(i) - \ln a(r)b(r)c(r)d(r) = 0, \quad i = 1, 2, \cdots, r-1 \tag{2.2}$$

or

$$a(i)b(i)c(i)d(i) = a(r)b(r)c(r)d(r) = \gamma_1, \qquad i = 1, 2, \cdots, r-1. \tag{2.3}$$

Hence (1.3) now becomes

$$p_1{}^*(ijkl) = \frac{a(i)b(j)c(k)}{a(l)b(l)c(l)}\,\pi(ijkl)\gamma_1, \tag{2.4}$$

$$\gamma_1 = 1/\sum\sum\sum\sum \frac{a(i)b(j)c(k)}{a(l)b(l)c(l)}\,\pi(ijkl),$$

where the parameters $a(i)$, $b(j)$, $c(k)$ must satisfy the marginal homogeneity restraints

$$a(i)\sum_j\sum_k\sum_l \frac{b(j)c(k)}{a(l)b(l)c(l)}\pi(ijkl) = b(i)\sum_s\sum_k\sum_l \frac{a(s)c(k)}{a(l)b(l)c(l)}\,\pi(sikl) \tag{2.5}$$

$$= c(i)\sum_s\sum_j\sum_l \frac{a(s)b(j)}{a(l)b(l)c(l)}\,\pi(sjil)$$

$$= \frac{1}{a(i)b(i)c(i)}\sum_s\sum_j\sum_k a(s)b(j)c(k)\pi(sjki), \qquad i = 1, \cdots, r,$$

and

$$I(p_1{}^*:\pi) = \ln \gamma_1. \tag{2.6}$$

There is presented in Section 3 a convergent iterative procedure for determining the values given in (2.4).

(b) *Two-way marginals.* Suppose there is now imposed the marginal homogeneity requirement that

$$p(ij\cdot\cdot) = p(i\cdot j\cdot) = \cdots = p(\cdot\cdot ij), \qquad \sum_i\sum_j p(ij\cdot\cdot) = 1, \; i,j = 1, 2, \cdots, r \tag{2.7}$$

then the minimum value of (1.1) over all $p(ijkl)$ satisfying (2.7) may be obtained by minimizing (1.7) subject to (2.7), that is

$$\frac{\partial I(p_2{}^*:\pi)}{\partial p(ij\cdot\cdot)} = \ln a(ij)b(ij)\cdots f(ij) - \ln a(rr)b(rr)\cdots f(rr) = 0, \tag{2.8}$$

or

$$a(ij)b(ij)\cdots f(ij) = a(rr)b(rr)\cdots f(rr) = \gamma_2, \qquad i = 1, \cdots, r, j = 1, \cdots, r-1. \tag{2.9}$$

Hence (1.6) now becomes

$$p_2^*(ijkl) = \frac{a(ij)b(ik)\cdots e(jl)}{a(kl)b(kl)\cdots e(kl)}\,\pi(ijkl)\gamma_2, \tag{2.10}$$

$$\gamma_2 = 1/\sum\sum\sum\sum \frac{a(ij)\cdots e(jl)}{a(kl)\cdots e(kl)}\,\pi(ijkl),$$

where the parameters $a(ij), \cdots, e(jl)$ must satisfy the marginal homogeneity restraints

$$\begin{aligned} a(ij)\textstyle\sum_k\sum_l & \frac{b(ik)\cdots e(jl)}{a(kl)\cdots e(kl)}\pi(ijkl) \\ &= b(ij)\textstyle\sum_s\sum_l \frac{a(is)c(il)\cdots e(sl)}{a(jl)b(jl)\cdots e(jl)}\,\pi(isjl) = \cdots = \\ &= e(ij)\textstyle\sum_s\sum_k \frac{a(si)b(sk)\cdots d(ik)}{a(kj)b(kj)\cdots e(kj)}\,\pi(sikj) \\ &= \frac{1}{a(ij)b(ij)\cdots e(ij)}\textstyle\sum_s\sum_t a(st)b(si)\cdots e(tj)\pi(stij), \quad i, j = 1, \cdots r, \end{aligned} \tag{2.11}$$

and

$$I(p_2^*:\pi) = \ln \gamma_2. \tag{2.12}$$

There is presented in Section 3 a convergent iterative procedure for determining the values given in (2.10).

(c) *Three-way marginals.* Suppose there is now imposed the marginal homogeneity requirement that

$$p(ijk\cdot) = p(ij\cdot k) = p(i\cdot jk) = p(\cdot ijk), \quad \textstyle\sum_i\sum_j\sum_k p(ijk\cdot) = 1, \qquad i, j, k = 1, \cdots r, \tag{2.13}$$

then the minimum value of (1.1) over all $p(ijkl)$ satisfying (2.13) may be obtained by minimizing (1.10) subject to (2.13), that is,

$$\frac{\partial I(p_3^*:\pi)}{\partial p(ijk\cdot)} = \ln a(ijk)b(ijk)c(ijk)d(ijk) - \ln a(rrr)b(rrr)c(rrr)d(rrr) = 0, \tag{2.14}$$

or

$$a(ijk)b(ijk)c(ijk)d(ijk) = a(rrr)b(rrr)c(rrr)d(rrr) = \gamma_3, \qquad i, j = 1, 2, \cdots, r,\ k = 1, 2, \cdots, r-1. \tag{2.15}$$

Hence (1.9) now becomes

$$p_3^*(ijkl) = \frac{a(ijk)b(ijl)c(ikl)}{a(jkl)b(jkl)c(jkl)}\pi(ijkl)\gamma_3, \tag{2.16}$$

$$\gamma_3 = 1/\sum\sum\sum\sum \frac{a(ijk)\cdots c(ikl)}{a(jkl)\cdots c(jkl)}\,\pi(ijkl),$$

where the parameters $a(ijk)\cdots, c(ikl)$ must satisfy the marginal homogeneity restraints

$$
\begin{aligned}
& a(ijk)\sum_l \frac{b(ijl)c(ikl)}{a(jkl)b(jkl)c(jkl)}\pi(ijkl) \\
(2.17)\qquad & = b(ijk)\sum_s \frac{a(ijs)c(isk)}{a(jsk)b(jsk)c(jsk)}\pi(ijsk) \\
& = \cdots = \\
& = \frac{1}{a(ijk)b(ijk)c(ijk)}\sum_s a(sij)b(sik)c(sjk)\pi(sijk), \quad i, j, k, = 1, \cdots, r,
\end{aligned}
$$

and

$$(2.18)\qquad I(p_3^*:\pi) = \ln \gamma_3 .$$

There is presented in Section 3 a convergent iterative procedure for determining the values given in (2.16).

3. Iterative procedure. (a) *One-way marginals.* The solution of (2.4) and (2.5) may be obtained by cycling through the iterations

$$
\begin{aligned}
& p^{(3n+1)}(ijkl) = \left(\frac{p^{(3n)}(\cdots i)p^{(3n)}(l\cdots)}{p^{(3n)}(i\cdots)p^{(3n)}(\cdots l)}\right)^{\frac{1}{2}} p^{(3n)}(ijkl)\gamma_1^{(3n)} \\
(3.1)\qquad & p^{(3n+2)}(ijkl) = \left(\frac{p^{(3n+1)}(\cdots j)p^{(3n+1)}p(\cdot l\cdot\cdot)}{p^{(3n+1)}(\cdot j\cdot\cdot)p^{(3n+1)}(\cdot\cdot\cdot l)}\right)^{\frac{1}{2}} p^{(3n+1)}(ijkl)\gamma_1^{(3n+1)} \\
& p^{(3n+3)}(ijkl) = \left(\frac{p^{(3n+2)}(\cdot\cdot\cdot k)p^{(3n+2)}(\cdot\cdot l\cdot)}{p^{(3n+2)}(\cdot\cdot k\cdot)p^{(3n+2)}(\cdot\cdot\cdot l)}\right)^{\frac{1}{2}} p^{(3n+2)}(ijkl)\gamma_1^{(3n+2)}
\end{aligned}
$$

where the γ's are normalizing factors so that for example

$$(3.2)\qquad \gamma_1^{(3n)} = 1/\sum\sum\sum\sum\left(\frac{p^{(3n)}(\cdot\cdot\cdot i)p^{(3n)}(l\cdot\cdot\cdot)}{p^{(3n)}(i\cdot\cdot\cdot)p^{(3n)}(\cdot\cdot\cdot l)}\right)^{\frac{1}{2}} p^{(3n)}(ijkl),$$

and in (3.1) and the similar iterations (3.4), (3.6)

$$(3.3)\qquad p^{(0)}(ijkl) = \pi(ijkl).$$

(b) *Two-way marginals.* The solution of (2.10) and (2.11) may be obtained by cycling through the iterations

$$
\begin{aligned}
& p^{(5n+1)}(ijkl) = \left(\frac{p^{(5n)}(\cdot\cdot ij)p^{(5n)}(kl\cdot\cdot)}{p^{(5n)}(ij\cdot\cdot)p^{(5n)}(\cdot\cdot kl)}\right)^{\frac{1}{2}} p^{(5n)}(ijkl)\gamma_2^{(5n)} \\
(3.4)\qquad & p^{(5n+2)}(ijkl) = \left(\frac{p^{(5n+1)}(\cdot\cdot ik)p^{(5n+1)}(k\cdot l\cdot)}{p^{(5n+1)}(i.k.)p^{(5n+1)}(..kl)}\right)^{\frac{1}{2}} p^{(5n+1)}(ijkl)\gamma_2^{(5n+1)} \\
& \cdots \qquad\qquad \cdots \qquad\qquad \cdots \\
& p^{(5n+5)}(ijkl) = \left(\frac{p^{(5n+4)}(\cdot\cdot jl)p^{(5n+4)}(\cdot k\cdot l)}{p^{(5n+4)}(\cdot j\cdot l)p^{(5n+4)}(\cdot\cdot kl)}\right)^{\frac{1}{2}} p^{(5n+4)}(ijkl)\gamma_2^{(5n+4)}
\end{aligned}
$$

where the γ's are normalizing factors so that for example

$$\gamma_2^{(5n)} = 1/\sum\sum\sum\sum \left(\frac{p^{(5n)}(\cdot\cdot ij)p^{(5n)}(kl\cdot\cdot)}{p^{(5n)}(ij\cdot\cdot)p^{(5n)}(\cdot\cdot kl)}\right)^{\frac{1}{2}} p^{(5n)}(ijkl). \tag{3.5}$$

(c) *Three-way marginals.* The solution of (2.16) and (2.17) may be obtained by cycling through the iterations

$$\begin{aligned} p^{(3n+1)}(ijkl) &= \left(\frac{p^{(3n)}(\cdot ijk)p^{(3n)}(jkl\cdot)}{p^{(3n)}(ijk\cdot)p^{(3n)}(\cdot jkl)}\right)^{\frac{1}{2}} p^{(3n)}(ijkl)\gamma_3^{(3n)} \\ p^{(3n+2)}(ijkl) &= \left(\frac{p^{(3n+1)}(\cdot ijl)p^{(3n+1)}(jk\cdot l)}{p^{(3n+1)}(ij\cdot l)p^{(3n+1)}(\cdot jkl)}\right)^{\frac{1}{2}} p^{(3n+1)}(ijkl)\gamma_3^{(3n+1)} \\ p^{(3n+3)}(ijkl) &= \left(\frac{p^{(3n+2)}(\cdot ikl)p^{(3n+2)}(j\cdot kl)}{p^{(3n+2)}(i\cdot kl)p^{(3n+2)}(\cdot jkl)}\right)^{\frac{1}{2}} p^{(3n+2)}(ijkl)\gamma_3^{(3n+2)} \end{aligned} \tag{3.6}$$

where the γ's are normalizing factors so that for example

$$\gamma_3^{(3n)} = 1/\sum\sum\sum\sum \left(\frac{p^{(3n)}(\cdot ijk)p^{(3n)}(jkl\cdot)}{p^{(3n)}(ijk\cdot)p^{(3n)}(\cdot jkl)}\right)^{\frac{1}{2}} p^{(3n)}(ijkl). \tag{3.7}$$

The fact that the iterations converge to the corresponding p*-distribution may be shown by the same procedure as the proof of the convergence for the two-way table given in [5] and consequently details are omitted here.

4. Analysis of information. Certain interrelationships among the various information values will now be derived.

THEOREM 4.1.

$$I(p_3^*:\pi) \geqq I(p_2^*:\pi) \geqq I(p_1^*:\pi)$$

PROOF. Theorem 4.1 is an immediate consequence of the fact that

$$(2.13) \Rightarrow (2.7) \Rightarrow (2.1). \tag{4.1}$$

THEOREM 4.2.

$$I(p_m:\pi) = I(p_m:p_m^*)+I(p_m^*:\pi)$$

where p_m, $m = 1, 2, 3$ *is any contingency table* $p(ijkl)$ *with homogeneous m-way marginals and* p_m^* *is the minimizing distribution as described in Section* 2.

PROOF. For notational convenience the proof of Theorem 4.2 is given for $m = 2$, the proof for the other cases being similar. Since

$$\begin{aligned} I(p_2^*:\pi) &= \sum\sum\sum\sum p_2^*(ijkl) \ln \frac{a(ij)b(ik)\cdots e(jl)}{a(kl)b(kl)\cdots e(kl)}\gamma_2 = \ln \gamma_2 \\ &= \sum\sum\sum\sum p_2(ijkl) \ln \frac{a(ij)b(ik)\cdots e(jl)}{a(kl)b(kl)\cdots e(kl)}\gamma_2 \\ &= \sum\sum\sum\sum p_2(ijkl) \ln \frac{p_2^*(ijkl)}{\pi(ijkl)} \end{aligned} \tag{4.2}$$

it follows that

$$I(p_2^*:\pi)+I(p_2:p_2^*) = \sum\sum\sum\sum p_2(ijkl)\ln\frac{p_2^*(ijkl)}{\pi(ijkl)} + \sum\sum\sum\sum p_2(ijkl)\ln\frac{p_2(ijkl)}{p_2^*(ijkl)} = I(p_2:\pi). \tag{4.3}$$

THEOREM 4.3.

$$I(p_r^*:\pi) = I(p_r^*:p_m^*)+I(p_m^*:\pi), \qquad m = 1, 2, \cdots, r-1.$$

PROOF. THEOREM 4.3 is an immediate consequence of Theorem 4.2 since the p_r^*-distribution has homogeneous m-way marginals for $m = 1, 2, \cdots, r-1$. In fact, in view of (2.4) and (2.6), (2.10) and (2.12), (2.16) and (2.18), it is seen that Theorem 4.3 may also be written as

$$\ln\gamma_r = \ln\gamma_m+\ln\frac{\gamma_r}{\gamma_m}, \qquad m = 1, 2, \cdots, r-1. \tag{4.4}$$

THEOREM 4.4.

$$I(p_3^*:p_1^*) = I(p_3^*:p_2^*)+I(p_2^*:p_1^*).$$

PROOF. Theorem 4.4 readily follows from THEOREM 4.3 by using different values for m, and indeed Theorem 4.4 may also be written as

$$\ln\frac{\gamma_3}{\gamma_1} = \ln\frac{\gamma_3}{\gamma_2}+\ln\frac{\gamma_2}{\gamma_1}. \tag{4.5}$$

It should be noted that (3.4) will yield the same values for $p_2^*(ijkl)$ if the iteration starts with $p^{(0)}(ijkl) = p_1^*(ijkl)$; the differences will appear in the associated parameters, that is

$$\begin{aligned} p_2^*(ijkl) &= \frac{a(ij)b(ik)\cdots e(jl)}{a(kl)b(kl)\cdots e(kl)}\pi(ijkl)\gamma_2 \\ &= \frac{A(ij)B(ik)\cdots E(jl)}{A(kl)B(kl)\cdots E(kl)}p_1^*(ijkl)\gamma_{21} \\ &= \frac{A(ij)B(ij)\cdots E(jl)}{A(kl)B(kl)\cdots E(kl)}\cdot\frac{a(i)b(j)c(k)}{a(l)b(l)c(l)}\pi(ijkl)\gamma_1\gamma_{21}, \end{aligned} \tag{4.6}$$

$$\gamma_{21} = 1/\sum\sum\sum\sum\frac{A(ij)B(ik)\cdots E(jl)}{A(kl)B(kl)\cdots E(kl)}p_1^*(ijkl), \tag{4.7}$$

$$\gamma_2 = \gamma_1\gamma_{21}. \tag{4.8}$$

Similarly (3.6) will yield the same value for $p_3^*(ijkl)$ if the iteration starts with $p^{(0)}(ijkl) = p_1^*(ijkl)$ or $p^{(0)}(ijkl) = p_2^*(ijkl)$; the differences will appear in the associated parameters, that is,

$$(4.9)\quad \begin{aligned} p_3^*(ijkl) &= \frac{a(ijk)b(ijl)c(ikl)}{a(jkl)b(jkl)c(jkl)}\pi(ijkl)\gamma_3 \\ &= \frac{A(ijk)B(ijl)C(ikl)}{A(jkl)B(jkl)C(jkl)}p_2^*(ijkl)\gamma_{32} \\ &= \frac{A(ijk)\cdots C(ikl)}{A(jkl)\cdots C(jkl)}\frac{A(ij)\cdots E(jl)}{A(kl)\cdots E(kl)}p_1^*(ijkl)\gamma_{21}\gamma_{32} \\ &= \frac{A(ijk)\cdots C(ikl)}{A(jkl)\cdots C(jkl)}\cdot\frac{A(ij)\cdots E(jl)}{A(kl)\cdots E(kl)}\cdot\frac{a(i)\cdots c(k)}{a(l)\cdots c(l)}\pi(ijkl)\gamma_1\gamma_{21}\gamma_{32}, \end{aligned}$$

$$(4.10)\quad \gamma_{32} = 1/\sum\sum\sum\sum\frac{A(ijk)\cdots C(ikl)}{A(jkl)\cdots C(jkl)}p_2^*(ijkl),$$

$$(4.11)\quad \gamma_3 = \gamma_1\gamma_{21}\gamma_{32} = \gamma_2\gamma_{32}.$$

Thus (4.5) may also be written as

$$(4.12)\quad \ln\gamma_{21}\gamma_{32} = \ln\gamma_{32}+\ln\gamma_{21}$$

and (4.4) as

$$(4.13)\quad \ln\gamma_1\gamma_{21}\gamma_{32}\cdots\gamma_{r(r-1)} = \ln\gamma_1\gamma_{21}\gamma_{32}\cdots\gamma_{m(m-1)}+\ln\gamma_{rm}, \quad m = 1, 2, \cdots, r-1.$$

A property of the iterative procedure similar to that just mentioned is also described in ([8] page 174).

Let us now take $\pi(ijkl) = \hat{p}(ijkl) = x(ijkl)/n$ and use the notation in Theorem 1.2, also using $\hat{\gamma}_m$ and $\hat{\gamma}_{mr}$ in this case for γ_m and γ_{mr}. Corresponding to Theorem 4.3 and Theorem 4.4 the following relations exist among the minimum discrimination information statistics,

$$(4.14)\quad 2I(x_r^*:x) = 2I(x_r^*:x_m^*)+2I(x_m^*:x), \qquad m = 1, 2, \cdots, r-1,$$

$$(4.15)\quad 2I(x_3^*:x_1^*) = 2I(x_3^*:x_2^*)+2I(x_2^*:x_1^*).$$

In (4.14) and (4.15) the terms of the form $2I(x_a^*:x)$ are interaction-type measures of the marginal homogeneity hypotheses and the terms of the form $2I(x_a^*:x_b^*)$ are effect-type measures. The interaction-type measure is a comparison of an estimated table with the original table and the effect-type measure is a comparison of estimated tables under different constraints, and is a measure of the effect of the differences in the constraints. For a discussion of interactions and effects as above see [8].

Note that an interaction-type measure may be analyzed into the sum of an interaction-type measure and an effect-type measure, while an effect-type measure may be analyzed into the sum of effect-type measures.

In accordance with Theorem 1.2 the minimum discrimination information statistics are asymptotically distributed as χ^2 under the null hypothesis and the relations may be summarized in the analysis of information Table 4.1. The degrees

TABLE 4.1

Information		d.f.
$2I(x_3^*:x)$	$= 2n \ln \hat{\gamma}_1 + 2n \ln \hat{\gamma}_{21} + 2n \ln \hat{\gamma}_{32}$	$3(r-1)+5(r-1)^2+3(r-1)^3$
$2I(x_3^*:x_2^*$	$= 2n \ln \hat{\gamma}_{32}$	$3(r-1)^3$
$2I(x_2^*:x)$	$= 2n \ln \hat{\gamma}_1 + 2n \ln \hat{\gamma}_{21}$	$3(r-1)+5(r-1)^2$
$2I(x_2^*:x_1^*)$	$= 2n \ln \hat{\gamma}_{21}$	$5(r-1)^2$
$2I(x_1^*:x)$	$= 2n \ln \hat{\gamma}_1$	$3(r-1)$

of freedom in Table 4.1 follow from the argument in ([4] page 928) (see also [6] page 187 or [8] Table 3.1]) setting $r = s = t = u$ and noting that in the four-way table, in the case of homogeneity, $\hat{\gamma}_1$ depends on three parameters $a(i)$, $b(j)$, $c(k)$, $\hat{\gamma}_{21}$ depends on five parameters $A(ij)$, $B(ik)$, $\cdots$, $E(jl)$, and $\hat{\gamma}_{32}$ depends on three parameters $A(ijk)$, $B(ijl)$, $C(ikl)$.

5. Examples. For the first example two observed $3 \times 3 \times 3$ contingency tables $x(ijk)$ and the corresponding $x_1{}^*(ijk)$ tables are given to illustrate the procedures and results. The observed tables were obtained by sampling from a table with nonhomogeneous marginals to get Table 5.1a and from a table with homogeneous one-way marginals to get Table 5.2a.

TABLE 5.1a
$x(ijk)$

j		1			2			3		
k	1	2	3	1	2	3	1	2	3	$x(i\,.\,.)$
1	223	24	6	40	'42	2	19	4	12	372
2	28	6	9	25	218	6	3	13	9	317
3	26	3	18	18	30	24	12	16	164	311
	277	33	33	83	290	32	34	33	185	1000
$x(\,.\,j\,.\,)$		343			405			252		
$x(\,.\,.\,k)$	394	356	250							

TABLE 5.1b
$x_1^*(ijk)$

j		1			2			3			
	k	1	2	3	1	2	3	1	2	3	$x_1^*(i\ldots)$
	1	228.8	30.1	8.9	26.5	33.9	1.9	19.0	4.9	17.4	371.4
i	2	36.5	9.6	17.0	21.0	223.7	7.3	3.8	20.2	16.6	355.7
	3	18.9	2.7	18.9	8.4	17.1	16.3	8.5	13.8	168.3	272.9
		284.2	42.4	44.8	55.9	274.7	25.5	31.3	38.9	202.3	1000.0
$x_1^*(.j.)$			371.4			356.1			272.5		
$x_1^*(..k)$		371.4	356.0	272.6							

For the values in Tables 5.1a and 5.1b it was found that $2I(x_1^*:x) = 52.55$, which as a χ^2 with $2(3-1) = 4$ degrees of freedom leads us to reject the null hypothesis of one-way marginal homogeneity, as we should for these tables. For the values in Tables 5.2a and 5.2b it was found that $2I(x_1^*:x) = 6.23$, which as a

TABLE 5.2a
$x(ijk)$

j		1			2			3			
	k	1	2	3	1	2	3	1	2	3	$x(i..)$
	1	229	27	15	22	39	2	18	5	16	373
i	2	34	15	17	18	212	13	6	26	16	357
	3	11	1	28	7	14	20	8	13	168	270
		274	43	60	47	265	35	32	44	200	1000
$x(.j.)$			377			347			276		
$x(..k)$		353	352	295							

TABLE 5.2b
$x_1^*(ijk)$

j		1			2			3			
	k	1	2	3	1	2	3	1	2	3	$x_1^*(i..)$
	1	229.7	24.4	11.6	23.9	38.2	1.7	19.8	4.9	13.6	367.8
i	2	34.9	13.9	13.5	20.1	212.7	11.1	6.8	26.4	13.9	353.3
	3	13.1	1.1	25.6	9.0	16.2	19.9	10.4	15.2	168.5	279.0
		277.7	39.4	50.7	53.0	267.1	32.7	37.0	46.5	196.0	1000.1
$x_1^*(.j.)$			367.8			352.8			279.5		
$x_1^*(..k)$		367.7	353.0	279.4							

χ^2 with 4 degrees of freedom leads us to accept the null hypothesis of one-way marginal homogeneity, as we should for these tables.

It is of interest to remark that the population table from which Table 5.2a was obtained by sampling was formed by starting with a table with nonhomogeneous marginals and using the iterative procedure to generate a population table with homogeneous marginals.

For the second example we use the data in Table 5.3a given by Mosteller [11]

TABLE 5.3a
Data from [11] [9] [10]

		Status category of son's occupation 1	2	3	4	5	
	1	50	45	8	18	8	129
Status category		18	17	16	4	2	57
of father's	2	28	174	84	154	55	495
occupation		24	105	109	59	21	318
	3	11	78	110	223	96	518
		23	84	289	217	95	708
	4	14	150	185	714	447	1510
		8	49	175	348	198	778
	5	3	42	72	320	411	848
		6	8	69	201	246	530
		106	489	459	1429	1017	3500
		79	263	658	829	562	2391

Upper numbers are British counts
Lower numbers are Danish counts

TABLE 5.3b
Estimate under marginal homogeneity

		Status category of son's occupation 1	2	3	4	5	
	1	50.236	38.569	7.088	15.113	5.775	116.781
Status category		18.069	22.947	19.780	4.471	2.194	67.461
of father's	2	32.978	174.821	87.242	151.572	46.542	493.156
occupation		17.917	105.404	100.216	49.041	17.132	289.711
	3	12.533	75.812	110:519	212.324	78.587	489.775
		18.748	92.068	290.113	196.937	84.622	682.489
	4	16.832	153.844	196.140	717.370	386.133	1470.319
		7.213	59.406	194.316	349.340	195.087	805.362
	5	4.195	50.102	88.785	373.947	412.940	929.969
		5.512	9.882	78.060	205.576	246.948	545.977
		116.774	493.148	489.775	1470.326	929.977	3500.000
		67.459	289.707	682.485	805.366	545.983	2391.000

with reference to Levine [9] [10] who studied Glass's British and Svalastoga's Danish occupational mobility data. The upper numbers in the cells are the British counts showing the joint distribution of father's and son's occupation distributed into five categories, the lower numbers are the Danish counts in the corresponding categories. In Table 5.3b are given the estimates under the hypothesis of marginal homogeneity. For both the British and Danish counts twelve iterations were needed to reach the criterion that the difference between corresponding marginals be less than 0.01.

For the British counts it was found that $2I(x_1^*:x) = 32.95$ and for the Danish counts $2I(x_1^*:x) = 18.38$, which as chi-squares with 4 degrees of freedom lead us to reject the null hypothesis of marginal homogeneity. This inference would be consistent with the existence of occupational mobility in both countries.

Acknowledgments. I am indebted to Mr. Yehuda Molk, Mr. E. W. Zedlewski, and Mrs. Marian Romer Fisher for programming and computation.

REFERENCES

[1] Bhapkar, V. P. (1966). A note on the equivalence of two criteria for hypotheses in categorical data. *J. Amer. Statist. Assoc.* **61** 228–35.

[2] Bowker, A. H. (1948). A test for symmetry in contingency tables. *J. Amer. Statist. Assoc.* **43** 572–4.

[3] Caussinus, H. (1965). Contribution a l'analyse statistique des tableaux de correlation. *Ann. Fac. Sci. Univ. Toulouse* **29** 77–182.

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

[5] Ireland, C. T., Ku, H. H. and Kullback, S. (1969). Symmetry and marginal homogeneity of an r x r contingency table. *J. Amer. Statist. Assoc.* **64** 1323–41.

[6] Ireland, C. T. and Kullback, S. (1968a). Contingency tables with given marginals. *Biometrika* **55** 179–88.

[7] Ireland, C. T. and Kullback, S. (1968b). Minimum discrimination information estimation. *Biometrics* **24** 707–13.

[8] Ku, H. H. and Kullback, S. (1968). Interaction in multidimensional contingency tables: an information theoretic approach. *J. Res. Nat. Bur. Standards Sect. B.***72** 159–199.

[9] Levine, J. H. (1967a). Measurement in the study of intergenerational status mobility. Ph.D. dissertation, Harvard Univ.

[10] Levine, J. H. (1967b). A measure of association for intergenerational status mobility. Unpublished manuscript.

[11] Mosteller, F. (1968). Association and estimation in contingency tables. *J. Amer. Statist. Assoc.* **63** 1–28.

[12] Neyman, J. (1949). Contributions to the theory of the χ^2 test. *Proc. First Berkeley Symp. Math. Statist. Prob.* Univ. California, 239–73.

[13] Stuart, A. (1955). A test for homogeneity of the marginal distributions in a two-way classification. *Biometrika* **42** 412–6.

[14] Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans. Amer. Math. Soc.* **54** 426–82.