Ann. Inst. Statist. Math.
Vol. 41, No. 1, 149-167 (1989)

ON THE TESTING OF MARGINAL HOMOGENEITY WITH A ONE-SIDED ALTERNATIVE IN THE ANALYSIS OF VARIANCE

CHU-IN CHARLES LEE[1*], TIM ROBERTSON[2*] AND F. T. WRIGHT[3*]

[1]*Department of Mathematics and Statistics, Memorial University of Newfoundland, St. John's, Newfoundland, Canada, A1C 5S7*
[2]*Department of Statistics and Actuarial Science, University of Iowa, Iowa City, IA 52242, U.S.A.*
[3]*Department of Mathematics and Statistics, University of Missouri-Rolla, Rolla, MO 65401, U.S.A.*

(Received October 19, 1987; revised July 18, 1988)

Abstract. We consider a two-factor experiment in which the factors have the same levels with a natural ordering among the levels. Likelihood ratio tests for testing equality of the main effects with a one-sided alternative and for testing the one-sided hypothesis as a null hypothesis are studied. Closed form expressions for the maximum likelihood estimates under the various hypotheses are obtained. The null hypothesis distributions for these test statistics are derived.

Key words and phrases: Chi-bar-squared distributions, convex duality, isotonic regression, likelihood ratio tests, one-sided hypotheses, order restricted inference, two-factor designs.

1. Introduction

Duality is a powerful tool for solving optimization problems and Luenberger (1969) contains an excellent discussion of the topic of duality. The concept of a simple (linear) order and the concept of stochastic ordering are related via Fenchel duality, as was first pointed out in Barlow and Brunk (1972). This duality between simple and stochastic orderings is also discussed in Section 1.7 and Chapters 5 and 6 of Robertson *et al.* (1988) and is further explored in order restricted testing problems in Robertson and Wright (1981, 1982).

In this paper we explore a further manifestation of this type of duality in an inference problem which has a different character than those discus-

*The efforts of the first author were supported by the NSERC of Canada. The efforts of the second author were supported by the Office of Naval Research under Contract ONR N00014-80-C-0321. The efforts of the third author were supported by the Office of Naval Research under Contract ONR N00014-80-C-0322.

sed above. This problem gives an elegant demonstration of the power of duality for solving optimization problems. We consider a "one-sided" analysis of a two-factor experiment in which the two factors each have the same levels and it is assumed that the levels have a natural ordering. For example, in a study of nutrition for plants, we may wish to compare fertilizer combinations of nitrogen and potassium at each of several levels. It is desired to compare the main effects of these two fertilizers on some variable of interest, when it is believed that the effect of one fertilizer dominates that of the other.

We assume the following standard parametric model:

$$(1.1) \qquad y_{ijr} = \mu_{ij} + e_{ijr}, \quad i, j = 1, 2, \ldots, k \quad \text{and} \quad r = 1, 2, \ldots, n,$$

where the e_{ijr} are assumed to be independent normal random variables with zero means and common variance, σ^2, and $\mu = [\mu_{ij}]$ is a matrix of unknown parameters. The main effect for the row (column) factor at level i is

$$\mu_{i+} = \sum_{j=1}^{k} \mu_{ij} \left(\mu_{+i} = \sum_{j=1}^{k} \mu_{ji} \right) \quad \text{for} \quad i = 1, 2, \ldots, k.$$

Within this context, we consider two order restricted testing problems. The first is to test homogeneity of the main effects corresponding to the two factors, specifically:

$$(1.2) \qquad H_0: \mu_{i+} = \mu_{+i}, \quad i = 1, 2, \ldots, k,$$

with a one-sided alternative given by

$$(1.3) \qquad H_1: \sum_{j=1}^{i} \mu_{j+} \geq \sum_{j=1}^{i} \mu_{+j} \quad \text{for} \quad i = 1, 2, \ldots, k.$$

In the fertilizer example, suppose that both the row effects, μ_{i+}, and the column effects, μ_{+i}, are nondecreasing. If the column factor, say potassium, were believed to have a more pronounced effect on the variable of interest than the row factor, then one could quantify the difference in the effects of the two factors by assuming that the values of μ_{+i} increase at a faster rate than do the values of μ_{i+}. Since $\Sigma\mu_{i+} = \Sigma\mu_{+i}$, this difference in effects would produce more variation in the column effects than in the row effects, and one quantification of greater dispersion is Schur-majorization. Because the row and column effects are nondecreasing, this is equivalent to (1.3).

The method used here for developing theory for the restriction (1.3) is applicable to any "cone restriction" which has (1.2) as a subhypothesis. The second testing problem we will consider is to test (1.3) as a null hypothesis.

To clarify the meaning of (1.3), we consider some possible configurations for the main effects. If μ_{ij} does not depend on i and is nondecreasing in j, then $\mu_{i+} = \mu_{++}/k$, μ_{+i} is nondecreasing with sum μ_{++}, and (1.3) holds. This is the extreme case in which the row factor has no effect on the variable of interest. With $k = 5$, suppose that

$$\mu_{+1} = -1, \quad \mu_{+5} = 1, \quad \mu_{+2} = \mu_{+3} = \mu_{+4} = 0 \quad \text{and} \quad \mu_{i+} = (i-3)/3 .$$

In this case, (1.3) holds even though the column factor has no effect except at its extreme values and the row factor has a linear effect. This is because there is more dispersion in the μ_{+i} values than in the μ_{i+} values. With $k = 4$, if $\mu_{+1} = \mu_{+2} = -1$, $\mu_{+3} = \mu_{+4} = 1$ and $\mu_{i+} = (2i-5)/3$, then (1.3) holds. In this example, $\mu_{1+} = \mu_{+1} = -1$ and $\mu_{4+} = \mu_{+4} = 1$, but again the μ_{+i} have the greater dispersion.

Of course, Schur-majorization is not the only way to characterize the concept of "greater dispersion". However, it is very useful since $\phi(x) \leq \phi(y)$ for every Schur-convex function ϕ and $y = (y_1, y_2, \dots, y_k)$ which Schur-majorizes $x = (x_1, x_2, \dots, x_k)$. This, in turn, has a number of important statistical applications (cf. Marshall and Olkin (1979)).

Robertson and Wright (1982), Pukelsheim (1984) and Cohen *et al.* (1989) study partial orderings on R^k which quantify the concept that one vector is more dispersed than another. In their terminology, (1.3) says that the vector $(\mu_{+1}, \mu_{+2}, \dots, \mu_{+k})$ is more isotonic than the vector $(\mu_{1+}, \mu_{2+}, \dots, \mu_{k+})$. They discuss tests of hypotheses involving this restriction for independent samples from the populations associated with the various parameter values. However, in the framework considered here, the observations on μ_{i+} are not independent of those on μ_{+i}.

Either of the likelihood ratio tests (LRT's) considered in this paper can be carried out under the additional assumption that the parameter matrix, μ, belongs to any one of the following linear models:

the additive model:

$$\mu_{ij} = m + \alpha_i + \beta_j ,$$

the interaction model:

$$\mu_{ij} = m + \alpha_i + \beta_j + (\alpha\beta)_{ij} ,$$

the cohort model:

$$\mu_{ij} = m + \alpha_i + \beta_j + \nu_{i+j-1} ,$$

the Latin square design:

$$\mu_{ij} = m + \alpha_i + \beta_j + \tau_{p(i,j)} + \rho_{q(i,j)} .$$

Additional models include the symmetric interaction model with $(\alpha\beta)_{ij} = (\alpha\beta)_{ji}$ as well as any of the above models without diagonal elements or with diagonal elements $\mu_{ii} = m + v_i$, since the hypotheses H_0 and H_1 do not involve the μ_{ii}. Within any one of these models, the maximum likelihood estimates (MLE's) subject to $H \cap H_0$ and $H \cap H_1$ can be found by first finding the MLE's subject to the model restriction, H, and then making a simple adjustment for the appropriate hypothesis, H_0 or H_1 (cf. Theorem 3.4). The null hypothesis distributions for the LRT statistics are found by making simple adjustments in the degrees of freedom for the distributions associated with the free model.

Fenchel duality is discussed in Section 2 and the remarkably simple forms of the duals of the subsets of the parameter space corresponding to H_0 and H_1 are derived. Closed form expressions for the estimates subject to the various restrictions are derived in Section 3 and we partition the sum of squares and derive distributions of quadratic forms in Section 4.

2. Parameter spaces

The set of unknown parameters, μ_{ij}, form a $k \times k$ matrix which we will think of as a point in the k^2-dimensional Euclidean space, $R^{k\times k}$, and we denote this point by $\mu = [\mu_{ij}]$. If μ and ν are two such matrices, we define the usual inner product in $R^{k\times k}$ by

$$\langle \mu, \nu \rangle = \sum_{i=1}^{k} \sum_{j=1}^{k} \mu_{ij} \nu_{ij} .$$

If C is a closed convex cone in $R^{k\times k}$, then the Fenchel dual or polar of C is defined by

$$C^* = \{\nu : \langle \nu, \mu \rangle \leq 0 \text{ for all } \mu \in C\} . \tag{2.1}$$

The duals of the subsets of $R^{k\times k}$ corresponding to the hypotheses H_0 and H_1 have a remarkably simple structure. Let the symbols H_0 and H_1 also denote the subsets of $R^{k\times k}$ corresponding to the hypotheses H_0 and H_1.

THEOREM 2.1. *The dual, H_0^*, of H_0 is the set of all matrices $\nu = [\nu_{ij}]$ of the form $\nu_{ij} = \delta_i - \delta_j$ for real numbers $\delta_1, \delta_2, \ldots, \delta_k$. The dual, H_1^*, of H_1 is the set of all matrices ν of the form $\nu_{ij} = \delta_i - \delta_j$ for nondecreasing real numbers $\delta_1 \leq \delta_2 \leq \cdots \leq \delta_k$.*

PROOF. Define the k matrices α^t; $t = 1, 2, \ldots, k$ so that the nondiagonal entries in the t-th row of α^t are 1; the nondiagonal entries in the t-th column of α^t are -1; and the other entries in α^t are 0. The linear space, H_0, can be written

$$H_0 = \{\mu : \langle \alpha^t, \mu \rangle = 0 \text{ for } t = 1, 2, \ldots, k\}\,.$$

Note that the matrix $\sum_{t=1}^{k} \alpha^t$ is identically zero so that it is sufficient to use only $k - 1$ constraints in characterizing H_0. Since H_0 is a linear space, its dual H_0^* is the orthogonal complement of H_0 and it follows from the definition of H_0^* that H_0^* is the set of all linear combinations of $\alpha^1, \alpha^2, \ldots, \alpha^k$. Thus,

$$H_0^* = \{\nu : \nu_{ij} = \delta_i - \delta_j \text{ for some } \delta_1, \delta_2, \ldots, \delta_k\}\,. \tag{2.2}$$

Now, $H_0 \subset H_1$, so that it follows from (2.1) that $H_1^* \subset H_0^*$ and thus that matrices in H_1^* are also of the form $[\delta_i - \delta_j]$. It remains to show that $\delta_1 \le \delta_2 \le \cdots \le \delta_k$ is the restriction characterizing H_1^*. Using the definition of α^t, the set H_1 is the closed convex cone,

$$H_1 = \left\{ \mu : \left\langle \sum_{t=1}^{h} \alpha^t, \mu \right\rangle \ge 0; h = 1, 2, \ldots, k-1 \right\}. \tag{2.3}$$

By (2.1), a matrix ν is in H_1^* if and only if it is a linear combination of $\alpha^1, \alpha^1 + \alpha^2, \ldots, \sum_{t=1}^{k-1} \alpha^t$ with nonpositive coefficients $\lambda_1, \lambda_2, \ldots, \lambda_{k-1}$. Using Abel's method of summation by parts and the fact that $\sum_{t=1}^{k} \alpha^t = 0$,

$$\nu = \sum_{h=1}^{k-1} \lambda_h \sum_{i=1}^{h} \alpha^i = \sum_{h=1}^{k} \delta_h \alpha^h\,,$$

with $\lambda_h = \delta_h - \delta_{h+1}; h = 1, 2, \ldots, k-1$. It follows that H_1^* is the closed convex cone

$$H_1^* = \{\nu : \nu_{ij} = \delta_i - \delta_j; \delta_1 \le \delta_2 \le \cdots \le \delta_k\}\,. \tag{2.4}$$ □

The sets in (2.2) and (2.4) are not changed if $\delta = (\delta_1, \delta_2, \ldots, \delta_k)$ is restricted so that $\Sigma \delta_i$ is a fixed value. This observation is useful when computing the MLE's of μ.

Remark. H_0^* and H_1^* can be characterized as follows:

$$H_0^* = \{\nu : \nu_{ij} = \delta_i - \delta_j \text{ with } \Sigma \delta_i = 0\}\,,$$

$$H_1^* = \{\nu : \nu_{ij} = \delta_i - \delta_j \text{ with } \delta_1 \le \delta_2 \le \cdots \le \delta_k \text{ and } \Sigma \delta_i = 0\}\,.$$

3. Maximum likelihood estimates

The purpose of this section is to derive the MLE's for the unknown parameter matrix, μ, under the restrictions $H \cap H_0$ and $H \cap H_1$ where H is any one of the linear models discussed in Section 1. Let $\bar{y} = [\bar{y}_{ij}]$ be the matrix of sample means $\left(\bar{y}_{ij} = \sum_{r=1}^{n} y_{ijr}/n\right)$. If C is any closed convex cone in $R^{k\times k}$, then the MLE of μ subject to the restriction $\mu \in C$ minimizes the sum of squares

$$\sum_{i=1}^{k}\sum_{j=1}^{k}(\bar{y}_{ij} - \mu_{ij})^2 , \tag{3.1}$$

subject to $\mu \in C$. The solution to (3.1) is the closest point of C to the matrix $\bar{y}$ in the least squares sense, and we denote this point by $E(\bar{y}|C)$. One of the keys to finding the MLE's subject to $\mu \in H \cap H_0$ and $\mu \in H \cap H_1$ is the following theorem together with the Remark at the end of Section 2. Moreover, this observation implies that both H_1^* and H_0^* are subsets of any one of the linear subspaces described in Section 1. The proof of the following theorem uses the characterization of $z = E(y|C)$ by the condition, $z \in C$, together with

$$\langle y - z, z\rangle = 0 \tag{3.2}$$

and

$$\langle y - z, w\rangle \le 0 \quad \text{for all} \quad w \in C \tag{3.3}$$

(cf. Theorem 1.3.2 in Robertson *et al.* (1988)). Note that if C is a linear subspace, then $z = E(y|C)$ is characterized by the two conditions, $z \in C$ and

$$\langle y - z, w\rangle = 0 \quad \text{for all} \quad w \in C . \tag{3.4}$$

THEOREM 3.1. *If C is a closed convex cone and H is a linear subspace such that $H \supset C^*$, then*

$$E(y|H \cap C) = E(y|H) - E(y|C^*) ,$$

for all y.

PROOF. Let $u = E(y|H) - E(y|C^*)$. We verify that u has the three properties characterizing $E(y|H \cap C)$. First note that $u \in H$ since H is a linear subspace, $E(y|H) \in H$, and $E(y|C^*) \in C^* \subset H$. In order to see that $u \in C = C^{**}$ suppose $w \in C^*$. Then

$$\begin{aligned}\langle u, w\rangle &= \langle E(y|H) - E(y|C^*), w\rangle \\ &= \langle y - E(y|C^*), w\rangle - \langle y - E(y|H), w\rangle .\end{aligned}$$

The first inner product is nonpositive by (3.3) and the second is zero by (3.4). Thus $u \in H \cap C$. Now consider

$$\begin{aligned}\langle y-u, u\rangle &= \langle y - E(y|H) + E(y|C^*), E(y|H) - E(y|C^*)\rangle \\ &= \langle E(y|C^*), E(y|H) - E(y|C^*)\rangle \\ &= \langle E(E(y|H)|C^*), E(y|H) - E(E(y|H)|C^*)\rangle = 0\end{aligned}$$

by (3.4); by Lemma 2.4 in Lee (1975) (or Lemma 2.2 in Raubertas *et al.* (1986) which states that if C_1 and C_2 are polyhedral cones and $C_1 \subset C_2$, then $E(E(y|C_2)|C_1) = E(y|C_1)$ if either is a linear subspace) and by (3.2). Finally, suppose $w \in H \cap C$ and consider

$$\begin{aligned}\langle y-u, w\rangle &= \langle y - E(y|H) + E(y|C^*), w\rangle \\ &= \langle y - E(y|H), w\rangle + \langle E(y|C^*), w\rangle \le 0 ,\end{aligned}$$

by (3.4), the definition of C^*, and the assumption that $w \in C$. □

Since H_0^* and H_1^* are subsets of H the projections of $\bar{y}$ onto $H \cap H_0$ and $H \cap H_1$ are determined once $E(\bar{y}|H_0^*)$ and $E(\bar{y}|H_1^*)$ are found.

THEOREM 3.2. *If y is any matrix, then*

$$E(y|H_0^*)_{ij} = (y_{i+} - y_{+i})/(2k) - (y_{j+} - y_{+j})/(2k) .$$

PROOF. Let $z_{ij} = \delta_i - \delta_j$ with $\delta_i = (y_{i+} - y_{+i})/(2k)$; $i = 1, 2, \ldots, k$. By Theorem 2.1, $z \in H_0^*$. Moreover, if $w \in H_0^*$ with $w_{ij} = \alpha_i - \alpha_j$ where $\sum_{i=1}^{k} \alpha_i = 0$, then

$$\begin{aligned}\langle y - z, w\rangle &= \sum_{i=1}^{k}\sum_{j=1}^{k} (y_{ij} - z_{ij}) w_{ij} \\ &= \left[\sum_{i=1}^{k} \alpha_i (y_{i+} + y_{+i}) - \sum_{j=1}^{k} \alpha_j (y_{j+} + y_{+j})\right]\Big/2 = 0 .\end{aligned}$$

Thus $z = E(y|H_0^*)$ by (3.4). □

Let $d = (d_1, d_2, \ldots, d_k)$ be defined by

$$d_i = (\bar{y}_{i+} - \bar{y}_{+i})/(2k), \qquad i = 1, 2, \ldots, k \,. \tag{3.5}$$

Clearly, $E(\bar{y}|H_0^*)_{ij} = d_i - d_j$ and

$$\sum_{i=1}^{k} d_i = 0 \,. \tag{3.6}$$

If $g = (g_1, g_2, \ldots, g_k)$ is any k-dimensional vector, let g^* be the isotonic regression of g whose components are in a nondecreasing order. Specifically, g^* solves

$$\min \sum_{i=1}^{k} (g_i - f_i)^2, \qquad f_1 \le f_2 \le \cdots \le f_k$$

(see Section 1.2 of Robertson *et al.* (1988) for computation algorithms and properties of g^*).

THEOREM 3.3. *If y is any matrix and if the k-dimensional vector g is defined by*

$$g_i = (y_{i+} - y_{+i})/(2k), \qquad i = 1, 2, \ldots, k \,,$$

then $E(y|H_1^*) = z$ with $z_{ij} = g_i^* - g_j^*$.

PROOF. Since $g_1^* \le g_2^* \le \cdots \le g_k^*$, $z \in H_1^*$, by Theorem 2.1. Also by Theorem 1.3.2 of Robertson *et al.* (1988),

$$\sum_{i=1}^{k} (g_i - g_i^*) g_i^* = 0 \tag{3.7}$$

and

$$\sum_{i=1}^{k} (g_i - g_i^*) f_i \le 0 \qquad \text{for} \qquad f_1 \le f_2 \le \cdots \le f_k \,. \tag{3.8}$$

Then (3.7) and (3.8) imply that $\sum_{i=1}^{k} g_i = \sum_{i=1}^{k} g_i^* = 0$ and $\langle y - z, z \rangle = 0$, and if $u \in H_1^*$ with $u_{ij} = f_i - f_j$ and $f_1 \le f_2 \le \cdots \le f_k$, then (3.8) implies that $\langle y - z, u \rangle \le 0$. Thus by (3.2) and (3.3), $z = E(y|H_1^*)$. □

This theorem implies that the MLE of μ subject to the restriction H_1^* is given by $E(\bar{y}|H_1^*)_{ij} = d_i^* - d_j^*$ where d is given by (3.5).

Let $\hat{y}$ be the maximum likelihood estimate of μ under H where H is any one of the linear models described in Section 1. Expressions for $\hat{y}$ are well known under any of these models. Combining the above theorems yields the following result.

THEOREM 3.4. *If $\hat{y}$ is the MLE of μ under the linear model H and if $\hat{y}^0$ and $\hat{y}^1$ are the MLE's of μ under $H \cap H_0$ and $H \cap H_1$, respectively, then*

$$\hat{y}^0_{ij} = E(\bar{y}|H \cap H_0)_{ij} = E(\bar{y}|H)_{ij} - E(\bar{y}|H_0^*)_{ij} = \hat{y}_{ij} - d_i + d_j \tag{3.9}$$

and

$$\hat{y}^1_{ij} = E(\bar{y}|H \cap H_1)_{ij} = E(\bar{y}|H)_{ij} - E(\bar{y}|H_1^*)_{ij} = \hat{y}_{ij} - d_i^* + d_j^* , \tag{3.10}$$

where d is given by (3.5) and d^* is the isotonic regression of d whose components are in a monotone nondecreasing order.

4. Distributions of quadratic forms

Consider the matrices $\alpha^t; t = 1,2,\dots,k$ defined in Section 2. Clearly, each of these matrices belongs to H_0^* and by assumption $H_0^* \subset H$. Thus, since $\hat{y} = E(\bar{y}|H)$,

$$\sum_{i=1}^{k}\sum_{j=1}^{k}(\bar{y}_{ij} - \hat{y}_{ij})\alpha^t_{ij} = 0, \qquad t = 1,2,\dots,k ,$$

which implies that

$$\hat{y}_{t+} - \hat{y}_{+t} = \bar{y}_{t+} - \bar{y}_{+t}, \qquad t = 1,2,\dots,k . \tag{4.1}$$

LEMMA 4.1. *The two projections $\hat{y}^0$ and $E(\hat{y}^1|H_0)$ are equal.*

PROOF. Note that

$$E(\hat{y}^1|H_0) = \hat{y}^1 - E(\hat{y}^1|H_0^*)$$

and using Theorem 3.2

$$E(\hat{y}^1|H_0)_{ij} = \hat{y}^1_{ij} - [(\hat{y}^1_{i+} - \hat{y}^1_{+i}) - (\hat{y}^1_{j+} - \hat{y}^1_{+j})]/(2k) . \tag{4.2}$$

However,

$$\hat{y}^1_{i+} = \sum_{j=1}^{k}(\hat{y}_{ij} - d_i^* + d_j^*) = \hat{y}_{i+} - kd_i^*$$

since $\sum_{j=1}^{k} d_j^* = 0$. Similarly, $\hat{y}^1_{+i} = \hat{y}_{+i} + kd_i^*$ and substituting into (4.2), we obtain

$$E(\hat{y}^1 \mid H_0)_{ij} = \hat{y}_{ij} - (\hat{y}_{i+} - \hat{y}_{+i})/(2k) + (\hat{y}_{j+} - \hat{y}_{+j})/(2k) .$$

By (4.1) and (3.9) this is equal to $\hat{y}^0$. □

Now consider the sum of squares,

$$\sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{n}(y_{ijr} - \hat{y}_{ij}^0)^2 = \sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{n}(y_{ijr} - \hat{y}_{ij}^1)^2 + 2\sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{n}(y_{ijr} - \hat{y}_{ij}^1)(\hat{y}_{ij}^1 - \hat{y}_{ij}^0) + n\sum_{i=1}^{k}\sum_{j=1}^{k}(\hat{y}_{ij}^1 - \hat{y}_{ij}^0)^2 . \tag{4.3}$$

The inner product term is equal to

$$2n\sum_{i=1}^{k}\sum_{j=1}^{k}(\bar{y}_{ij} - \hat{y}_{ij}^1)(\hat{y}_{ij}^1 - \hat{y}_{ij}^0) = 2n\sum_{i=1}^{k}\sum_{j=1}^{k}(\bar{y}_{ij} - \hat{y}_{ij}^1)\hat{y}_{ij}^1 - 2n\sum_{i=1}^{k}\sum_{j=1}^{k}(\bar{y}_{ij} - \hat{y}_{ij}^1)\hat{y}_{ij}^0 .$$

The first term is zero by (3.2) and the second term is equal to

$$-2n\sum_{i=1}^{k}\sum_{j=1}^{k}(\bar{y}_{ij} - \hat{y}_{ij}^0)\hat{y}_{ij}^0 - 2n\sum_{i=1}^{k}\sum_{j=1}^{k}(\hat{y}_{ij}^0 - \hat{y}_{ij}^1)\hat{y}_{ij}^0 .$$

The first term is zero by (3.2) and the second term is zero by Lemma 4.1 together with (3.2). Thus, the inner product term in (4.3) is zero and the sum of squares is partitioned as follows:

$$\sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{n}(y_{ijr} - \hat{y}_{ij}^0)^2 = \sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{r=1}^{n}(y_{ijr} - \hat{y}_{ij}^1)^2 + n\sum_{i=1}^{k}\sum_{j=1}^{k}(\hat{y}_{ij}^1 - \hat{y}_{ij}^0)^2 . \tag{4.4}$$

Now consider the problem of testing the null hypothesis, $H \cap H_0$ against the alternative $H \cap H_1 - H \cap H_0$ (i.e., $H \cap H_1$ but not $H \cap H_0$). If Λ_{01} is the likelihood ratio and σ^2 is known, then the LRT rejects for large values of $T_{01} = -2 \ln \Lambda_{01}$ and by (4.4)

$$T_{01} = (n/\sigma^2)\sum_{i=1}^{k}\sum_{j=1}^{k}(\hat{y}_{ij}^1 - \hat{y}_{ij}^0)^2 .$$

Using the expressions for $\hat{y}^1$ and $\hat{y}^0$ given in Theorem 3.4

$$T_{01} = (n/\sigma^2) \sum_{i=1}^{k} \sum_{j=1}^{k} [(d_i - d_j) - (d_i^* - d_j^*)]^2$$
$$= (2kn/\sigma^2) \sum_{i=1}^{k} (d_i - d_i^*)^2$$

since $\sum_{i=1}^{k} (d_i - d_i^*) = 0$.

Clearly, the random vector d has a multivariate normal distribution. If H_0 is true, then the mean vector is zero and the covariance matrix is $\sigma^2(I - k^{-1}J)/(2kn)$ where I is the $k \times k$ identity matrix and J is the $k \times k$ matrix with each entry one. Let $u_1, u_2, \ldots, u_k$ be independent normal random variables with zero means and common variance $\sigma^2/(2kn)$, and let $\bar{u} = \sum_{i=1}^{k} u_i/k$. The two random vectors d and $x = (u_1 - \bar{u}, u_2 - \bar{u}, \ldots, u_k - \bar{u})$ are identically distributed. Suppose x^* is the isotonic regression of x with nondecreasing components. Then $x^* = u^* - \bar{u}$ and the random variable T_{01} has the same distribution as

$$T'_{01} = (2kn/\sigma^2) \sum_{i=1}^{k} (u_i - u_i^*)^2 .$$

The distribution of T'_{01} is given by Corollary 2.6 in Robertson and Wegman (1978), which proves the following theorem.

THEOREM 4.1. *If H_0 is satisfied, then*

$$P[T_{01} \geq t] = \sum_{l=1}^{k} P(l,k) P[\chi^2_{k-l} \geq t] ,$$

where χ_i^2 is a standard chi-square variable with i degrees of freedom ($\chi_0^2 = 0$) and the $P(l,k)$ are the equal-weights level probabilities given by Corollary A *on p.* 81 *of Robertson et al.* (1988).

Now consider the problem of testing $H \cap H_1$ as a null hypothesis and again assume σ^2 is known. Using (4.1) and Theorem 3.3 one shows that $E(\hat{y}|H_1^*) = E(\bar{y}|H_1^*)$ and then by Theorem 3.1, we obtain the following lemma.

LEMMA 4.2. *The two projections $\hat{y}^1 = E(\bar{y}|H \cap H_1)$ and $E(\hat{y}|H_1)$ are equal.*

If Λ_{12} is the likelihood ratio for testing $H \cap H_1$ against $H - H_1$ and $T_{12} = -2 \ln \Lambda_{12}$, then an argument similar to that leading to (4.4) yields

$$
\begin{aligned}
T_{12} &= \left[\sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{r=1}^{n} (y_{ijr} - \hat{y}_{ij}^{1})^2 - \sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{r=1}^{n} (y_{ijr} - \hat{y}_{ij})^2 \right] \Big/ \sigma^2 \\
&= \left[n \sum_{i=1}^{k} \sum_{j=1}^{k} (\hat{y}_{ij} - \hat{y}_{ij}^{1})^2 + 2 \sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{r=1}^{n} (y_{ijr} - \hat{y}_{ij})(\hat{y}_{ij} - \hat{y}_{ij}^{1}) \right] \Big/ \sigma^2 \\
&= (n/\sigma^2) \sum_{i=1}^{k} \sum_{j=1}^{k} (\hat{y}_{ij} - \hat{y}_{ij}^{1})^2 = (n/\sigma^2) \sum_{i=1}^{k} \sum_{j=1}^{k} (d_i^* - d_j^*)^2 \\
&= (2kn/\sigma^2) \sum_{i=1}^{k} (d_i^*)^2 ,
\end{aligned}
$$

recall that $\Sigma d_i^* = \Sigma d_i = 0$. Modifying the argument given for Theorem 4.1, the distribution of T_{12} can be shown to be the same as that of

$$(2kn/\sigma^2) \sum_{i=1}^{k} (u_i^* - \bar{u})^2 .$$

The distribution of this random variable under H_0 is given by the corollary to Theorem 2.3.1 of Robertson *et al.* (1988). Moreover, by the first corollary to Theorem 3.6 of Raubertas *et al.* (1986), $H \cap H_0$ is least favorable within $H \cap H_1$ for this test statistic. Denoting probability with a mean vector μ, by $P_\mu(\cdot)$, we obtain the following theorem.

THEOREM 4.2. *For any real* t,

$$\sup_{\mu \in H \cap H_1} P_\mu[T_{12} \geq t] = \sup_{\mu \in H \cap H_0} P_\mu[T_{12} \geq t] = \sum_{l=1}^{k} P(l,k) P[\chi_{l-1}^2 \geq t] ,$$

with $P(l,k)$ *as in Theorem* 4.1.

σ^2 *Unknown.* First consider the problem of testing $H \cap H_0$ against the alternative $H \cap H_1 - H \cap H_0$. If Λ_{01} is the likelihood ratio, then Λ_{01} can be written as

$$\Lambda_{01} = (\hat{\sigma}_1^2 / \hat{\sigma}_0^2)^{(N/2)} ,$$

where $N = k^2 n$, or $N = k(k-1)n$ if the model has no diagonal elements, and $\hat{\sigma}_1^2$ and $\hat{\sigma}_0^2$ are the MLE's of σ^2 under $H \cap H_1$ and $H \cap H_0$, respectively. Specifically, applying (4.4) and its analogue developed in the proof of Theorem 4.2,

$$
\begin{aligned}
(4.5) \qquad \hat{\sigma}_1^2 / \sigma^2 &= \sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{r=1}^{n} (y_{ijr} - \hat{y}_{ij}^{1})^2 / (N\sigma^2) \\
&= \left[\sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{r=1}^{n} (y_{ijr} - \hat{y}_{ij})^2 / \sigma^2 + T_{12} \right] \Big/ N
\end{aligned}
$$

and

$$\hat{\sigma}_0^2/\sigma^2 = \left[\sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{r=1}^{n} (y_{ijr} - \hat{y}_{ij})^2/\sigma^2 + T_{12} + T_{01} \right] \Big/ N . \tag{4.6}$$

The LRT rejects $H \cap H_0$ for small values of Λ_{01} or equivalently for large values of $\bar{E}_{01}^2 = 1 - \Lambda_{01}^{(2/N)}$. This statistic can be written

$$\bar{E}_{01}^2 = \frac{T_{01}}{T_{01} + T_{12} + \sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{r=1}^{n} (y_{ijr} - \hat{y}_{ij})^2/\sigma^2}$$

$$= \frac{2kn \sum_{i=1}^{k} (d_i - d_i^*)^2}{2kn \sum_{i=1}^{k} (d_i - d_i^*)^2 + 2kn \sum_{i=1}^{k} (d_i^*)^2 + \sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{r=1}^{n} (y_{ijr} - \hat{y}_{ij})^2} .$$

This statistic is similar to the two statistics discussed in Theorem 2.3.1 of Robertson *et al.* (1988) and Theorem 2.7 of Robertson and Wegman (1978). While it is not equal to either one, a careful comparison shows that $\bar{E}_{01}^2$ is the statistic discussed in Theorem 2.7 of Robertson and Wegman with the roles of T_{01} and T_{12} interchanged. The term $\sum_i \sum_j \sum_r (y_{ijr} - \hat{y}_{ij})^2/\sigma^2$ has a chi-squared distribution with γ degrees of freedom, where γ depends on the model, H, of interest. The values of γ are well known for the models considered here. Furthermore by (4.1), this term is independent of T_{01} and T_{12}. For $\nu = 1, 2, \dots$, let $Q(\nu)$ be a sequence of chi-square variables which are independent of T_{01} and T_{12} (it may be necessary to enlarge the probability space to construct such a sequence). Define

$$\bar{E}_{01}^2(\nu) = T_{01}/[T_{01} + T_{12} + Q(\nu)] ,$$

and note that $\bar{E}_{01}^2 \overset{D}{=} \bar{E}_{01}^2(\gamma)$.

An argument like that given for Theorem 2.3.1 of Robertson *et al.* (1988) and Theorems 2.5 and 2.7 of Robertson and Wegman (1978) yields the null hypothesis distribution of $\bar{E}_{01}^2(\nu)$.

THEOREM 4.3. *Let ν be a positive integer. Under the hypothesis $H \cap H_0$,*

$$P[\bar{E}_{01}^2(\nu) \geq t] = \sum_{l=1}^{k} P(l, k) P[B_{(1/2)(k-l),\,(1/2)(\nu+l-1)} \geq t] ,$$

where $B_{a,b}$ is a Beta random variable having the standard Beta distribution with parameters a, b and $B_{0,b} = 0$. The level probabilities $P(l, k)$ are the same as those given in Theorem 4.1.

It is more convenient to table the critical values for an increasing function of $\bar{E}_{01}^2(\nu)$. Define

$$S_{01}(\nu) = \nu\bar{E}_{01}^2(\nu)/[1 - \bar{E}_{01}^2(\nu)]$$

and note that $S_{01}(\nu) = \nu T_{01}/[T_{12} + Q(\nu)]$. As $\nu \to \infty$, T_{12}/ν converges to zero at each point in the probability space and $Q(\nu)/\nu \xrightarrow{D} 1$ and thus $S_{01}(\nu) \xrightarrow{D} T_{01}$ as $\nu \to \infty$. Tables 1 and 2 contain the $\alpha = 0.05$ and $\alpha = 0.01$ critical values of

Table 1. Critical values of $S_{01}(\nu)$ with significance level $\alpha = 0.05$.

ν	k							
	3	4	5	6	7	8	9	10
1	53.026	56.636	58.065	58.957	59.707	60.475	61.263	62.133
2	14.413	18.706	22.254	25.363	28.201	30.825	33.300	35.653
3	9.652	12.990	15.959	18.696	21.272	23.727	26.084	28.357
4	7.950	10.817	13.434	15.899	18.253	20.527	22.733	24.883
5	7.093	9.684	12.081	14.362	16.564	18.700	20.785	22.831
6	6.580	8.992	11.240	13.393	15.478	17.515	19.507	21.471
7	6.239	8.528	10.669	12.724	14.725	16.681	18.603	20.496
8	5.997	8.194	10.253	12.237	14.169	16.063	17.927	19.767
9	5.817	7.942	9.939	11.865	13.743	15.586	17.402	19.197
10	5.675	7.747	9.692	11.572	13.406	15.206	16.983	18.739
11	5.565	7.589	9.494	11.334	13.131	14.898	16.642	18.364
12	5.473	7.461	9.331	11.139	12.904	14.641	16.354	18.051
13	5.397	7.355	9.196	10.975	12.713	14.424	16.113	17.786
14	5.333	7.263	9.080	10.835	12.551	14.238	15.906	17.555
15	5.279	7.186	8.980	10.714	12.410	14.079	15.726	17.358
16	5.231	7.119	8.894	10.610	12.286	13.939	15.568	17.183
17	5.190	7.060	8.818	10.517	12.179	13.814	15.430	17.029
18	5.153	7.007	8.751	10.436	12.084	13.704	15.306	16.893
19	5.121	6.961	8.692	10.363	11.998	13.607	15.195	16.769
20	5.092	6.920	8.638	10.297	11.921	13.519	15.097	16.659
21	5.065	6.882	8.589	10.239	11.851	13.438	15.005	16.557
22	5.042	6.849	8.546	10.186	11.789	13.366	14.924	16.466
23	5.021	6.819	8.506	10.137	11.731	13.300	14.849	16.383
24	5.001	6.790	8.470	10.093	11.679	13.238	14.781	16.307
25	4.984	6.764	8.437	10.052	11.630	13.183	14.717	16.236
26	4.967	6.740	8.406	10.014	11.586	13.131	14.659	16.170
27	4.952	6.719	8.378	9.980	11.544	13.084	14.604	16.110
28	4.939	6.699	8.352	9.947	11.507	13.040	14.554	16.054
29	4.925	6.680	8.327	9.917	11.471	12.999	14.507	16.000
30	4.913	6.663	8.304	9.890	11.437	12.959	14.463	15.952
40	4.826	6.537	8.141	9.688	11.198	12.682	14.146	15.594
60	4.742	6.415	7.980	9.489	10.958	12.404	13.827	15.235
120	4.658	6.294	7.821	9.290	10.721	12.125	13.506	14.872
200	4.626	6.246	7.758	9.212	10.627	12.014	13.379	14.726
∞	4.578	6.175	7.665	9.095	10.485	11.846	13.185	14.505

$S_{01}(\nu)$ for $k = 3, 4, \ldots, 10$ and various values of ν. With $S_{01} = \gamma\bar{E}_{01}^2/[1 - \bar{E}_{01}^2]$, the $\alpha = 0.05$ and $\alpha = 0.01$ critical values of S_{01} are found in these tables with $\nu = \gamma$. The rows labeled $\nu = \infty$ give critical values for T_{01}.

For testing $H \cap H_1$ against $H - H_1$, the likelihood ratio, Λ_{12} can be written

$$\Lambda_{12} = (\hat{\sigma}^2 / \hat{\sigma}_1^2)^{(N/2)},$$

where $\hat{\sigma}_1^2$ is given in (4.5) and

Table 2. Critical values of $S_{01}(\nu)$ with significance level $\alpha = 0.01$.

ν	k							
	3	4	5	6	7	8	9	10
1	1157.754	1104.000	1021.901	947.938	886.871	835.811	795.185	760.493
2	71.716	86.799	98.204	107.770	116.077	123.541	130.388	136.838
3	31.517	40.002	47.308	53.954	60.014	65.839	71.246	76.532
4	21.441	27.538	33.026	38.145	42.979	47.603	52.061	56.401
5	17.188	22.126	26.654	30.930	35.039	38.995	42.850	46.587
6	14.889	19.155	23.118	26.877	30.490	34.026	37.459	40.856
7	13.465	17.298	20.864	24.284	27.586	30.801	33.960	37.077
8	12.506	16.041	19.329	22.496	25.574	28.571	31.503	34.418
9	11.815	15.126	18.216	21.192	24.077	26.912	29.692	32.443
10	11.289	14.432	17.371	20.195	22.952	25.648	28.298	30.909
11	10.882	13.893	16.710	19.412	22.056	24.631	27.183	29.701
12	10.562	13.461	16.175	18.787	21.323	23.825	28.280	28.706
13	10.293	13.108	15.736	18.267	20.733	23.149	25.530	27.897
14	10.074	12.809	15.370	17.831	20.235	22.583	24.904	27.195
15	9.885	12.558	15.059	17.465	19.800	22.101	24.372	26.612
16	9.726	12.346	14.790	17.141	19.434	21.686	23.912	26.105
17	9.587	12.159	14.558	16.868	19.116	21.323	23.507	25.654
18	9.469	11.995	14.358	16.630	18.836	21.010	23.149	25.268
19	9.362	11.852	14.180	16.415	18.587	20.726	22.830	24.919
20	9.268	11.721	14.020	16.224	18.370	20.474	22.556	24.607
21	9.181	11.607	13.874	16.052	18.169	20.255	22.300	24.325
22	9.105	11.505	13.745	15.902	17.996	20.049	22.075	24.070
23	9.038	11.414	13.628	15.761	17.836	19.861	21.866	23.846
24	8.971	11.329	13.528	15.631	17.685	19.700	21.680	23.639
25	8.919	11.254	13.431	15.514	17.551	19.541	21.503	23.445
26	8.865	11.181	13.341	15.414	17.423	19.404	21.347	23.269
27	8.813	11.115	13.259	15.316	17.317	19.273	21.199	23.110
28	8.771	11.056	13.181	15.221	17.206	19.153	21.064	22.961
29	8.730	11.001	13.115	15.141	17.112	19.042	20.941	22.814
30	8.690	10.946	13.048	15.061	17.017	18.937	20.824	22.687
40	8.423	10.584	12.589	14.509	16.375	18.197	19.993	21.762
60	8.163	10.232	12.144	13.974	15.747	17.478	19.182	20.855
120	7.913	9.890	11.718	13.457	15.140	16.780	18.388	19.965
200	7.813	9.759	11.550	13.257	14.901	16.505	18.076	19.618
∞	7.673	9.565	11.305	12.958	14.550	16.098	17.611	19.096

$$\hat{\sigma}^2 = \sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{r=1}^{n} (y_{ijr} - \hat{y}_{ij})^2 / N .$$

The LRT rejects for large values of

$$\bar{E}_{12}^2 = 1 - \Lambda_{12}^{2/N} = \frac{T_{12}}{T_{12} + \sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{r=1}^{n} (y_{ijr} - \hat{y}_{ij})^2 / \sigma^2} .$$

Table 3. Critical values of $S_{12}(\nu)$ with significance level $\alpha = 0.05$.

ν	k							
	3	4	5	6	7	8	9	10
1	93.288	140.680	181.647	217.226	248.839	276.916	302.320	325.354
2	14.978	19.831	23.710	26.947	29.721	32.151	34.307	36.257
3	8.947	11.376	13.265	14.822	16.136	17.281	18.292	19.200
4	7.068	8.815	10.158	11.252	12.174	12.974	13.674	14.304
5	6.181	7.624	8.724	9.616	10.365	11.012	11.580	12.084
6	5.670	6.943	7.910	8.690	9.343	9.907	10.399	10.839
7	5.339	6.505	7.388	8.096	8.691	9.201	9.648	10.045
8	5.107	6.201	7.024	7.686	8.239	8.713	9.128	9.497
9	4.936	5.978	6.759	7.386	7.908	8.356	8.749	9.097
10	4.806	5.806	6.556	7.156	7.656	8.084	8.459	8.791
11	4.702	5.670	6.395	6.975	7.457	7.871	8.232	8.551
12	4.618	5.561	6.265	6.829	7.297	7.698	8.048	8.358
13	4.549	5.470	6.159	6.708	7.165	7.555	7.896	8.199
14	4.491	5.394	6.069	6.607	7.053	7.436	7.770	8.065
15	4.441	5.330	5.993	6.521	6.960	7.335	7.661	7.952
16	4.398	5.274	5.927	6.447	6.879	7.247	7.569	7.854
17	4.360	5.226	5.870	6.382	6.808	7.171	7.488	7.769
18	4.327	5.183	5.820	6.326	6.746	7.105	7.418	7.695
19	4.299	5.146	5.775	6.276	6.691	7.046	7.355	7.629
20	4.273	5.112	5.736	6.231	6.643	6.994	7.300	7.571
21	4.250	5.082	5.700	6.191	6.599	6.947	7.249	7.518
22	4.229	5.055	5.668	6.156	6.559	6.904	7.205	7.470
23	4.209	5.030	5.639	6.123	6.524	6.866	7.163	7.428
24	4.192	5.008	5.613	6.093	6.491	6.831	7.126	7.388
25	4.176	4.988	5.589	6.066	6.462	6.799	7.093	7.353
26	4.161	4.968	5.666	6.041	6.434	6.769	7.062	7.320
27	4.148	4.951	5.546	6.017	6.409	6.743	7.033	7.290
28	4.136	4.935	5.527	5.997	6.385	6.717	7.007	7.262
29	4.124	4.920	5.510	5.977	6.364	6.695	6.982	7.237
30	4.114	4.907	5.493	5.959	6.344	6.673	6.959	7.212
40	4.037	4.808	5.377	5.829	6.202	6.520	6.796	7.042
60	3.963	4.712	5.265	5.703	6.063	6.371	6.639	6.876
120	3.891	4.619	5.155	5.579	5.930	6.228	6.487	6.715
200	3.862	4.582	5.113	5.532	5.877	6.172	6.427	6.653
∞	3.820	4.528	5.049	5.460	5.800	6.088	6.339	6.560

With $\{Q(\nu)\}$ defined as before let $\bar{E}_{12}^2(\nu) = T_{12}/[T_{12} + Q(\nu)]$ and $S_{12}(\nu) = \nu\bar{E}_{12}^2(\nu)/[1 - \bar{E}_{12}^2(\nu)]$. The null hypothesis distribution of $\bar{E}_{12}^2(\nu)$ is given in the following theorem.

THEOREM 4.4. *For ν a positive integer*

$$\sup_{\mu \in H \cap H_1} P_\mu[\bar{E}_{12}^2(\nu) \geq t] = \sup_{\mu \in H \cap H_0} P_\mu[\bar{E}_{12}^2(\nu) \geq t]$$

Table 4. Critical values of $S_{12}(\nu)$ with significance level $\alpha = 0.01$.

ν	k							
	3	4	5	6	7	8	9	10
1	2355.351	3541.487	4558.026	5460.333	6240.542	6943.212	7597.377	8159.125
2	81.589	106.498	126.494	142.981	157.056	169.546	180.536	190.735
3	31.181	38.689	44.535	49.344	53.366	56.940	60.014	62.798
4	20.183	24.420	27.690	30.348	32.571	34.505	36.206	37.716
5	15.813	18.856	21.172	23.055	24.617	25.983	27.176	28.220
6	13.536	15.982	17.837	19.336	20.597	21.676	22.610	23.450
7	12.153	14.270	15.846	17.124	18.195	19.113	19.909	20.622
8	11.241	13.127	14.536	15.668	16.619	17.421	18.131	18.760
9	10.588	12.315	13.609	14.646	15.502	16.241	16.888	17.464
10	10.103	11.712	12.921	13.883	14.675	15.362	15.967	16.494
11	9.725	11.250	12.387	13.289	14.045	14.688	15.256	15.755
12	9.426	10.883	11.965	12.824	13.547	14.159	14.691	15.163
13	9.182	10.582	11.618	12.453	13.140	13.724	14.237	14.690
14	8.979	10.329	11.340	12.137	12.796	13.365	13.857	14.297
15	8.809	10.124	11.100	11.877	12.515	13.068	13.544	13.967
16	8.665	9.945	10.892	11.652	12.273	12.807	13.277	13.688
17	8.539	9.792	10.720	11.456	12.068	12.587	13.040	13.447
18	8.426	9.655	10.566	11.292	11.886	12.397	12.842	13.234
19	8.331	9.538	10.434	11.141	11.730	12.230	12.662	13.046
20	8.243	9.436	10.313	11.013	11.587	12.075	12.508	12.886
21	8.168	9.341	10.205	10.893	11.459	11.944	12.365	12.738
22	8.098	9.256	10.114	10.792	11.350	11.826	12.244	12.605
23	8.035	9.180	10.026	10.694	11.245	11.718	12.126	12.490
24	7.979	9.110	9.945	10.608	11.153	11.621	12.022	12.382
25	7.931	9.048	9.877	10.531	11.069	11.532	11.934	12.284
26	7.884	8.991	9.809	10.459	10.997	11.452	11.845	12.198
27	7.838	8.942	9.754	10.394	10.926	11.373	11.770	12.113
28	7.801	8.892	9.695	10.332	10.858	11.307	11.698	12.038
29	7.764	8.847	9.648	10.280	10.800	11.245	11.631	11.967
30	7.728	8.806	9.600	10.226	10.743	11.187	11.570	11.903
40	7.486	8.513	9.268	9.860	10.350	10.772	11.132	11.449
60	7.258	8.234	8.947	9.512	9.977	10.373	10.717	11.019
120	7.037	7.967	8.645	9.182	9.625	9.997	10.321	10.606
200	6.947	7.860	8.528	9.053	9.487	9.850	10.169	10.449
∞	6.822	7.709	8.356	8.865	9.284	9.639	9.946	10.216

$$= \sum_{l=1}^{k} P(l,k) P[B_{(1/2)(l-1),\,(1/2)\nu} \geq t] \,.$$

Because $S_{12}(\nu) = \nu T_{12}/Q(\nu) \xrightarrow{D} T_{12}$ as $\nu \to \infty$, it is more convenient to table critical values for $S_{12}(\nu)$. They are given in Tables 3 and 4 for $\alpha = 0.05, 0.01$, $k = 3, 4, \ldots, 10$ and various values of ν. Of course, the critical value for $S_{12} = \gamma \bar{E}_{12}^2/[1 - \bar{E}_{12}^2]$ can be obtained from these tables with $\nu = \gamma$ and the critical values for T_{12} are found in the $\nu = \infty$ rows.

The authors have been unable to obtain closed form expressions for the appropriate weighted l_2 projections of $\bar{y}$ onto $H \cap H_0$ and $H \cap H_1$ for an unbalanced design, i.e., one with cells having different sample sizes. If $p = (p_{ij})$ is a $k \times k$ matrix of multinomial probabilities, then the hypotheses H_0 and H_1 are marginal homogeneity and a stochastic ordering between the marginal distributions, respectively. Tests of these hypotheses are of interest in this setting also. However, the restricted MLE's must be obtained by iterative techniques (cf. Bishop *et al.* (1975)). Weighted l_2 projections could provide reasonable approximations for the restricted MLE's and one would anticipate that test statistic like those studied here could be based on these approximations. Further research in this area is needed.

Acknowledgement

The authors wish to express their appreciation for the referee's helpful comments, particularly the comments which have helped to clarify the meaning of the one-sided alternative, (1.3).

REFERENCES

Barlow, R. E. and Brunk, H. D. (1972). The isotonic regression problem and its dual, *J. Amer. Statist. Assoc.*, **67**, 140–147.

Bishop, Y. M. M., Fienberg, S. E. and Holland, P. W. (1975). *Discrete Multivariate Analysis: Theory and Practice*, MIT Press, Cambridge, Massachusetts.

Cohen, A., Perlman, M. D. and Sackrowitz, H. B. (1989). Unbiasedness of tests of homogeneity when alternatives are ordered, *Proc. Conf. on Depend. Statist.*, I. M. S. Monograph Series (to appear).

Lee, C. C. (1975). Conditional expectation given a σ-lattice. Isotonic regression, univariate and multivariate, Ph. D. Dissertation, Oregon State University.

Luenberger, D. G. (1969). *Optimization by Vector Space Methods*, Wiley, New York.

Marshall, A. W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York.

Pukelsheim, F. (1984). A note on nonparametric trend conformity, *Ann. Statist.*, **12**, 775–777.

Raubertas, R. F., Lee, C. C. and Nordheim, E. V. (1986). Hypothesis tests for normal means constrained by linear inequalities, *Comm. Statist. A—Theory Methods*, **15**, 2809–2833.

Robertson, T. and Wegman, E. J. (1978). Likelihood ratio tests for order restrictions in exponential families, *Ann. Statist.*, **6**, 485–505.

Robertson, T. and Wright, F. T. (1981). Likelihood ratio tests for and against a stochastic ordering between multinomial populations, *Ann. Statist.*, **9**, 1248–1257.

Robertson, T. and Wright, F. T. (1982). On measuring the conformity of a parameter set to a trend with applications, *Ann. Statist.*, **10**, 1234–1245.

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*, Wiley, New York.