Journal of the Korean
Statistical Society
Vol. 25, No. 4, 1996

Bounds for the Full Level Probabilities with Restricted Weights and Their Applications

Chul Gyu Park[1]

Abstract

Lower bounds for the full level probabilities are derived under order restrictions in weights. Discussions are made on typical isotonic cones such as linear order, simple tree order, and unimodal order cones. We also discuss applications of these results for constructing conditional likelihood ratio tests for ordered hypotheses in a contingency table. A real data set on torus mandibularis will be analyzed for illustrating the testing procedure.

Key Words : Isotonic regression; Level probability; Likelihood ratio statistic; Least favorable conditional test.

1. INTRODUCTION

Orthant probability is a very important concept in the area of one-sided tests in multivariate probability models. This probability is defined as $P[\mathbf{X} \in \mathcal{O}]$ where $\mathbf{X}$ is a random vector following $N_k(\mathbf{u}, \mathbf{W}^{-1})$ and $\mathcal{O} = \{\mathbf{x} \in \mathbf{R}^k :$

[1]Department of Applied Statistics, University of Suwon, Suwon, Kyunggi-Do 445-743, Korea.

$x_1 \geq 0,\ x_2 \geq 0, \cdots, x_k \geq 0\}$. There is a vast literature on the computation of orthant probabilities: see Gupta(1963), Sun(1988), Kepner et al.(1989), and Nomakuchi and Shi(1992). Even though orthant probabilities are expressed in closed forms for very particular structures of $\mathbf{W}^{-1}$, it is generally the case that no closed expressions are known to us unless $k \leq 4$. Another concept related with orthant probability is level probability which appears in chi-bar-square distributions. We begin with discussions on some preliminary concepts and results regarding level probabilities.

Consider a subset $\mathcal{C}$ of Euclidian space $\mathbf{R}^k$ generated by all possible isotonic functions with respect to a partial order $\preceq$ on index set $\mathcal{N} = \{1, 2, \cdots, k\}$. The set $\mathcal{C}$ is called an isotonic cone associated with partial order $\preceq$ on $\mathcal{N}$. Typical examples of isotonic cones are simple linear order, simple tree order, and unimodal order cones which are of types $\{\mathbf{x} \in \mathbf{R}^k : x_1 \leq x_2 \leq \cdots \leq x_k\}$, $\{\mathbf{x} \in \mathbf{R}^k : x_1 \leq x_j, j = 2, 3, \cdots, k\}$, and $\{\mathbf{x} \in \mathbf{R}^k : x_1 \leq \cdots \leq x_h \geq \cdots \geq x_k\}$, respectively. Let $\mathbf{X}$ follow $N_k(\mathbf{u}, \mathbf{W}^{-1})$ where $\mathbf{W} = \text{diag}\{w_1, w_2, \cdots, w_k\}$ is known. Then, the unrestricted maximum likelihood estimator, $\hat{\mathbf{u}}$, of $\mathbf{u}$ is $\mathbf{X}$. It is also well known that the maximum likelihood estimator(MLE) of $\mathbf{u}$ constrained by the cone $\mathcal{C}$ is the isotonic regression of $\hat{\mathbf{u}}$ with weight vector $\mathbf{w} = (w_1, w_2, \cdots, w_k)'$. The isotonic regression of $\hat{\mathbf{u}}$ is the least squares projection of $\hat{\mathbf{u}}$ onto the cone $\mathcal{C}$, and is usually denoted by $\mathbf{u}^* = P_{\mathbf{w}}(\hat{\mathbf{u}}|\mathcal{C})$.

Let T_{12} be the likelihood ratio test(LRT) statistic for testing $H_1 : \mathbf{u} \in \mathcal{C}$ against $H_2 - H_1$ where H_2 imposes no restriction on $\mathbf{u}$. Then, the LRT statistic is expressed as $T_{12} = \sum_{i=1}^{k}(u_i^* - \hat{u}_i)^2 w_i$. Under the composite null hypothesis H_1, the distribution of T_{12} depends on $\mathbf{u}$ in H_1 and is unknown. However, in order to guarantee the desired significance level α, we usually appeal to the least favorable distribution of T_{12} which is attainable at $u_1 = u_2 = \cdots = u_k$. This particular distribution is given by

$$P[T_{12} > t] = \sum_{l=1}^{k} p(l, k; \mathbf{w}) P[\chi^2_{k-l} > t],\ t > 0, \tag{1.1}$$

where χ^2_d is a usual chi-square variable with d degrees of freedom. The level probability, denoted by $p(l, k; \mathbf{w})$ in (??), is interpreted as the probability that $\mathbf{u}^*$ takes on exactly l distinct levels. These probabilities depend not only on the type of isotonic cone but also on the structure of weight vector $\mathbf{w}$. In general, the level probabilities are unknown even though we have a closed form of recursive expressions in the simple linear order case with equal weights. Refer to Robertson et al.(1988) for details.

Let M denote the number of distinct component values in $\mathbf{u}^*$. Then, as is well known, the conditional variable T_{12} , given $M = l$, follows a chi-square distribution with $k-l$ degrees of freedom, regardless of the type of isotonic cone and the structure of weight vector. From this fact, one might suggest a conditional chi-square test for H_1, based on the outcome of M. However, since $T_{12}=0$ if $M=k$, the use of chi-square test of size α^* given the outcome of M results in the overall size $\alpha^*(1-P[M=k])$. Thus, in order to meet the desired size α, each conditional test given the outcome of M should be conducted at level $\alpha^*=\alpha/(1-P[M=k])$. Therefore, the computation of the full level probability $P[M=k]=p(k,k;\mathbf{w})$ is crucial in constructing the conditional test.

As is often the case in the asymptotic analysis of ordered categorical data, the weight vector might be dependent on the underlying parameter values, and possibly has an ordering among its components in accordance with the null hypothesis. In this situation, we should find the least favorable parameter configuration within the null hypothesis to guarantee the desired significance level. This problem motivates us to investigate the lower bound for the probability $p(k,k;\mathbf{w})$ whose the weight vector has an order restriction. The following section derives the lower bounds for the level probabilities in the simple linear order, simple tree order, and unimodal order cases. Section 3 discusses applications of these results in the analysis of $2\times k$ contingency tables, and provides an example of analyzing real data on torus mandibularis.

2. BOUNDS FOR THE FULL LEVEL PROBABILITIES

In this section, we will derive the lower bound for the level probability, $p(k,k;\mathbf{w})$, when the components of $\mathbf{w}$ are also isotonically ordered. Detailed derivations are provided for the simple linear order case, but the proofs for other order cones will be omitted because the basic ideas are similar. First, it should be noted that $P[M=k]=P[\mathbf{X}\in\mathcal{C}]$ where $\mathbf{X}$ follows $N_k(\mathbf{0},\mathbf{W}^{-1})$ and $\mathbf{W}=\operatorname{diag}\{w_1,w_2,\cdots,w_k\}$. The following lemma states when the level probability is minimized with respect to w_k in the linear order case.

Lemma 1. Let $\mathbf{X}$ follow a k-dimensional multivariate normal distribution $N_k(\mathbf{0},\mathbf{W}^{-1})$ where $\mathbf{W}=\operatorname{diag}\{w_1,w_2,\cdots,w_k\}$ and define $\mathcal{C}=\{\mathbf{x}\in\mathbf{R}^k: x_1\le x_2\le\cdots\le x_k\}$. Then, we have

$$\inf_{w_k>0} P[\mathbf{X}\in\mathcal{C}]=P[\mathbf{X}\in\mathcal{D}]$$

where $\mathcal{D}=\{\mathbf{x}\in\mathbf{R}^k: x_1\le x_2\le\cdots\le x_{k-1}\le 0\}$. This infimum occurs when $w_k\uparrow\infty$.

Proof. Consider a linear transformation $\mathbf{Y} = \mathbf{TX}$ such that

$$Y_i = \begin{cases} (X_{i+1} - X_i)/(w_i^{-1} + w_{i+1}^{-1})^{1/2}, & i = 1, 2, \cdots, k-1 \\ \sum_{j=1}^{k} w_j X_j / (\sum_{j=1}^{k} w_j)^{1/2}, & i = k. \end{cases} \tag{2.1}$$

Then, $\mathbf{Y}$ follows a multivariate normal distribution $N_k(\mathbf{0}, \mathbf{R})$ where $\mathbf{R} = \mathbf{TW}^{-1}\mathbf{T}'$. The upper diagonal elements of $\mathbf{R}$ are

$$r_{ij} = \begin{cases} 1 & \text{if } \ i = j \\ -[\frac{w_i w_{i+2}}{(w_i + w_{i+1})(w_{i+1} + w_{i+2})}]^{1/2} & \text{if } \ j = i+1, i = 1, 2, \cdots, k-2 \\ 0 & \text{if } \ i = k-1, j = k \text{ or if } j > i+1. \end{cases}$$

The image of $\mathcal{C}$ under transformation $\mathbf{T}$ is $\mathbf{T}(\mathcal{C}) = \{\mathbf{y} \in \mathbf{R}^k : y_i \geq 0, i = 1, 2, \cdots, k-1\}$. Now, consider another transformation $\mathbf{Z} = \mathbf{SX}$ such that

$$Z_i = \begin{cases} (X_{i+1} - X_i)/(w_i^{-1} + w_{i+1}^{-1})^{1/2}, & i = 1, 2, \cdots, k-2 \\ -[(\sum_{j=1}^{k-1} w_j X_j + \beta X_{k-1})/\{(w_{k-1}^{-1} + w_k^{-1})^{1/2}\beta\}], & i = k-1 \\ \sqrt{w_k} X_k, & i = k, \end{cases} \tag{2.2}$$

where $\beta = w_k + (w_k \sum_{i=1}^{k} w_i)^{1/2}$. Then, we can show with rather tedious algebra that $\mathbf{SW}^{-1}\mathbf{S}' = \mathbf{R}$, and hence, that $\mathbf{Y}$ and $\mathbf{Z}$ are identically distributed. Let $\mathcal{C}_{w_k} = \mathbf{S}^{-1}(\mathbf{T}(\mathcal{C})) = \{\mathbf{x} \in \mathbf{R}^k : \mathbf{x} = \mathbf{S}^{-1}\mathbf{z}, \mathbf{z} \in \mathbf{T}(\mathcal{C})\} = \{\mathbf{x} \in \mathbf{R}^k : \mathbf{Sx} \in \mathbf{T}(\mathcal{C})\} = \{\mathbf{x} \in \mathbf{R}^k : x_1 \leq x_2 \leq \cdots \leq x_{k-1}, \sum_{j=1}^{k-1} w_j x_j + \beta x_{k-1} \leq 0\}$. Now, define $\mathcal{D}_1 = \{\mathbf{x} \in \mathbf{R}^k : x_1 \leq x_2 \leq \cdots \leq x_{k-1}, \sum_{j=1}^{k-1} w_j x_j + \beta x_{k-1} \leq 0, x_{k-1} \leq 0\}$ and $\mathcal{D}_2 = \{\mathbf{x} \in \mathbf{R}^k : x_1 \leq x_2 \leq \cdots \leq x_{k-1}, \sum_{j=1}^{k-1} w_j x_j + \beta x_{k-1} \leq 0, x_{k-1} > 0\}$. Then, $\mathcal{C}_{w_k}$ is the union of disjoint sets $\mathcal{D}_1$ and $\mathcal{D}_2$. Since $\beta > 0$, $\mathcal{D}_1 = \mathcal{D}$ regardless of w_i's. When w_i's other than w_k are fixed, $\beta \uparrow \infty$ as $w_k \uparrow \infty$. And thus, $\mathcal{D}_2 \downarrow \emptyset$ as $w_k \uparrow \infty$, which implies that $\mathcal{C}_{w_k} \downarrow \mathcal{D}$. Since $P[\mathbf{X} \in \mathcal{C}] = P[\mathbf{Y} \in \mathbf{T}(\mathcal{C})] = P[\mathbf{S}^{-1}\mathbf{Y} \in \mathbf{S}^{-1}\mathbf{T}(\mathcal{C})] = P[\mathbf{X} \in \mathcal{C}_{w_k}]$, the probability $P[\mathbf{X} \in \mathcal{C}]$ converges downward to $P[\mathbf{X} \in \mathcal{D}]$ as $w_k \uparrow \infty$. This completes the proof. □

Lemma 1 is used to prove the following theorem on the lower bound of $P[\mathbf{X} \in \mathcal{C}]$ with linear order restriction on weights.

Theorem 1. Suppose that $\mathbf{X}$ follow a k-dimensional multivariate normal distribution $N_k(\mathbf{0}, \mathbf{W}^{-1})$ where $\mathbf{W} = \text{diag}\{w_1, w_2, \cdots, w_k\}$ and define $\mathcal{C} = \{\mathbf{x} \in \mathbf{R}^k : x_1 \leq x_2 \leq \cdots \leq x_k\}$. Then, we have

$$\inf_{w_1 \leq \cdots \leq w_k} P[\mathbf{X} \in \mathcal{C}] = \frac{1}{2^{k-1}(k-1)!} \tag{2.3}$$

and this infimum is obtained by letting $w_1 = w_2 = \cdots = w_{k-1}$ and $w_k/w_1 \uparrow \infty$.

Proof. Let $w_1, w_2, \cdots, w_{k-1}$ be fixed. Then, by Lemma 1, $\inf_{w_k>0} P[\mathbf{X} \in \mathcal{C}] = P[\mathbf{X} \in \mathcal{D}]$ which is achieved when $w_k \to \infty$. Thus, we can say that $\inf_{w_1\leq\cdots\leq w_k} P[\mathbf{X} \in \mathcal{C}] = \inf_{w_1\leq\cdots\leq w_{k-1}} P[\mathbf{X} \in \mathcal{D}]$. Now, let $\mathbf{Z} = \mathbf{W}^{1/2}\mathbf{X}$ and $\mathcal{D}_{\mathbf{W}} = \mathbf{W}^{1/2}\mathcal{D}$. Then, $\mathbf{Z}$ follows $N_k(\mathbf{0}, \mathbf{I}_k)$ where $\mathbf{I}_k = \text{diag}\{1, 1, \cdots, 1\}$, and we have $\mathcal{D}_{\mathbf{W}} = \{\mathbf{z} \in \mathbf{R}^k : \mathbf{z} = \mathbf{W}^{1/2}\mathbf{x}, \mathbf{x} \in \mathcal{D}\} = \{\mathbf{z} \in \mathbf{R}^k : \mathbf{W}^{-1/2}\mathbf{z} \in \mathcal{D}\} = \{\mathbf{z} \in \mathbf{R}^k : z_1 {w_1}^{-1/2} \leq z_2 {w_2}^{-1/2} \leq \cdots \leq z_{k-1} {w_{k-1}}^{-1/2} \leq 0\}$. Suppose $\mathbf{x} \in \mathcal{D}$. Then, $x_1 \leq x_2 \leq \cdots \leq x_{k-1} \leq 0$. Since $0 < w_i/w_{i+1} \leq 1, i = 1, \cdots, k-1$ and $x_i \leq 0, i = 1, \cdots, k-1$, we have $x_i \leq x_{i+1} \leq (w_i/w_{i+1})^{1/2} x_{i+1}$ and thus $x_i w_i^{-1/2} \leq x_{i+1} w_{i+1}^{-1/2}, i = 1, 2, \cdots, k-2$. This implies that $\mathbf{x} \in \mathcal{D}_{\mathbf{W}}$, and thus, $\mathcal{D} \subset \mathcal{D}_{\mathbf{W}}$. Therefore, it follows that $P[\mathbf{X} \in \mathcal{D}] = P[\mathbf{Z} \in \mathcal{D}_{\mathbf{W}}] \geq P[\mathbf{Z} \in \mathcal{D}]$. Here, the equality holds when $\mathcal{D}_{\mathbf{W}} = \mathcal{D}$, or equivalently $w_1 = w_2 = \cdots = w_{k-1}$. Using the result of Chase(1974), we get $P[\mathbf{Z} \in \mathcal{D}] = [2^{k-1}(k-1)!]^{-1}$. □

The infimum level probability in Theorem 1 is the same as $Q(k,k) = \lim_{w_k\to\infty} p(k,k;\mathbf{w})$ where $w_1 = w_2 = \cdots = w_{k-1} (< \infty)$, and it does not depend on the common value of $w_i, i = 1, 2, \cdots, k-1$. The infima of the full level probabilities for other cones are provided in the following theorems. The proofs are similar to that in the linear order case, and we will omit them.

Theorem 2. Let $\mathbf{X}$ follow a k-dimensional multivariate distribution $N_k(\mathbf{0}, \mathbf{W}^{-1})$ where $\mathbf{W} = \text{diag}\{w_1, w_2, \cdots, w_k\}$ and define $\mathcal{C} = \{\mathbf{x} \in \mathbf{R}^k : x_1 \leq x_i, i = 2, \cdots, k\}$. Then, we have

$$\inf_{w_1 \leq w_i, 2\leq i\leq k} P[\mathbf{X} \in \mathcal{C}] = \frac{1}{k}. \tag{2.4}$$

This infimum is achieved when $w_1 = w_2 = \cdots = w_k$.

Theorem 3. Suppose that $\mathbf{X}$ follow a k-dimensional multivariate normal distribution $N_k(\mathbf{0}, \mathbf{W}^{-1})$ where $\mathbf{W} = \text{diag}\{w_1, w_2, \cdots, w_k\}$ and define $\mathcal{C} = \{\mathbf{x} \in \mathbf{R}^k : x_1 \leq \cdots \leq x_h \geq \cdots \geq x_k\}$. Then,

$$\inf_{w_1\leq\cdots\leq w_h\geq\cdots\geq w_k} P[\mathbf{X} \in \mathcal{C}] = \frac{1}{2^{k-1}(h-1)!(k-1)!}. \tag{2.5}$$

This infimum is obtained by letting $w_1 = \cdots = w_{k-1}, w_{h+1} = \cdots = w_{k-1}$, and $w_h/w_i \uparrow \infty$ for $h \neq i$.

3. APPLICATION TO CONDITIONAL TESTS

This section discusses how the results on the full level probabilities in Section 2 can be used in testing problems. Consider a contingency table with positive cell probabilities $p_{ij} = P[(X,Y) = (i,j)]$, $i = 1,2$, $j = 1,2,\cdots,k$. Let $\theta_{1j} = p_{2j}/[p_{1j} + p_{2j}]$ and $\theta_{2j} = p_{1j} + p_{2j}, j = 1,2,\cdots,k$. Various types of dependence between X and Y discussed by Lehmann(1966) can be defined in terms of order relations of θ_{1j}'s. Since other cases are implicatively straightforward, we focus on the negative regression dependence between X and Y which is expressed as a linear order cone

$$\mathcal{C} = \{\theta_1 : \theta_{11} \geq \theta_{12} \geq \cdots \geq \theta_{1k}\}. \tag{3.1}$$

Most researchers have been interested merely in order relations as given in (??), and have paid little attention to the impact of the shape of marginal distribution of Y. However, as is often the case in clinical data, the marginal distribution of Y may have also a shape restriction. For example, suppose that Y represents the severity of a certain symptom. Since the sources of patients are usually ordered in size according to the level of severity, the marginal probabilities of Y would have a descending order under multinomial sampling scheme. Thus, in this case, the whole parameter space would be limited to a class of contingency tables with $\theta_{21} \geq \theta_{22} \geq \cdots \geq \theta_{2k}$.

In this limited parameter space, we might be interested in testing the goodness-of-fit of the positive regression dependence between X and Y. If we reparameterize by setting $\theta_{1j} = p_{2j}/[p_{1j} + p_{2j}]$ and $\theta_{2j} = p_{1j} + p_{2j}$, the problem becomes that of testing $H_1 = \{\theta : \theta_{11} \geq \theta_{12} \geq \cdots \geq \theta_{1k},\ \theta_{21} \geq \theta_{22} \geq \cdots \geq \theta_{2k}\}$ against $H_2 - H_1$ where $H_2 = \{\theta : \theta_{21} \geq \theta_{22} \geq \cdots \geq \theta_{2k}\}$. Let n_{ij} be the (i,j)th cell frequency under multinomial sampling. The likelihood function is

$$L(\theta) = [\prod_{j=1}^{k} \theta_{1j}^{n_{2j}}(1-\theta_{1j})^{n_{1j}}][\prod_{j=1}^{k} \theta_{2j}^{n_{1j}+n_{2j}}]. \tag{3.2}$$

As is discussed in Example 1.5.1 of Robertson et al.(1988), the MLE of $\theta_1 = (\theta_{11}, \theta_{12}, \cdots, \theta_{1k})'$ under H_1 is the isotonic regression of the unrestricted MLE, $\hat{\theta}_1$, where $\hat{\theta}_{1j} = n_{2j}/[n_{1j} + n_{2j}], j = 1,2,\cdots,k$, with weights $w_j = n_{1j} + n_{2j}$. This isotonic regression, θ_1^*, is the least squares projection of $\hat{\theta}_1$ onto the set $\mathcal{C} = \{\mathbf{x} \in \mathbf{R}^k : x_1 \geq x_2 \geq \cdots \geq x_k\}$. Now, the unrestricted MLE's of θ_{2j}'s are $\hat{\theta}_{2j} = [n_{1j} + n_{2j}]/n, j = 1,2,\cdots,k$, where $n = \sum_{i=1}^{2}\sum_{j=1}^{k} n_{ij}$. As shown in Example 1.5.7 of Robertson et al.(1988), the MLE of the nuisance parameter $\theta_2 = (\theta_{21}, \theta_{22}, \cdots, \theta_{2k})'$ under both hypotheses is also the isotonic regression of $\hat{\theta}_2$ with equal weights.

Based on these estimators, we can construct the likelihood ratio test which rejects H_1 in favor of $H_2 - H_1$ for the large values of

$$T_{12} = -2\sum_{j=1}^{k} [n_{2j} \ln(\theta^*_{1j}/\hat{\theta}_{1j}) + n_{2j} \ln\{(1-\theta^*_{1j})/(1-\hat{\theta}_{1j})\}]. \qquad (3.3)$$

It is well known that the asymptotic distribution of T_{12} under the null hypothesis H_1 is a chi-bar-square distribution which is a certain mixture of chi-square distributions. This complicated null distribution results in much work for computing critical values or p-values.

A simpler approach for this problem is to construct a conditional test based on the outcome of M, the number of levels in θ^*_1. Recall that the asymptotic conditional distribution of T_{12} given $M = l$ is the chi-square distribution with $k - l$ degrees of freedom under H_1. Thus, once the event $[M = l]$ occurs, it would be reasonable to reject H_1 when $T_{12} > \chi^2_{k-l}(\alpha^*)$. Since we can not reject H_1 in the case $M = k$, the overall size of this test will be $\alpha^*(1 - P[M = k])$. Hence, in order to meet the overall size α, we should use $\alpha^* = \alpha/(1 - P[M = k])$ in each conditional test. However, since $P[M = k]$ depends on the parameter θ, we must determine the parameter configuration for which the conditional test is least favorable. This will be possible by finding the infimum of $P[M = k]$ over the parameters in H_1.

First, define $\mathcal{I}_{\theta_1} = \{\mathbf{x} \in \mathbf{R}^k : x_i \geq x_{i+1} \text{ if } \theta_{1i} = \theta_{1,i+1}\}$. It is not hard to show that $\sqrt{n}(\hat{\theta}_1 - \theta_1)$ converges weakly to $\mathbf{U} = (U_1, \cdots, U_k)'$ where $U_j, j = 1, \cdots, k$ are independent normal variables whose mean and variance are $E[U_j] = 0$ and $Var[U_j] = \theta_{1j}(1 - \theta_{1j})/\theta_{2j}$, respectively. Since $\hat{\theta}_1$ converges almost surely to θ_1 as $n \to \infty$, we have $\hat{\theta}_{1i} > \hat{\theta}_{1,i+1}$ eventually if $\theta_{1i} > \theta_{1,i+1}$. Thus, when θ_2 is fixed, it follows that

$$P[M = k] = P[\hat{\theta}_{11} > \hat{\theta}_{12} > \cdots > \hat{\theta}_{1k}] \to P[\mathbf{U} \in \mathcal{I}_{\theta_1}] \text{ as } n \to \infty.$$

Since $\mathcal{I}_{\theta_1} \supset \mathcal{C}$, we have that

$$\inf_{\theta_{11} \geq \cdots \geq \theta_{1k}} \lim_{n \to \infty} P[M = k] = P[\mathbf{U} \in \mathcal{C}]. \qquad (3.4)$$

This infimum occurs when $\mathcal{I}_{\theta_1} = \mathcal{C}$ or equivalently $\theta_{11} = \cdots = \theta_{1k}$.

Now, we need to maximize the probability in (??) with respect to θ_2. Let $w_j = 1/Var[U_j], j = 1, \cdots, k$. Then, since $w_1 \geq w_2 \geq \cdots \geq w_k$ under H_1 with $\theta_{11} = \theta_{12} = \cdots = \theta_{1k}$, it is the immediate result of Theorem 1 that

$$\inf_{\theta \in H_1} \lim_{n \to \infty} P[M = k] = \inf_{\theta_{21} \geq \cdots \geq \theta_{2k}} P[\mathbf{U} \in \mathcal{C}] = \frac{1}{2^{k-1}(k-1)!}. \qquad (3.5)$$

Table 3.1. Incidence of Torus Mandibularis in Aleutians

Incidence	child	young adult	old adult	total
present	7	6	7	20
absent	16	15	4	35
total	23	21	11	55

From this result, the conditional chi-square test should be of size $\alpha^* = \alpha/[1 - 1/\{2^{k-1}(k-1)!\}]$ so as to meet the desired overall size α.

As an illustration of this conditional testing procedure, we analyze 'torus mandibularis' data from an Eskimo population which are tabulated in Table 3.1. These data are obtained by pooling Aleutian groups over sex in Table 9.7-1 of Bishop et al. (1975). Since α^* is close to α for large k, we consider only three groups: child$(1-10)$, younger adult $(21-30)$, and old adult(over 50). Suppose we test whether or not the incidence of torus mandibularis occurs more likely in older groups. As we did in the earlier part of the section, let θ_{1j} and θ_{2j} denote the rate of absence of the incidence in the jth group and the marginal probability of the jth age group, respectively. Since most of Eskimo populations are of pyramid structure, it is natural to assume that the whole parameter space is restricted by $\theta_{21} \geq \theta_{22} \geq \theta_{23}$. Thus, the problem becomes that of testing the null hypothesis $H_1 : \theta_{11} \geq \theta_{12} \geq \theta_{13}$ within the restricted parameter space. The unconstrained MLE's of θ_{ij}'s are $\hat{\theta}_{11} = .6957, \hat{\theta}_{12} = .7143, \hat{\theta}_{13} = .3636, \hat{\theta}_{21} = .4182, \hat{\theta}_{22} = .3818$, and $\hat{\theta}_{23} = .2000$. The H_1-constrained MLE's are computed as $\theta^*_{11} = \theta^*_{12} = .7045$, and $\theta^*_{13} = .3636$. For the test of overall size $\alpha = .05$, we must use $\alpha^* = .0375$ in the conditional test. Noting that θ^*_1 has two distinct levels $(M = 2)$, we do not reject H_1 because $T_{12} = .0183 < \chi^2_{3-2}(.0375) = 4.3276$.

REFERENCES

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

(2) Chase, G. R. (1974). On testing for ordered alternatives with increased sample size for a control. *Biometrika*, **61**, 569–578.

(3) Gupta, S. S. (1963). Bibliography on the multivariate normal integrals and related topics. *Annals of Mathematical Statistics*, **34**, 829–838.

(4) Lehmann, E. L. (1966). Some concepts of dependence. *Annals of Mathematical Statistics*, **26**, 1137–1153.

(5) Kepner, J. L., Harper, J. D., and Keith, S. Z. (1989). A note on evaluating a certain orthant probability. *The American Statistician*, **43**, 48–49.

(6) Nomakuchi, K. and Shi, N. Z. (1992). A note on the orthant probability of the elliptically countoured distribution. *Communications in Statistics — Theory and Methods*, **21**, 161–167.

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

(8) Sun, H. (1988). A general reduction method for n-variate normal orthant probability. *Communications in Statistics—Theory and Methods*, **17**, 3913–3921.