Vietnam Journal of Mathematics **36**:3(2008) 291–304

Vietnam Journal
of
MATHEMATICS

Nonparametric Statistical Inference for Umbrella Trend Alternative among Multinomial Populations *

Yanqin Feng

School of Mathematics and Statistics,
Wuhan University, Wuhan 430072, P. R. China

Received January 06, 2008

Abstract. Umbrella ordering is important in dose-response experiment. This paper treats the problem of comparing umbrella pattern treatment effects. Under an $m \times r$ table, for testing the equality against the umbrella ordering alternative, this article introduces a model-free test method by using likelihood ratio test statistic and gives the asymptotic distribution of this statistic, which is a chi-bar-squared distribution. A real example will be used to illustrate our theoretical result.

2000 Mathematics Subject Classification: 62J10.
Keywords: Stochastic ordering, chi-bar-squared distribution, likelihood ratio test, umbrella ordering, asymptotic distribution.

1. Introduction

A commonly occurring problem in statistics is that of testing for equality of two probability vectors corresponding to independent multinomial distributions against an alternative that they are not equal. Sometimes it is reasonable to assume that these vectors satisfy some type of a stochastic ordering and it may be of interest to test for equality against such an assumption. For example, Robertson and Wright [17] considered the problem of testing for equality of two multinomial distributions against the alternative that they are stochastically ordered and obtained the maximum likelihood estimates. Clearly, the problem

* The research is supported in part by NSFC #10771163.

of testing simple stochastic ordering among many treatment groups is very important, and this problem has been noticed in Chuang-Stein and Agresti [6], Agresti and Coull [1]. Wang [24] extended their work to more than two multinomial populations. In his paper, the limit distribution of likelihood ratio test statistic is characterized by minimization problems and has no closed form. Dardanoni and Forcina [8] considered the same hypothesis test problem and gave the asymptotic distribution of likelihood ratio test statistic. Colombi and Forcina [7] provided a general method for testing arbitrary sets of equality and inequality constraints on logits, log-odds ratios and higher order of interactions. Some results are cited in the book by Silvapulle and Sen [21].

For many objective causes, order-restricted problems for umbrella trend are commonly existed in practice. Umbrella ordering is an important concept in dose-response experiment, see, e.g., Simpson and Margolin [19]. In the cases where the mode of action of a drug is related to its toxic effects, for example, or in the cases of life saving digitalization therapy of heart failure, umbrella behavior is anticipated and careful dosage planning is required. For testing the null hypothesis against the umbrella ordering with at least one strict inequality, Bartholomew [2, 3] proposed the likelihood ratio test and Hayter and Liu [11] proposed another test for the ANOVA model. More detailed references are available in Marcus and Genizi [14] and Marcus and Talpaz [15]. Singh and Liu [20] proposed a test method for testing the equality of scale parameters of several distribution functions against an umbrella ordering with at least one strict inequality. All the results listed above are devoted to the scale parameters in distribution functions or population means. This paper presents an approach to testing the equality against an umbrella ordering placed on $m \geq 3$ probability vectors corresponding to m independent multinomial populations. Moreover our approach is model-free. The hypothesis test model studied here extends those in Wang [23]. While the method used in this paper is based on that in our earlier papers Feng and Wang [9, 10], although there are several other ways to derive the required distributions of the testing statistics (see e.g. Silvapulle and Sen [21]), and it is completely different from those used in the literatures listed above.

Let $\mathbf{p}_i = (p_{i1}, p_{i2}, \ldots, p_{ir})'$ denote the probability vectors corresponding to the i-th population for $1 \leq i \leq m$, and consider testing $\mathcal{H}_0$ against $\mathcal{H}_1 - \mathcal{H}_0$, where

$$\mathcal{H}_0 : \mathbf{p}_1 = \mathbf{p}_2 = \cdots = \mathbf{p}_m, \tag{1}$$

$$\mathcal{H}_1 : \mathbf{p}_1 \leq_{st} \cdots \leq_{st} \mathbf{p}_h \geq_{st} \cdots \geq_{st} \mathbf{p}_m. \tag{2}$$

Here 'st' means a stochastic ordering restriction. The theoretical results obtained in this paper are fit for well-defined stochastic ordering subject to equation (2), which can be expressed as

$$\mathcal{H}_1 : \theta \in S = \{\theta : (E \otimes C)\theta \geq \mathbf{0}\}, \tag{3}$$

where $\theta = (\theta_1', \theta_2', \ldots, \theta_m')'$ with $\theta_i = (p_{i1}, p_{i2}, \ldots, p_{i,r-1})'$ for $i = 1, 2, \ldots, m$, $E = (e_{ij})$ is an $(m-1) \times m$-order full rank matrix with $e_{ii} = 1$ if $i \leq h - 1$ and

-1 otherwise, $e_{i,i+1} = -1$ if $i \leq h-1$ and 1 otherwise, $e_{ij} = 0$ for $|i-j| > 1$, and C is any full rank matrix of order $r-1$.

The particular cases are that $h = 1$ and $h = m$, i.e., there exists monotone increasing or decreasing trend among the probability parameter vectors. This kind of order-restricted hypothesis test models has been studied by many statisticians. For example, Wang [24], Dardanoni and Forcina [8] and Feng and Wang [9] considered the testing $\mathcal{H}_0$ against $\mathcal{H}_1 - \mathcal{H}_0$ with $h = m$ in the means of simple stochastic ordering. However, umbrella order-restricted problem arises in many situations, in particular, it very commonly arises in medical research. This paper extends the hypothesis test model to the case that h can be any number among $1, 2, \ldots, m$ for any stochastic ordering with the alternative hypothesis that can be expressed as the form of equation (3), and gives the test statistic for testing $\mathcal{H}_0$ against $\mathcal{H}_1 - \mathcal{H}_0$ as well as obtains the limit distribution of the test statistic, which is shown to be of a chi-bar-squared type.

To illustrate our theoretical result, we consider the problem of the testing against simple stochastic ordering. This type of ordering of distribution is very important and most commonly considered ordering by statisticians. A random variable X with distribution function F is simply stochastically smaller than a random variable Y with distribution function G denoted by $F \leq_d G$ or $X \leq_d Y$ if

$$F(x) \geq G(x) \quad \text{for any real } x.$$

Among various formally defined notions of stochastic ordering used in the literature (see for example, Stoyan [22]), the least stringent one is the simple stochastic ordering. Other commonly used stochastic orderings, such as likelihood ratio ordering, hazard rate ordering and so on, are more restrictive and yield parameter spaces which are subsets of the simple stochastic ordering. Moreover, among various kinds of stochastic orderings, the simple stochastic ordering may be the best choice and most suitable for medical problems or other practical problems. As noted in Chuang-Stein and Agresti [6], it has the advantage of specifying the alternative in a broader and more realistic manner.

The rest of the paper is organized as follows. In Sec. 2, we transform the problem of testing simple stochastic ordering into a polyhedral cone constrained problem, form likelihood ratio test statistic, and derive the null asymptotic distribution of the test statistic. Section 3 includes a real example to illustrate the practical use of this test method.

2. Hypothesis Test

2.1. Equivalent Form of Test Model and Likelihood Ratio Test Statistic

This section gives an equivalent form of the hypothesis test model discussed in Sec. 1 and constructs the required likelihood ratio test statistic.

Suppose that we have interval censored data, which is often the case in medical research. For example, with checking times $-\infty = s_0 < s_1 < \cdots < s_r = \infty$, an experiment is checked at the times $s_1, s_2, \ldots, s_{r-1}$, and only the number of failures (e.g. deaths) between the checking time are observed. Another example is when a categorical response is determined by quantitative character, and corresponding to a different index and only the primary outcomes corresponding to certain events (e.g., tumor occurrence, death and so on) are observed. Consider independent multinomial random variables X_i for $i = 1, 2, \ldots, m$. Assume that the i-th variable summarizes the results of observing the outcome of n_i independent random experiments, each of which can result in anyone of r mutually exclusive outcomes, $1, 2, \ldots, r$, with observations $n_{i1}, n_{i2}, \ldots, n_{ir}$ and positive probabilities $p_{i1}, p_{i2}, \ldots, p_{ir}$, respectively. Let $n = \sum_{i=1}^{m} n_i$ denote the observations from all the m populations, and throughout the rest of this paper, we suppose that $\lim_{n\to\infty} n_i/n = \varphi_i > 0$ for $i = 1, 2, \ldots, m$.

By definition of simple stochastic ordering, under $\mathcal{H}_1$, we have

$$\sum_{t=1}^{j} p_{it} \geq \sum_{t=1}^{j} p_{i+1,t} \quad \text{for} \quad i = 1, 2, \ldots, h-1; j = 1, 2, \ldots, r-1,$$

$$\sum_{t=1}^{j} p_{it} \leq \sum_{t=1}^{j} p_{i+1,t} \quad \text{for} \quad i = h, h+1, \ldots, m-1; j = 1, 2, \ldots, r-1,$$

and

$$\sum_{t=1}^{r} p_{it} = 1, \quad \text{for} \quad i = 1, 2, \ldots, m. \tag{4}$$

Because (4) holds trivially for any multinomial distribution, the probability distribution of population X_i is completely determined by the first $(r-1)$-dimensional parameter vector $(p_{i1}, p_{i2}, \ldots, p_{i,r-1})'$ for $i = 1, 2, \ldots, m$, so, under $\mathcal{H}_1$, it is only required to consider the following inequality constraints

$$\sum_{t=1}^{j} p_{it} \geq \sum_{t=1}^{j} p_{i+1,t} \quad \text{for} \quad i = 1, 2, \ldots, h-1; j = 1, 2, \ldots, r-1,$$

and

$$\sum_{t=1}^{j} p_{it} \leq \sum_{t=1}^{j} p_{i+1,t} \quad \text{for} \quad i = h, h+1, \ldots, m-1; j = 1, 2, \ldots, r-1.$$

Let $\theta = (\theta_1', \theta_2', \cdots, \theta_m')'$ with $\theta_i = (p_{i1}, p_{i2}, \ldots, p_{i,r-1})'$, thus the hypothesis

$$\mathcal{H}_1 : \mathbf{p}_1 \leq_d \cdots \leq_d \mathbf{p}_h \geq_d \cdots \geq_d \mathbf{p}_m,$$

can be written equivalently as

$$\mathcal{H}_1 : \theta \in S = \{\theta : A\theta \geq \mathbf{0}\}, \tag{5}$$

and the null hypothesis test

$$\mathcal{H}_0 : \theta \in S_0 = \{\theta : A\theta = \mathbf{0}\},$$

where $(m-1)(r-1) \times m(r-1)$-order matrix $A = E \otimes D$, where $\otimes$ is the Kronecker product, $E = (e_{ij})$ is the same as that defined in (3), $D = (d_{ij})$ is a full rank matrix of order $r-1$, with $d_{ij} = 1$ if $i \geq j$ and 0 otherwise, and S is a polyhedral cone. Thus the hypotheses test problem for $\mathcal{H}_0$ against $\mathcal{H}_1 - \mathcal{H}_0$ can be written equivalently as

$$\mathcal{H}_0 : \theta \in S_0 \quad \text{versus} \quad \mathcal{H}_1 - \mathcal{H}_0 : \theta \in (S \setminus S_0), \tag{6}$$

where $S_0 \subset S \subseteq \Theta$, and

$$\Theta = \left\{ \theta : p_{ij} > 0, \sum_{i=1}^{r-1} p_{ij} < 1, i = 1, 2, \dots m \right\}.$$

Let $\mathbf{p} = (\mathbf{p}_1', \mathbf{p}_2', \dots, \mathbf{p}_m')'$ with $\mathbf{p}_i = (p_{i1}, p_{i2}, \dots, p_{ir})'$. The Log-likelihood function, omitting the constant term, for any outcome $(n_{i1}, n_{i2}, \dots, n_{ir})'$ for $i = 1, 2, \dots, m$ can be written as

$$\begin{aligned}\log L(p) &= \sum_{i=1}^{m} \sum_{j=1}^{r} n_{ij} \log p_{ij} \\ &= \sum_{i=1}^{m} \left\{ \sum_{j=1}^{r-1} n_{ij} \log p_{ij} + \left(n_i - \sum_{t=1}^{r-1} n_{it} \right) \log \left(1 - \sum_{t=1}^{r-1} p_{it} \right) \right\} \\ &\triangleq \log(\theta).\end{aligned}$$

Let $\mathbf{p}_0 = (p_{01}, p_{02}, \dots, p_{0r})'$ denote the common values of $\mathbf{p}_1 = \mathbf{p}_2 = \cdots = \mathbf{p}_m$ under $\mathcal{H}_0$, the unrestricted maximum likelihood estimate of p_{ij} is given by $\tilde{p}_{ij} = n_{ij}/n$, and $\bar{p}_{ij} = (n_1 \tilde{p}_{1j} + \cdots + n_m \tilde{p}_{mj})/n$ is the maximum likelihood estimate of p_{ij} under $\mathcal{H}_0$, thus $\bar{\theta} = (\bar{\theta}_1', \dots, \bar{\theta}_m')' = (\bar{p}_{11}, \dots, \bar{p}_{1,r-1}, \dots, \bar{p}_{m1}, \dots, \bar{p}_{m,r-1})'$ is the maximum likelihood estimate of θ under $\mathcal{H}_0$. In general, no closed presentation is available for the maximum likelihood estimate of p_{ij} under $\mathcal{H}_1$, and it is an optimal solution of the following mathematical programming problem

$$\begin{aligned} \min \quad & -\log L(\theta_1, \theta_2, \dots, \theta_m) \\ \text{s.t.} \quad & \theta \in S. \end{aligned} \tag{7}$$

Some usual algorithms in mathematical programming, for example, the feasible direction method or penalty function method, can be directly used to compute the maximum likelihood estimate of θ under the order-restricted condition (refer to Bazaraa and Shetty [4]). In practice, numerical solution of the maximum likelihood estimate of θ under $\mathcal{H}_1$ can be quickly computed by a series of Matlab functions.

Let Λ_{01} denote the likelihood ratio test statistic for testing $\mathcal{H}_0$ against $\mathcal{H}_1 - \mathcal{H}_0$. The likelihood ratio test will reject $\mathcal{H}_0$ in favor of $\mathcal{H}_1$ for large values of $T_{01} = -2\log\Lambda_{01}$.

2.2. Distribution Theory for the Likelihood Ratio Test Statistic

This section gives the desired null asymptotic distribution of the likelihood ratio statistic T_{01}, and which is a chi-bar-squared distribution. First of all, we summarize some key results about the chi-bar-squared variable as follows (refer to Shapiro [18] and Wang [23]).

Let $Y \sim N(0,\Sigma)$ be a k-dimensional normal distribution random vector, V be a closed convex cone, then

$$\bar{\chi}^2 = Y'\Sigma^{-1}Y - \min_{\beta\in V}(Y-\beta)'\Sigma^{-1}(Y-\beta) = \min_{\beta\in V^0}(Y-\beta)'\Sigma^{-1}(Y-\beta),$$

where V^0 is the polar (or dual) cone of V. The basic result about statistic $\bar{\chi}^2$ is distributed as a mixture of chi-squared distributions, that is,

$$P(\bar{\chi}^2 \geq c) = \sum_{i=0}^{k}\omega_i P(\chi_i^2 \geq c),$$

where χ_i^2 is a chi-squared random variable with degree i of freedom, $\chi_0^2 \equiv 0$ and ω_i's are nonnegative weights such that $\sum_{i=0}^{k}\omega_i = 1$. The weights $\omega_i = \omega_i(k,\Sigma,V)$ depend on Σ and V. Kudô [12] proposed a formula for the weights $\omega_i(k,\Sigma,R_+^k)$, in the case of $V = R_+^k = \{\mathbf{x} : \mathbf{x} > \mathbf{0}\}$ for $k \leq 4$ and the expression in a closed form of $\omega_i(k,\Sigma,R_+^k)$ is available. For $k > 4$, there are no closed expression form for the weights, but reasonably accurate estimates of the weights can be obtained easily by Monte Carlo simulations. Since the distribution of $\bar{\chi}^2$ is determined by covariance matrix Σ and convex cone V, we write $\bar{\chi}^2 \sim \bar{\chi}^2(\Sigma, V)$.

For the hypothesis test problem (6), the likelihood ratio test statistic is

$$\begin{aligned} T_{01} =& -2\min_{\theta\in S}\{\log L(\bar{\theta}) - \log L(\theta)\} = 2\{\log L(\hat{\theta}) - \log L(\bar{\theta})\} \\ =& \, 2\sum_{i=1}^{m}\sum_{j=1}^{r} n_{ij}(\log\hat{p}_{ij} - \log\bar{p}_{ij}), \end{aligned} \tag{8}$$

where $\hat{p}_{ij}$ and $\hat{\theta}$ denote the maximum likelihood estimates of p_{ij} and θ under $\mathcal{H}_1$, respectively.

The following theoretical results are helpful for us to derive the null asymptotic distribution of T_{01} (for the proofs, we refer to Feng and Wang [9]).

Lemma 1. *Suppose that $p_{ij} > 0$ for $i = 1,2,\ldots,m$ and $j = 1,2,\ldots,r$. Then, under $\mathcal{H}_0$, $\sqrt{n}(\hat{\theta} - \theta_0)$ is bounded in probability, that is,*

$$\lim_{c\to\infty}\lim_{n\to\infty} P\{\mid \hat{\theta} - \theta_0 \mid > cn^{-1/2}\} = 0,$$

where θ_0 is the unknown true value of θ under null hypothesis restriction.

Lemma 1 shows that, the maxima

$$\log L(\hat{\theta}) = \max\{\log L(\theta) : \theta \in S\},$$

is achieved in a $n^{-1/2}-$shrinking neighborhood of θ_0. So we consider $\beta = \sqrt{n}(\theta - \theta_0)$ as the optimization variables (refer to Prakasa Rao [16] and Wang [23]). Thus optimization parameter variables β could be any real numbers, and the constraint set Θ will be slackened in this case.

For the test problem (6), θ_0 is the true value of θ under $\mathcal{H}_0$, so it can be easily seen that $\theta \in S$ if and only if $\beta \in S$. Let

$$F_n(\theta) = 2\{\log L(\bar{\theta}) - \log L(\theta)\} \quad \text{and} \quad \mathcal{F}_n(\beta) = F_n(\beta n^{-\frac{1}{2}} + \theta_0),$$

thus (8) can be equivalently rewritten as

$$T_{01} = -\min_{\beta \in S} \mathcal{F}_n(\beta) = -\mathcal{F}_n(\hat{\beta}_n), \tag{9}$$

where $\hat{\beta}_n$ is the optimal solution.

For any fixed β, the following theorem (due to Feng and Wang [9]) gives the limit form of $\mathcal{F}_n(\beta)$, which plays an important role in obtaining the final limit distribution of the statistic T_{01}.

Theorem 2. *Suppose that $n_i \to \infty$, $p_{0j} > 0$ and $n_i/n \to \varphi_i > 0$ for $i = 1, 2, \dots, m; j = 1, 2, \dots, r$. Then under $\mathcal{H}_0$, $\mathcal{F}_n(\beta)$ converges in distribution to*

$$\mathcal{F}(\beta) = (\mathbf{Z} - \beta)' Q (\mathbf{Z} - \beta) - \frac{1}{2} \sum_{t,l=1}^{m} \varphi_t \varphi_l (\mathbf{Z}_t - \mathbf{Z}_l)' M (\mathbf{Z}_t - \mathbf{Z}_l), \tag{10}$$

where $\mathbf{Z} \sim N(\mathbf{0}, Q^{-1})$, $Q = \operatorname{diag}(Q_1, \dots, Q_m)$ is such that $Q_i = \varphi_i M$, $M = (m_{ij})_{(r-1)\times(r-1)}$ with $m_{ii} = 1/p_{0i} + 1/p_{0r}$ and $m_{ij}(i \neq j) = 1/p_{0r}$ for $i, j = 1, \dots, r-1$.

Theorem 2 has shown that $\mathcal{F}_n(\beta) \xrightarrow{L} \mathcal{F}(\beta)$ for any fixed $\beta \in S$. It is expected to show that $\mathcal{F}_n(\hat{\beta}_n) \to \mathcal{F}(\hat{\beta})$, where $\hat{\beta}$ is defined as

$$\hat{\beta} = \operatorname{argmin}_{\beta \in S} \mathcal{F}(\beta). \tag{11}$$

The following lemma (see Feng and Wang [9]) gives a necessary result for the validity of this result.

Lemma 3. *The stochastic processes $\{\mathcal{F}_n(\beta),\, \beta \in S \cap G\}$ converge in distribution to $\{\mathcal{F}(\beta),\, \beta \in S \cap G\}$, where $G = \{\beta : ||\beta|| \leq K,\, \beta \in \mathbb{R}^{m(r-1)}\}$.*

Now it is ready to derive the desired convergence result, which is a chi-bar-squared distribution, a mixture of chi-squared distributions.

Theorem 4. *Let Q and $\mathbf{Z}$ be as defined in Theorem 2, and let A be as in* (5). *Denote $\triangle = -Q^{-1}A'$ and $W = \triangle' Q \triangle$. Then under $\mathcal{H}_0$*

$$T_{01} \xrightarrow{L} T = \min_{\eta \in \mathbb{R}_+^{(m-1)(r-1)}} (\mathbf{X} - \eta)' W (\mathbf{X} - \eta) \sim \bar{\chi}^2 \left(W^{-1}, (\mathbb{R}_+^{(m-1)(r-1)})^0 \right),$$

where $\mathbf{X} = (\triangle' Q \triangle)^{-1} \triangle' Q \mathbf{Z}$, $\mathbf{X} \sim N(\mathbf{0}, W^{-1})$, and $(\mathbb{R}_+^{(m-1)(r-1)})^0$ is the polar cone of $\mathbb{R}_+^{(m-1)(r-1)}$.

Proof. The sample functions of the stochastic processes $\{\mathcal{F}_n(\beta), \beta \in S \cap G\}$ and $\{\mathcal{F}(\beta), \beta \in S \cap G\}$ are continuous functions over $S \cap G$. Let $\mathbb{C}(S \cap G)$ denote the space of all continuous functions over $S \cap G$ with a metric defined by

$$d(h_1, h_2) = \sup_{\beta \in S \cap G} \mid h_1(\beta) - h_2(\beta) \mid, \quad h_1, h_2 \in \mathbb{C}(S \cap G).$$

Let μ_n and μ denote the probability measures induced respectively by $\{\mathcal{F}_n(\beta), \beta \in S \cap G\}$ and $\{\mathcal{F}(\beta), \beta \in S \cap G\}$. By Theorem 2, we have $\mu_n \Rightarrow \mu$.

Define a mapping $H(\cdot)$ on $\mathbb{C}(S \cap G)$ as follows

$$H(\mathcal{F}_n) = \min_{\beta \in S \cap G} \mathcal{F}_n(\beta), \quad H(\mathcal{F}) = \min_{\beta \in S \cap G} \mathcal{F}(\beta).$$

By convexity of the function $\mathcal{F}(\beta)$ and of the set $S \cap F$, we obtain that $H(\mathcal{F})$ is unique. Hence the Continuous Mapping Theorem (Billingsley [5]) implies that $H(\mathcal{F}_n) \xrightarrow{L} H(\mathcal{F})$.

Since $\hat{\beta}$ is the optimal solution of (11) and $\beta = 0 \in S$, we obtain

$$\mathcal{F}(\hat{\beta}) - \mathcal{F}(\mathbf{0}) = (\mathbf{Z} - \hat{\beta})' Q (\mathbf{Z} - \beta) - \mathbf{Z}' Q \mathbf{Z} = \hat{\beta}' Q \hat{\beta} - 2\beta' Q \mathbf{Z} \leq 0.$$

Since Q is a positive definite matrix and $\mathbf{Z}$ is distributed as $N(\mathbf{0}, Q^{-1})$, for any $\varepsilon > 0$ there exists a constant C_ε such that $\| \hat{\beta} \| \leq C_\varepsilon$ with probability larger than $1 - \varepsilon$. Note that $\mathcal{F}_n(\hat{\beta}_n) = H(\mathcal{F}_n)$ and $\mathcal{F}(\hat{\beta}) = H(\mathcal{F})$, if $\| \hat{\beta}_n \| \leq C_\varepsilon$ and $\| \hat{\beta} \| \leq C_\varepsilon$. Therefore, we have

$$P\{\mathcal{F}_n(\hat{\beta}_n) \neq H(\mathcal{F}_n)\} < \varepsilon, \quad P\{\mathcal{F}(\hat{\beta}) \neq H(\mathcal{F})\} < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary and $H(\mathcal{F}_n) \xrightarrow{L} H(\mathcal{F})$, we obtain $\mathcal{F}_n(\hat{\beta}_n) \xrightarrow{L} \mathcal{F}(\hat{\beta})$, thus $T_{01} \xrightarrow{L} T$ holds.

We rewrite the limit expression (10) as

$$T = -\min_{\beta \in S} (\mathbf{Z} - \beta)' Q (\mathbf{Z} - \beta) + \frac{1}{2} \sum_{t,l=1}^{m} \varphi_t \varphi_l (\mathbf{Z}_t - \mathbf{Z}_l)' M (\mathbf{Z}_t - \mathbf{Z}_l) \triangleq T_1 - T_2,$$

where

$$T_1 = \mathbf{Z}'Q\mathbf{Z} - \min_{\beta\in S}(\mathbf{Z}-\beta)'Q(\mathbf{Z}-\beta),$$

and

$$T_2 = \mathbf{Z}'Q\mathbf{Z} - \frac{1}{2}\sum_{t,l=1}^{m}\varphi_t\varphi_l(\mathbf{Z}_t-\mathbf{Z}_l)'M(\mathbf{Z}_t-\mathbf{Z}_l).$$

It is obvious that $T_1 \sim \bar{\chi}^2(Q^{-1}, S)$. The polar cone of S is given by $S^0 = \{\beta : \beta = \triangle\eta, \eta \in \mathbb{R}_+^{(m-1)(r-1)}\}$, where $\triangle = -Q^{-1}A'$ is an $m(r-1)\times(m-1)(r-1)$ full rank matrix (we refer to Shapiro [18, p. 54]). It can be easily seen that

$$\begin{aligned} T_1 &= \min_{\beta\in S^0}(\mathbf{Z}-\beta)'Q(\mathbf{Z}-\beta) \\ &= \min_{\eta\in\mathbb{R}_+^{(m-1)(r-1)}}(\mathbf{Z}-\triangle\eta)'Q(\mathbf{Z}-\triangle\eta) \\ &= \min_{\eta\in\mathbb{R}_+^{(m-1)(r-1)}}(X-\eta)'W(X-\eta)+T_3, \end{aligned}$$

where

$$W = \triangle'Q\triangle, \quad X = (\triangle'Q\triangle)^{-1}\triangle'Q\mathbf{Z}, \quad X \sim N(\mathbf{0}, W^{-1}),$$

and

$$T_3 = \mathbf{Z}'\{I_{m(r-1)} - \triangle(\triangle'Q\triangle)^{-1}\triangle'Q\}'Q\{I_{m(r-1)} - \triangle(\triangle'Q\triangle)^{-1}\triangle'Q\}\mathbf{Z},$$

where $I_{m(r-1)}$ is a unit matrix of order $m(r-1)$ and

$$\triangle'Q\triangle = T\otimes(DM^{-1}D'), \quad (\triangle'Q\triangle)^{-1} = T^{-1}\otimes(DM^{-1}D')^{-1},$$

$T = (t_{ij})$ is a tridiagonal matrix of order $r-1$ with $t_{ii} = 1/\varphi_i + 1/\varphi_{i+1}$; $t_{i,i+1} = t_{i+1,i} = 1/\varphi_{i+1}$ if $i = h-1$ and $-1/\varphi_{i+1}$ otherwise, and $T^{-1} = (k_{ij})_{m\times m}$ is given by

$$k_{ij} = \begin{cases} (\varphi_1+\cdots+\varphi_j)(\varphi_{i+1}+\cdots+\varphi_m), \text{for } h-1\ge i\ge j \text{ or } i\ge j\ge h-1; \\ (\varphi_1+\cdots+\varphi_i)(\varphi_{j+1}+\cdots+\varphi_m), \text{for } h-1\ge j\ge i \text{ or } j\ge i\ge h-1; \\ -(\varphi_1+\cdots+\varphi_j)(\varphi_{i+1}+\cdots+\varphi_m), \text{for } i\ge h-1\ge j; \\ -(\varphi_1+\cdots+\varphi_i)(\varphi_{j+1}+\cdots+\varphi_m), \text{for } j\ge h-1\ge i. \end{cases}$$

Since $\varphi_1+\cdots+\varphi_m = 1$, a straightforward computation then yields

$$\triangle(\triangle'Q\triangle)^{-1}\triangle'Q = (I_m - B)\otimes I_{r-1},$$

where I_m and I_{r-1} are unit matrices of order m and $r-1$, respectively, and the matrix $B = (1,\ldots,1)'(\varphi_1,\ldots,\varphi_m)$. Thus we have

$$\begin{aligned} T_3 =& \mathbf{Z}'\{I_{m(r-1)} - \triangle(\triangle'Q\triangle)^{-1}\triangle'Q\}'Q\{I_{m(r-1)} - \triangle(\triangle'Q\triangle)^{-1}\triangle'Q\}\mathbf{Z} \\ =& \mathbf{Z}'(B\otimes I_{r-1})'Q(B\otimes I_{r-1})\mathbf{Z} \\ =& (\varphi_1\mathbf{Z}_1+\cdots+\varphi_m\mathbf{Z}_m)'M(\varphi_1\mathbf{Z}_1+\cdots+\varphi_m\mathbf{Z}_m). \end{aligned}$$

On the other hand,

$$
\begin{aligned}
T_2 &= \mathbf{Z}'Q\mathbf{Z} - \frac{1}{2}\sum_{t,l=1}^{m} \varphi_t\varphi_l(\mathbf{Z}_t - \mathbf{Z}_l)'M(\mathbf{Z}_t - Z_l) \\
&= \mathbf{Z}'QZ - \frac{1}{2}\sum_{t,l=1}^{m} \varphi_t\varphi_l(\mathbf{Z}_t'M\mathbf{Z}_t + \mathbf{Z}_lM\mathbf{Z}_l - 2\mathbf{Z}_t'M\mathbf{Z}_l) \\
&= \sum_{i=1}^{m}\left\{\varphi_i\mathbf{Z}_i'M\mathbf{Z}_i - \left(\sum_{l\neq i}\varphi_l\right)\varphi_i\mathbf{Z}_i'M\mathbf{Z}_i\right\} + \sum_{t\neq l}\varphi_t\varphi_l\mathbf{Z}_t'M\mathbf{Z}_l \\
&= \sum_{i=1}^{m}\varphi_i^2\mathbf{Z}_i'M\mathbf{Z}_i + \sum_{t\neq l}\varphi_t\varphi_l\mathbf{Z}_t'M\mathbf{Z}_l \\
&= (\varphi_1\mathbf{Z}_1 + \cdots + \varphi_m\mathbf{Z}_m)'M(\varphi_1\mathbf{Z}_1 + \cdots + \varphi_m Z_m) = T_3.
\end{aligned}
$$

Thus we obtain that

$$
T = \min_{\eta\in\mathbb{R}_+^{(m-1)(r-1)}} (X-\eta)'W(X-\eta),
$$

hence the conclusion holds. ■

Here, we have considered the tests for the case that the peak point (h) of the umbrella is known. For more common practical setting where the peak point of the umbrella is unknown, it is enough to apply our approach to the test models corresponding to all $h \in \{1, 2, \ldots, m\}$, that is, the idea of a direct plug-in method is used for $h = 1, 2, \ldots, m$, respectively. Order the p values p_i: $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(m)}$, let $\mathcal{H}_{(1)}, \mathcal{H}_{(2)}, \ldots, \mathcal{H}_{(m)}$ be the corresponding hypotheses, and the corresponding peaks are $h^1, h^2, \ldots, h^m$. For the significance level α, if $p_{(r)} \leq \alpha$, we suggest $h = \min\{h^1, h^2, \ldots, h^r\}$. Practically, it needs not consider all the probable values of h, but it needs only consider the values of h the researcher is interested in.

There is still a problem for applying the testing procedure, since the chi-bar-squared distribution depends on the unknown parameters θ_0, that how to handle the unknown parameter θ_0. For this problem one may compute the weights based on the estimate $\bar{\theta}$. Since $\bar{\theta}$ converges to θ_0 in probability, then $W(\bar{\theta})$ also converges to W in probability, therefore it is very reasonable to use $\bar{\theta}$ for the unknown θ_0.

3. Example

In this section, in order to illustrate the theoretical results given in Sec. 2, simulation studies based a real data set are carried out. All of the following results are obtained by a system of Matlab functions, and one may refer to Feng and Wang [9] for the computation of weights corresponding to the chi-bar-square distribution.

Table 1 includes five ordered categories describing the clinical trial outcomes of a sample of patients who experienced trauma due to subarachnoid hemorrhage. This table has four treatment groups, i.e., 'placebo' $\triangleq$ $Pop.1$, 'low dose' $\triangleq$ $Pop.2$, 'medium dose' $\triangleq$ $Pop.3$ and 'high dose' $\triangleq$ $Pop.4$, and five ordered categories ranging from 'death' to 'good recovery'(death$\triangleq$ 1, 'vegetative state' $\triangleq$ 2, 'major disability' $\triangleq$ 3, 'minor disability' $\triangleq$ 4 and 'good recovery' $\triangleq$ 5), which are often referred to as the Glasgow Outcome Scale. For these variables, the alternative is that the response improves in simple stochastic ordering sense as dose level increases. For $i = 1, 2, \ldots, 4$ and $j = 1, 2, \ldots, 5$, let

$$p_{ij} = P\{\text{Outcome} = j \mid \text{Treatment group} = i\} \quad \text{and} \quad \mathbf{p}_i = (p_{i1}, p_{i2}, \ldots, p_{i5})'.$$

Consider testing the null hypothesis of no treatment effect against the alternative of existing treatment effect. Specially, for $h = 2, 3$ and 4, we want to test $\mathcal{H}_0$ against $\mathcal{H}_1^h - \mathcal{H}_0$, where

$$\begin{aligned} &\mathcal{H}_0 : \mathbf{p}_1 = \mathbf{p}_2 = \mathbf{p}_3 = \mathbf{p}_4, \\ &\mathcal{H}_1^h : H_0 : \mathbf{p}_1 \leq_d \cdots \leq_d \mathbf{p}_h \geq_d \cdots \geq_d \mathbf{p}_4. \end{aligned} \tag{12}$$

In general, one may think that curative effect will be improved as the dose level increases. However, in fact, for the reason of the side-effect of drug, it is common for the response variable to increase with an increase of the treatment level up to a certain point then decrease with further increase in the treatment level. Therefore, we consider three kinds of alternative hypotheses listed in equation (12) for $h = 2, 3$ and $h = 4$.

Table 1. Results of a clinical trial comparing four treatments extent of trauma due to subarachnoid hemorrhage

Pop.	Death	Vegetative state	Major disability	Minor disability	Good	
1	59	25	46	48	32	210
2	48	21	44	47	30	190
2	44	14	54	64	31	207
4	43	4	49	58	41	195
	194	64	193	217	134	802

The fitted values under $\mathcal{H}_0$ and $\mathcal{H}_1^h$ for $h = 3, 4$ are given in Tables 2, 3 and 4, respectively. The p values of the test statistics for testing $\mathcal{H}_0$ against

$\mathcal{H}_1^h$ for $h = 2, 3$ and 4, are given in Table 5. We obtain that the peak is $h = 3$ or $h = 4$ for given significant level $\alpha \leq 0.01$, and we suggest $h = 3$.

Table 2. Fitted values of the clinical data under the hypothesis $\mathcal{H}_0$

Pop.	Death	Vegetative state	Major disability	Minor disability	Good
1	50.7980	16.7581	50.5362	56.8204	35.0873
2	45.9601	15.1621	45.7232	51.4090	31.7456
3	50.0723	16.5187	49.8142	56.0087	34.5860
4	47.1696	15.5611	46.9264	52.7618	32.5810

Table 3. Fitted values of the clinical data under the hypothesis $\mathcal{H}_1^3$

Pop.	Death	Vegetative state	Major disability	Minor disability	Good
1	59.0100	24.9900	45.9900	48.0060	32.0040
2	47.9940	20.9950	44.0040	47.0060	30.0010
3	41.0274	13.0617	53.0127	62.8452	37.0530
4	46.6050	4.3290	49.9785	59.1630	34.9245

Table 4. Fitted values of the clinical data under the hypothesis $\mathcal{H}_1^4$

Pop.	Death	Vegetative state	Major disability	Minor disability	Good
1	58.2120	24.9480	47.0820	48.1530	31.6050
2	47.6900	20.4440	45.4100	47.0440	29.4120
3	44.3601	14.1381	54.1305	62.3277	32.0436
4	41.7885	5.0700	49.0230	60.1185	39.0000

Table 5. Test statistics for testing $\mathcal{H}_0$ v.s. $\mathcal{H}_1^h$ for $h = 3, 4$ and the p values ($\times 10^{-4}$)

	$\mathcal{H}_0$ v.s. $\mathcal{H}_1^2$	$\mathcal{H}_0$ v.s. $\mathcal{H}_1^3$	$\mathcal{H}_0$ v.s. $\mathcal{H}_1^4$
Statistics	4.40	25.20	27.20
P-value	3130	0.9969	2.5519

The exact conditional test approach is valid irrespective of sample size and can be considered as a good competitor to asymptotic results. For the data listed in Table 1, Agresti and Coull also considered the hypothesis test problem $\mathcal{H}_0$ against $\mathcal{H}_1^4 - \mathcal{H}_0$, and they obtained a 95% confidence interval for the exact conditional p value is (0.0002, 0.0004), based on 100,000 randomly generated tables with margins equal to the observed values. For the same hypothesis test model, $\mathcal{H}_0$ against $\mathcal{H}_1^4 - \mathcal{H}_0$, the p value 2.5519×10^{-4} listed in Table 5 tends to conservative, all p values are small enough to provide a strong evidence of a positive association, that is, in simple stochastic ordering sense, a more favorable outcome tends to occur as the dose increases.

4. Conclusions

In this paper, we have shown how to test a stochastic ordering against umbrella trend upon the probability vectors of independent multinomials by using likelihood ratio test approach. Our result extends those of Wang [23], and Liu and Wang [13]. Neither the hypothesis models proposed in this paper, even for the particular case of $h = 1$ or $h = m$ in (2), nor the models studied by Dardanoni and Forcina [8] can uniformly include the other. The efficiency and the practicality of our method are demonstrated by a real example and the numerical studies offered desired p values.

References

1. A. Agresti and B. A. Coull, Order-restricted inference using odds ratios for monotone trend alternatives in contingency tables, *Comput. Stat. Data Anal.* **28** (1998), 139-155.
2. H. Bartholomew, A test of homogeneity for ordered alternative, *Biometrika* **46** (1959), 36-48.
3. D. J. Bartholomew, A test of homogeneity of means under restricted alternatives, *J. Royal Stat. Assoc, Ser B* **18** (1961), 239-281.
4. M. S. Bazaraa and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms, Wiley*, New York, 1979.
5. P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.
6. C. Chuang-Stein and A. Agresti, A review of tests for detecting a monotone dose-response relationship with ordinal response data, *Stat Med* **16** (1997), 2599-2618.
7. R. Colombi and A. Forcina, Marginal regression models for the analysis of positive association of ordinal response variables, *Biometrika* **88** (2001), 1007-1019.
8. V. Dardanoni and A. Forcina, A unified approach to likelihood inference on stochastic in a nonparametric context, *J. Amer. Stat. Assoc.* **93** (1998), 1112-1123.
9. Y. Feng and J. Wang, Likelihood ratio test against simple stochastic ordering among several multinomial populations, *J. Stat. Plann Inference* **137** (2007), 1362-1374.
10. Y. Feng and J. Wang, Test for increasing order in multivariate response, *Biometrical Journal* **49** (2007), 7-17.
11. A. J. Hayter and W. Liu, Exact calculations for the one-sided studentized range test for testing against a simple ordered alternative, *Comput. Stat. Data Anal.* **22** (1996), 17-25.
12. A. Kudô, A multivariate analogue of the one-sided test, *Biometrika* **50** (1963), 403-418.

13. X. Liu and J. Wang, Testing the equality of multinomial populations ordered by increasing convexity under the alternative, *Canad. J. Stat.* **32** (2004), 159-168.
14. R. Marcus and A. Genizi, Simultaneous confidence intervals for umbrella contrasts of normal means, *Comput. Stat. Data Anal.* **17** (1994), 393-407.
15. R. Marcus and H. Talpaz, Further results on testing homogeneity of normal means against simple tree alternatives, *Commun. Stat-Theory Meth.* **21** (1992), 2135-2149.
16. Prakasa Rao BLS, *Asymptotic Theory of Statistical Inference*, Wiley, New York, 1987.
17. T. Roberston and F. T. Wright, Likelihood ratio tests for and against a stochastic ordering between multinomial population, *Ann. Stat.* **9** (1981), 1248-1257.
18. A. Shapiro, Towards a unified theory of inequality constrained testing in multivariate analysis, *Int. Stat. Rev.* **56** (1988), 49-62.
19. D. G. Simpson and B. H. Margolin, Recursive nonparametric testing for dose-response relationships subject to down-turns at high dose, *Biometrika* **73** (1986), 589-596.
20. P. Singh and W. Liu, A test against an umbrella ordered alternative, *Comput. Stat. Data Anal.* **51** (2006), 1957-1964.
21. M. J. Silvapulle and P. K. Sen, *Constrained Statist Inference*, Wiley, New York, 2005.
22. D. Stoyan, *Comparsion Methods for Queues and Other Stochastic Models*, Wiley, New York, 1983.
23. J. Wang, Asymptotics of least-squares estimators for constrained nonlinear regression, *Ann. Statist.* **24** (1996), 1316-1326.
24. Y. Wang, A Likelihood ratio test against stochastic ordering in several populations, *J. Amer. Stat. Assoc.* **91** (1996), 1676-1683.