Ann. Inst. Statist. Math.
Vol. 49, No. 2, 237–254 (1997)

ON TESTS OF SYMMETRY AGAINST ONE-SIDED ALTERNATIVES

BHASKAR BHATTACHARYA

Department of Mathematics, Southern Illinois University, Carbondale, IL 62901-4408, U.S.A.

(Received May 31, 1995; revised September 13, 1996)

Abstract. To test that an experimental treatment is better than an existing one (or control), one can equivalently consider the difference in their response and test if the distribution of the difference is symmetric (about zero) versus it exhibits positive bias (skewness to the right). In this paper, we test the symmetry (about zero) of a discrete distribution against two particular classes of one sided alternatives. We obtain the maximum likelihood estimators under each alternative. The asymptotic null distributions of the likelihood ratio statistics are shown to have chi-bar square type distributions. A power study is performed to compare these one-sided alternatives with other one-sided tests. The theory developed is illustrated by an example.

Key words and phrases: Chi-bar square distribution, consistency, isotonic regression, likelihood ratio test, multinomial, positive bias, stochastic order relations.

1. Introduction

A random variable X has a symmetric distribution about θ if $X - \theta$ and $\theta - X$ have identical distribution. The concept of symmetry has played an important role in statistics, especially in nonparametric literature. In parametric inference, there are often situations when the assumption of normality as the underlying distribution can be replaced by the assumption of any symmetric distribution (e.g. Chaffin and Rhiel (1993)).

Tests of symmetry come up in several real life situations. The following paired-data example is given by Hettmansperger (1984). To determine the effect of environment on brain anatomy, an experiment is conducted as follows. Rats from several litters are randomly assigned to an enriched cage containing a variety of toys and an impoverished environment in which the rats lived in isolation. The variable of interest is the weight gain of the cortex over a specific period of time. The pairs are formed by litter mates with the same genetic makeup of the rats in an impoverished environment to the rats in an enriched environment. Let $X(Y)$ denote the impoverished (enriched) measurement, the random variable of interest is $D = Y - X$. If there is no difference in the effects of the two environments,

D has a distribution symmetric about 0. If we let θ denote the center of the distribution of D, then it is of interest to test $H_0 : \theta = 0$ versus $H_1 : \theta > 0$. The Wilcoxon signed rank test is used for this test of location of the center of the underlying symmetric distribution. Tests of symmetry (about 0) versus one-sided alternatives are also of interest in medical studies where patients are monitored before and after the administration of a drug.

One-sided hypotheses arise naturally in many situations. When testing against such hypotheses, it is desirable to take the available one-sided information into account, rather than simply applying a two sided test. This helps to increase the power of the test. When testing for symmetry of a distribution against both one-sided and two-sided alternatives, many nonparametric tests are available in the literature. However, most of the work on developing nonparametric tests against one-sided alternatives are directed to testing for possible changes in the location of the center of an underlying symmetric distribution (as in the above example). Many of the suggested tests for symmetry can be described as variations of either sign tests, Wilcoxon tests, Kolmogorov-Smirnov tests or Cramer-von Mises tests (Gastwirth (1971); Rothman and Woodroofe (1972); Hettmansperger (1984), for other related references). Since symmetry (about 0) is equivalent to $F(t-)+F(-t) = 1\ \forall t$, many tests are phrased in terms of the empirical cumulative distribution function (CDF) and take advantage of the rich literature on this topic (see Shorack and Wellner (1986), for details). Aki (1993) discussed nonparametric tests for symmetry in R^m and also provides many related references.

Of course, instead of testing for the location of the underlying distribution, it would be more useful to know the specific type of one-sided bias or skewness property of the underlying distribution that can lead to the rejection of the null hypothesis of symmetry for a given data set. In this paper we consider two one-sided alternatives to symmetry and consider the likelihood ratio tests against such alternatives.

We suppose the data are discrete, and assume without loss of generality that $\theta = 0$. We let X take on the $(2k+1)$ values $-k, -k+1, \ldots, -1, 0, 1, \ldots, k-1, k$ with $P(X = i) = p_i$, $-k \le i \le k$. We will denote the probability vector (PV) $(p_{-k}, p_{-k+1}, \ldots, p_{-1}, p_0, p_1, \ldots, p_{k-1}, p_k)$ by $\boldsymbol{p}$. We assume that $p_i > 0, \forall i$.

Suppose n_i, $-k \le i \le k$ are observed random variables from a multinomial distribution with $\sum_{i=-k}^{k} n_i = n$ and PV $\boldsymbol{p}$. Also let $\boldsymbol{n}$ denote the vector of observations.

The null hypothesis of symmetry about 0 can be expressed as

$$H_0 : p_i = p_{-i}, \qquad i = 1, \ldots, k.$$

Some possible alternatives to symmetry can be obtained by considering partial order relations between X and $-X$. For example, X is stochastically greater (in usual sense) than $-X$, or

$$\sum_{j=i}^{k} p_j \ge \sum_{j=i}^{k} p_{-j}, \qquad i = 1, \ldots, k, \tag{1.1}$$

with at least one strict inequality. This would imply that $E(g(X)) \geq E(g(-X))$ for all nondecreasing functions g. A stronger stochastic order relation is obtained by directly comparing the i-th cell with the $-i$-th cell, for all i, as follows

$$p_i \geq p_{-i}, \qquad i = 1, \ldots, k, \tag{1.2}$$

with at least one strict inequality. Both (1.1) and (1.2) are examples of positive bias (or, skewness to the right) in the sense that the probability of positive X is larger than the probability of negative X (referred to as the positive bias of first kind later in the paper). However, positive biasedness can also occur due to the factor that the random variable X under the condition $X > 0$ is stochastically larger than $-X$ under the condition $X < 0$ (Yanagimoto and Sibuya (1972)). An example of this latter (or, second) kind of positive bias is the following

$$H_1 : \frac{p_i}{p_{-i}} \text{ nondecreasing in } i, \qquad i = 1, \ldots, k. \tag{1.3}$$

An appropriate extension of (1.3) for a general univariate random variable, not necessarily discrete, is (with some algebra)

$$P(a_2 < X \leq a_3 \mid a_1 < X < a_3) > P(a_2 < -X < a_3 \mid a_1 < -X < a_3), \tag{1.4}$$

for all $0 < a_1 < a_2 < a_3$. The types of positive bias given by (1.1), (1.2), (1.3) and related properties have been discussed by Yanagimoto and Sibuya (1972).

Here we introduce a new type of positive bias (corresponding to the second kind) as follows

$$H_2 : \frac{\sum_{i=j}^{k} p_i}{\sum_{i=j}^{k} p_{-i}} \text{ nondecreasing in } j, \qquad j = 1, \ldots, k. \tag{1.5}$$

Also, for a general univariate random variable, not necessarily discrete, an appropriate extension of (1.5) is (with some algebra)

$$P(X > s + t \mid X > t) \geq P(-X > s + t \mid -X > t), \qquad \forall s \geq 0, \;\; t \geq 0. \tag{1.6}$$

It can be seen by some algebra that (1.3) is more restrictive than (1.5) (for $k > 2$). The likelihood ratio tests of symmetry versus the order relations given in (1.1) and (1.2) have been considered by Dykstra *et al.* (1995*a*). In this paper we are interested in the order relations given by (1.3) and (1.5) as alternatives to symmetry. The corresponding restrictions of negative bias can be handled by minor changes to the results of this paper.

In Section 2, we derive the maximum likelihood estimator (MLE) of $\boldsymbol{p}$ under the hypotheses H_0 and H_1, and use these estimates to obtain the likelihood ratio statistic for testing H_0 versus $H_1 - H_0$. The asymptotic distribution of the test statistic is shown to be of the chi-bar square type (a weighted combination of chi-square random variables mixed over their degrees of freedom). In Section 3, we consider the corresponding problem for the order restriction in H_2. Dykstra *et al.* (1995*a*) have shown that in the class of all univariate distributions, the generalized MLE (in the sense of Kiefer and Wolfowitz (1956)) under the restriction (1.1) is strongly consistent. However, the generalized MLE under the restriction (1.2) is not consistent when considered in the class of all univariate distributions. In Section 4, we show that the generalized MLE of F, the distribution of X under H_1, is strongly consistent when considered in the class of all univariate distributions. However, the generalized MLE of F, the distribution of X under H_2, is not consistent when considered in the class of all univariate distributions. In Section 5, we perform a Monte Carlo study to compare the power of different tests. As in Dykstra *et al.* (1995*a*), we consider the shifted binomial distribution for which the uniformly most powerful (UMP) test for testing symmetry against any of the one-sided alternatives is readily available. Using this UMP test as a benchmark, it is shown that the tests considered in this paper are quite powerful. In Section 6, we illustrate the procedures developed with an example. A discussion by comparing the results obtained in this paper with those in Dykstra *et al.* (1995*a*) is given in Section 7.

2. Testing H_0 versus $H_1 - H_0$

We first obtain the MLE of the vector $\boldsymbol{p}$ under hypotheses H_0 and H_1. The likelihood function is proportional to

$$L(\boldsymbol{p} \mid \boldsymbol{n}) = p_0^{n_0} \prod_{i=1}^{k} \{p_{-i}^{n_{-i}} p_i^{n_i}\}.$$

The unrestricted MLE of p_i is $\hat{p}_i = n_i/n$, $i = 0, \pm 1, \ldots, \pm k$. Let $\hat{\boldsymbol{p}} = (\hat{p}_{-k}, \ldots, \hat{p}_k)$. We consider a reparametrization as follows. Let

$$\theta_i = \frac{p_i}{p_i + p_{-i}}, \quad \phi_i = p_i + p_{-i}, \qquad 1 \le i \le k.$$

Then

$$p_i = \theta_i \phi_i, \quad p_{-i} = (1-\theta_i)\phi_i, \qquad 1 \le i \le k. \tag{2.1}$$

Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$ and $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_k)$. The likelihood function, in terms of the parameters θ_i's and ϕ_i's, is proportional to

$$L(\boldsymbol{\theta}, \boldsymbol{\phi} \mid \boldsymbol{n}) = \left[\prod_{i=1}^{k} \theta_i^{n_i}(1-\theta_i)^{n_{-i}}\right] \left[\prod_{i=1}^{k} \phi_i^{n_i + n_{-i}} \left(1 - \sum_{i=1}^{k} \phi_i\right)^{n_0}\right]. \tag{2.2}$$

It is easily seen that the MLE's of θ_i's and ϕ_i's under H_0 are

$$\hat{\theta}_i^0 = \frac{1}{2}, \quad \hat{\phi}_i^0 = \hat{p}_{-i} + \hat{p}_i, \qquad 1 \le i \le k.$$

Under H_1, the constraints on p_i can equivalently be expressed in terms of the θ_i's and ϕ_i's as $0 \le \theta_i, \phi_i \le 1, \forall i$, and

$$\theta_1 \le \theta_2 \le \cdots \le \theta_k. \tag{2.3}$$

To find the MLE's of θ_i's and ϕ_i's under (2.3), we can maximize the terms in two brackets in (2.2) individually, since there are no constraints connecting the θ_i's and ϕ_i's. Maximizing the quantity in first bracket subject to the restrictions in (2.3) is a bioassay problem (as described in Robertson *et al.* (1988)), and hence the MLE of $\boldsymbol{\theta}$ under (2.3) is $\boldsymbol{\theta}^* = E_{\hat{\boldsymbol{v}}}(\hat{\boldsymbol{\theta}} \mid \mathcal{I}_k)$, where $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k)$, $\hat{\boldsymbol{v}} = (\hat{v}_1, \ldots, \hat{v}_k)$,

$$\hat{\theta}_i = \frac{\hat{p}_i}{\hat{p}_{-i} + \hat{p}_i}, \quad \hat{v}_i = \hat{p}_{-i} + \hat{p}_i, \quad 1 \le i \le k, \tag{2.4}$$

and $E_{\boldsymbol{w}}(\boldsymbol{x} \mid \mathcal{I}_k)$ is the least square projection with weights $\boldsymbol{w}$ of the vector $\boldsymbol{x}$ onto $\mathcal{I}_k$, the cone of nondecreasing vectors of length k. Since there is no restriction on the ϕ_i's, the MLE of $\boldsymbol{\phi}$ under (2.3) is the same as the corresponding unrestricted MLE's. Using (2.1), we obtain the MLE of $\boldsymbol{p}$ under H_1, as given in the following theorem.

THEOREM 2.1. *The MLE of* $\boldsymbol{p}$ *under* H_1 *is given by*

$$\begin{aligned} \hat{p}_0^* &= \frac{n_0}{n}, \\ \hat{p}_i^* &= \frac{n_{-i} + n_i}{n} E_{\hat{\boldsymbol{v}}}(\hat{\boldsymbol{\theta}} \mid \mathcal{I}_k)_i, \quad 1 \le i \le k, \\ \hat{p}_{-i}^* &= \frac{n_{-i} + n_i}{n} - \frac{n_{-i} + n_i}{n} E_{\hat{\boldsymbol{v}}}(\hat{\boldsymbol{\theta}} \mid \mathcal{I}_k)_i, \quad 1 \le i \le k, \end{aligned} \tag{2.5}$$

where $\hat{v}_i$ *and* $\hat{\theta}_i$ *are defined in* (2.4).

Robertson *et al.* (1988) described several algorithms for computing the least square projection $E_{\boldsymbol{w}}(\boldsymbol{x} \mid \mathcal{I}_k)$. The pool adjacent violators algorithm (PAVA) is one of the simplest and can be used for our purpose.

2.1 *The asymptotic null distribution of the test statistic*

Let Λ_1 be the likelihood ratio when testing H_0 versus $H_1 - H_0$ and we reject H_0 for large values of $T_1 = -2 \ln \Lambda_1$. Using the MLE as constructed earlier, it is straightforward to show that

$$T_1 = 2n \sum_{i=1}^{k} \left[\hat{p}_i \ln\left(\frac{\theta_i^*}{\theta_i^0}\right) + \hat{p}_{-i} \ln\left(\frac{1-\theta_i^*}{1-\theta_i^0}\right) \right].$$

Expanding $\ln \theta_i^*$, $\ln \theta_i^0$ about $\hat{\theta}_i$, and $\ln(1-\theta_i^*)$, $\ln(1-\theta_i^0)$ about $(1-\hat{\theta}_i)$ via Taylor's theorem with a second degree remainder term, we see that the linear terms drop out, and combining appropriate terms, we obtain

$$T_1 = n \sum_{i=1}^{k} \left\{ (\theta_i^0 - \hat{\theta}_i)^2 \left[\frac{\hat{p}_i}{(1-\alpha_i)^2} + \frac{\hat{p}_{-i}}{\beta_i^2} \right] - (\theta_i^* - \hat{\theta}_i)^2 \left[\frac{\hat{p}_i}{\gamma_i^2} + \frac{\hat{p}_{-i}}{(1-\delta_i)^2} \right] \right\}$$

where α_i, β_i are between θ_i^0 and $\hat{\theta}_i$, and γ_i, δ_i are between θ_i^* and $\hat{\theta}_i$. Under H_0, each of α_i, β_i, γ_i, δ_i converges to 1/2, and hence it follows that, the asymptotic distribution of T_1 is the same as the asymptotic distribution of

$$\sum_{i=1}^{k} 4(\hat{p}_i + \hat{p}_{-i})\{[\sqrt{n}(\theta_i^0 - \hat{\theta}_i)]^2 - [\sqrt{n}(\theta_i^* - \hat{\theta}_i)]^2\}.$$

Using the multinomial central limit theorem to $\sqrt{n}(\hat{\boldsymbol{p}} - \boldsymbol{p})$, and using the delta method, it follows that

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{\mathcal{L}} MVN(\mathbf{0}, \mathbf{W}),$$

where $\mathbf{W}$ is a diagonal matrix with i-th entry equal to $p_i p_{-i}(p_i + p_{-i})^{-3}$. Let $\mathbf{U} = (U_1, \ldots, U_k) \sim MVN(\mathbf{0}, \mathbf{W})$. Then, under H_0, using continuity of the projection operator, it follows that $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*) \xrightarrow{\mathcal{L}} \mathbf{U} - E_{\boldsymbol{v}}(\mathbf{U} \mid \mathcal{I}_k)$, where $\boldsymbol{v} = (v_1, \ldots, v_k)$ and $v_i = 2p_i, \forall i$. Then, under H_0, the asymptotic distribution of T_1 is the same as the distribution of

$$\begin{aligned}
T_1' &= \sum_{i=1}^{k} 8p_i[U_i^2 - (U_i - E_{\boldsymbol{v}}(\mathbf{U} \mid \mathcal{I}_k)_i)^2] \\
&= \sum_{i=1}^{k} 8p_i U_i E_{\boldsymbol{v}}(\mathbf{U} \mid \mathcal{I}_k)_i^2 + \sum_{i=1}^{k} 8p_i[U_i E_{\boldsymbol{v}}(\mathbf{U} \mid \mathcal{I}_k)_i - E_{\boldsymbol{v}}(\mathbf{U} \mid \mathcal{I}_k)_i^2] \\
&= \sum_{i=1}^{k} 8p_i E_{\boldsymbol{v}}(\mathbf{U} \mid \mathcal{I}_k)_i^2
\end{aligned}$$

since $\sum_{i=1}^{k} 8p_i U_i E_{\boldsymbol{v}}(\mathbf{U} \mid \mathcal{I}_k)_i = \sum_{i=1}^{k} 8p_i E_{\boldsymbol{v}}(\mathbf{U} \mid \mathcal{I}_k)_i^2$.

Let $\boldsymbol{p} \in H_0$, and let c be a real number. Also, let $G(B_1, \ldots, B_\ell)$ be the event on which $B_1, \ldots, B_\ell$ are the ordered level sets of $E_{\boldsymbol{v}}(\mathbf{U} \mid \mathcal{I}_k)$ (level sets are subsets of $\{1, \ldots, k\}$ where $E_{\boldsymbol{v}}(\mathbf{U} \mid \mathcal{I}_k)$ has constant value). Then

$$\begin{aligned}
P(T_1' \geq c) &= \sum_{\ell=1}^{k} \sum_{B_1,\ldots,B_\ell} P[T_1' \geq c, G(B_1, \ldots, B_\ell)] \\
&= \sum_{\ell=1}^{k} \sum_{B_1,\ldots,B_\ell} P\left[\sum_{i=1}^{\ell} \left(\sum_{j \in B_i} 8p_j\right) (\mathrm{Av}(B_i))^2 \geq c, G(B_1, \ldots, B_\ell)\right] \\
&= \sum_{\ell=1}^{k} \sum_{B_1,\ldots,B_\ell} P\left[\sum_{i=1}^{\ell} \left(\sum_{j \in B_i} 8p_j\right) (\mathrm{Av}(B_i))^2 \geq c \,\middle|\, G(B_1, \ldots, B_\ell)\right] \\
&\quad \times P(G(B_1, \ldots, B_\ell)) \\
&= \sum_{\ell=1}^{k} \left(\sum_{B_1,\ldots,B_\ell} P(G(B_1, \ldots, B_\ell))\right) P(\chi_\ell^2 \geq c) \\
&= \sum_{\ell=1}^{k} P(\ell, k, \boldsymbol{v}) P(\chi_\ell^2 \geq c)
\end{aligned}$$

(the equality before the last follows by using Lemma B on p. 71 of Robertson *et al.* (1988)) where Av is the (weighted) average function and $P(\ell, k, \boldsymbol{v}) = \sum_{B_1,\ldots,B_\ell} P(G(B_1,\ldots,B_\ell))$ is the probability that $E_{\boldsymbol{v}}(\mathbf{U} \mid \mathcal{I}_k)$ has exactly ℓ level sets.

The above developments are summarized in the following theorem. The asymptotically least favorable distribution of T_1 (corresponds to the case when the probability of rejecting H_0 is maximum) can be obtained from Theorem 3.6.1 of Robertson *et al.* (1988).

THEOREM 2.2. *When H_0 is true and $p_i > 0, \forall i$, then for any real number c*

$$\lim_{n\to\infty} P(T_1 \geq c) = \sum_{\ell=1}^{k} P(\ell, k, \boldsymbol{v}) P(\chi_\ell^2 \geq c),$$

where χ_ν^2 is a chi-square random variable with ν degrees of freedom. The asymptotically least favorable distribution is given by

$$\sup_{p\in H_0} \lim_{n\to\infty} P(T_1 > c) = \sum_{\ell=1}^{k} \binom{k-1}{\ell-1} 2^{-k+1} P(\chi_\ell^2 > c). \tag{2.6}$$

The asymptotic distribution of T_1 depends on $\boldsymbol{p}$ through the level probabilities. When v_i's are equal (which corresponds to an essentially uniform distribution under H_0), then the level probabilities can be calculated recursively (Corollary A on p. 81 of Robertson *et al.* (1988)). For arbitrary weights and $k \leq 4$, the calculations for level probabilities are given in Robertson *et al.* (1988). For arbitrary weights with $k \geq 5$, no closed form expressions for the level probabilities exist. Monte-Carlo techniques often provide a good alternative due to the difficulty of calculations. Another possibility is to use pattern approximations (see Robertson *et al.* (1988), for details).

Note that according to Lee *et al.* (1991), bounds identical to the right side of (2.6) are obtained for the case of nonincreasing weights which corresponds to a symmetric, unimodal distribution under H_0.

The asymptotic critical values of T_1 based on the least favorable distribution of (2.6) are given in Table 1 for $k-1 = 2, \ldots, 14$ and $\alpha = 0.10, 0.05, 0.01$. However, a test based on the least favorable distribution is likely to be conservative. An alternative is to approximate the level probabilities by $P(\ell, k, \hat{\boldsymbol{p}}^0)$, where $\hat{\boldsymbol{p}}^0$ is the estimate of $\boldsymbol{p}$ under H_0 (see, e.g., Bohrer and Chow (1978)). The resulting expression has the same asymptotic distribution as T_1 under H_0 and provides a good approximation (Oh (1994)).

3. Testing H_0 versus $H_2 - H_0$

To find the MLE of $\boldsymbol{p}$ under H_2, we reparametrize as follows. Let

$$\theta_{-1} = \sum_{i=1}^{k} p_{-i}, \quad \theta_1 = \sum_{i=1}^{k} p_i,$$
$$\theta_{-j} = \frac{\sum_{i=j}^{k} p_{-i}}{\sum_{i=j-1}^{k} p_{-i}}, \quad \theta_j = \frac{\sum_{i=j}^{k} p_i}{\sum_{i=j-1}^{k} p_i}, \qquad 2 \le j \le k.$$

Then it can be seen that the p_j's can be expressed in terms of the new parameters as follows

$$\begin{aligned} p_{-j} &= (1-\theta_{-j-1}))\prod_{i=1}^{j}\theta_{-i}, \quad p_j = (1-\theta_{j+1})\prod_{i=1}^{j}\theta_i, \qquad 1 \le j \le k-1, \\ p_{-k} &= \prod_{i=1}^{k}\theta_{-i}, \quad p_k = \prod_{i=1}^{k}\theta_i. \end{aligned} \tag{3.1}$$

Let $m_j = \sum_{i=j}^{k} n_i$, $m_{-j} = \sum_{i=j}^{k} n_{-i}$, $1 \le j \le k$. The likelihood function in terms of the new parameters is proportional to

$$\begin{aligned} L(\boldsymbol{\theta} \mid \boldsymbol{n}) &= (1-\theta_{-1}-\theta_1)^{n-m_{-1}-m_1}\theta_{-1}^{m_{-1}}\theta_1^{m_1} \\ &\quad \times \prod_{i=2}^{k}[\{\theta_{-i}^{m_{-i}}(1-\theta_{-i})^{m_{-i+1}-m_{-i}}\}\{\theta_i^{m_i}(1-\theta_i)^{m_{i-1}-m_i}\}]. \end{aligned} \tag{3.2}$$

In terms of the parameters θ_j's, the hypotheses H_0 and H_2 can be rewritten as

$$\begin{aligned} &H_0 : \theta_{-j} = \theta_j, \qquad 1 \le j \le k, \\ &H_2 : \theta_{-j} \le \theta_j, \qquad 2 \le j \le k. \end{aligned} \tag{3.3}$$

Note that, for both H_0 and H_2, there is no constraint over θ_j for different values of j, only θ_{-j} and θ_j are related, $1 \le j \le k$. Then it follows that, under H_0, the MLE of θ_j is given by

$$\begin{aligned} \theta_{-1}^0 &= \theta_1^0 = \frac{m_{-1}+m_1}{2n}, \\ \theta_{-j}^0 &= \theta_j^0 = \frac{m_{-j}+m_j}{m_{-j+1}+m_{j-1}}, \qquad 2 \le j \le k. \end{aligned} \tag{3.4}$$

Under H_2, there is no constraint on θ_1 and θ_{-1}, so the estimates of these parameters under H_2 are the same as the corresponding unrestricted MLE's. Maximizing the second line of (3.2) subject to the constraints H_2 in (3.3) is equivalent to $(k-1)$ bioassay problems, which can be solved independently. Thus the MLE of θ_j, under H_2, is given by

$$\begin{aligned} &\theta_{-1}^* = \frac{m_{-1}}{n}, \quad \theta_1^* = \frac{m_1}{n}, \\ &\theta_{-j}^* = E_{\hat{\vartheta}_j}((\hat{\theta}_{-j}, \hat{\theta}_j) \mid \mathcal{I}_2)_1, \quad \theta_j^* = E_{\hat{\vartheta}_j}((\hat{\theta}_{-j}, \hat{\theta}_j) \mid \mathcal{I}_2)_2, \qquad 2 \le j \le k, \end{aligned} \tag{3.5}$$

where

$$\hat{\theta}_{-j} = \frac{m_{-j}}{m_{-j+1}}, \quad \hat{\theta}_j = \frac{m_j}{m_{j-1}}, \quad \hat{v}_j = (m_{-j+1}, m_{j-1}), \qquad 2 \le j \le k. \tag{3.6}$$

Using equations (3.1), (3.5) and (3.6), we obtain the MLE of $\boldsymbol{p}$ under H_2 as given by the following theorem.

THEOREM 3.1. *The MLE of* $\boldsymbol{p}$ *under* H_2 *is given by*:

$$\hat{p}^*_{-j} = (1 - \theta^*_{-j-1}) \prod_{i=1}^{j} \theta^*_{-i}, \quad \hat{p}^*_j = (1 - \theta^*_{j+1}) \prod_{i=1}^{j} \theta^*_i, \qquad 1 \le j \le k-1,$$

$$\hat{p}^*_{-k} = \prod_{i=1}^{k} \theta^*_{-i}, \quad \hat{p}^*_k = \prod_{i=1}^{k} \theta^*_i,$$

where θ^*_{-i} *and* θ^*_i *are as defined in* (3.5).

3.1 *The asymptotic null distribution of the test statistic*

Let Λ_2 be the likelihood ratio when testing H_0 versus $H_2 - H_0$ and we reject H_0 for large values of $T_2 = -2 \ln \Lambda_2$. Using (3.2) and the MLE's as constructed in (3.4) and (3.5), it is straightforward to show that

$$T_2 = A + B$$

where

$$A = 2\left[m_{-1} \ln\left(\frac{\theta^*_{-1}}{\theta^0_{-1}}\right) + m_1 \ln\left(\frac{\theta^*_1}{\theta^0_1}\right)\right]$$

and

$$B = 2\left[\sum_{j=2}^{k} m_{-j} \ln\left(\frac{\theta^*_{-j}}{\theta^0_{-j}}\right) + \sum_{j=2}^{k} m_j \ln\left(\frac{\theta^*_j}{\theta^0_j}\right) + \sum_{j=2}^{k} (m_{-j+1} - m_{-j}) \ln\left(\frac{1 - \theta^*_{-j}}{1 - \theta^0_{-j}}\right) + \sum_{j=2}^{k} (m_{j-1} - m_j) \ln\left(\frac{1 - \theta^*_j}{1 - \theta^0_j}\right)\right].$$

We will find the asymptotic distributions of A and B separately. First consider the asymptotic distribution of A. Expanding $\ln \theta^0_i$ about θ^*_i, for $i = -1, 1$, we find the linear terms drop out, and the asymptotic distribution of A is the same as the distribution of

$$\theta^*_{-1} \frac{(\theta^0_{-1} - \theta^*_{-1})^2}{\psi_1^2} + \theta^*_1 \frac{(\theta^0_1 - \theta^*_1)^2}{\psi_2^2}, \tag{3.7}$$

where ψ_1 and ψ_2 are obtained from the Taylor expansion. Under H_0, the asymptotic distribution of the quantity in (3.7) is chi-square with 1 degree of freedom (also equivalent to the chi-square goodness-of-fit test of symmetry on a 3-cell multinomial).

Now consider the asymptotic distribution of B. For negative and positive indices j, expanding $\ln \theta_j^*$, $\ln \theta_j^0$ about $\hat{\theta}_j$ and $\ln(1-\theta_j^*)$, $\ln(1-\theta_j^0)$ about $(1-\hat{\theta}_j)$ via Taylor's theorem with a second degree remainder term, we note that all the linear terms cancel out and we obtain

$$
\begin{aligned}
B = & \sum_{j=2}^{k} m_{-j} \left\{ \frac{(\theta_{-j}^0 - \hat{\theta}_{-j})^2}{\psi_{1j}^2} - \frac{(\theta_{-j}^* - \hat{\theta}_{-j})^2}{\psi_{2j}^2} \right\} \\
& + \sum_{j=2}^{k} (m_{-j+1} - m_{-j}) \left\{ \frac{(\theta_{-j}^0 - \hat{\theta}_{-j})^2}{(1-\psi_{3j})^2} - \frac{(\theta_{-j}^* - \hat{\theta}_{-j})^2}{(1-\psi_{4j})^2} \right\} \\
& + \sum_{j=2}^{k} m_j \left\{ \frac{(\theta_j^0 - \hat{\theta}_j)^2}{\psi_{5j}^2} - \frac{(\theta_j^* - \hat{\theta}_j)^2}{\psi_{6j}^2} \right\} \\
& + \sum_{j=2}^{k} (m_{j-1} - m_j) \left\{ \frac{(\theta_j^0 - \hat{\theta}_j)^2}{(1-\psi_{7j})^2} - \frac{(\theta_j^* - \hat{\theta}_j)^2}{(1-\psi_{8j})^2} \right\}
\end{aligned}
$$

where ψ_{1j} to ψ_{8j} are obtained from the Taylor expansion.

After combining similar terms, we obtain

$$
\begin{aligned}
B = & \sum_{j=2}^{k} \left\{ \frac{m_{-j}}{\psi_{1j}^2} + \frac{m_{-j+1} - m_{-j}}{(1-\psi_{3j})^2} \right\} (\theta_{-j}^0 - \hat{\theta}_{-j})^2 \\
& - \sum_{j=2}^{k} \left\{ \frac{m_{-j}}{\psi_{2j}^2} + \frac{m_{-j+1} - m_{-j}}{(1-\psi_{4j})^2} \right\} (\theta_{-j}^* - \hat{\theta}_{-j})^2 \\
& + \sum_{j=2}^{k} \left\{ \frac{m_j}{\psi_{5j}^2} + \frac{m_{j-1} - m_j}{(1-\psi_{7j})^2} \right\} (\theta_j^0 - \hat{\theta}_j)^2 \\
& - \sum_{j=2}^{k} \left\{ \frac{m_j}{\psi_{6j}^2} + \frac{m_{j-1} - m_j}{(1-\psi_{8j})^2} \right\} (\theta_j^* - \hat{\theta}_j)^2.
\end{aligned}
$$

Then it also follows that, under H_0, the asymptotic distribution of B is the same as that of

$$
(3.8) \quad \sum_{j=2}^{k} \frac{\sum_{i=j-1}^{k} p_i}{\theta_j(1-\theta_j)} \{[(\theta_{-j}^0 - \hat{\theta}_{-j})^2 - (\theta_{-j}^* - \hat{\theta}_{-j})^2] + [(\theta_j^0 - \hat{\theta}_j)^2 - (\theta_j^* - \hat{\theta}_j)^2]\}.
$$

Using the multinomial central limit theorem to $\sqrt{n}(\hat{\boldsymbol{p}} - \boldsymbol{p})$, and the delta method, it follows that, for $2 < j \leq k$,

$$
(3.9) \quad \sqrt{n}((\hat{\theta}_{-j}, \hat{\theta}_j) - (\theta_{-j}, \theta_j)) \xrightarrow{\mathcal{L}} MVN\left((0,0), \begin{bmatrix} \frac{\theta_{-j}(1-\theta_{-j})}{\sum_{i=j-1}^{k} p_{-i}} & 0 \\ 0 & \frac{\theta_j(1-\theta_j)}{\sum_{i=j-1}^{k} p_i} \end{bmatrix} \right).
$$

Table 1. The $100(1-\alpha)$ percentiles for the chi-bar square distribution given in (2.6) and (3.11).

	α		
$k-1$	0.10	0.05	0.01
2	3.81	5.14	8.27
3	4.78	6.25	9.64
4	5.68	7.27	10.86
5	6.54	8.22	11.99
6	7.35	9.12	13.05
7	8.14	10.00	14.07
8	8.91	10.84	15.05
9	9.66	11.66	16.00
10	10.40	12.47	16.93
11	11.13	13.25	17.83
12	11.84	14.03	18.72
13	12.55	14.80	19.59
14	13.25	15.55	20.45

Let $\mathbf{Y}_j = (Y_{j1}, Y_{j2})$ has the distribution given by the right side of (3.9). Then, using continuity of the projection operator, under H_0, and some algebra, it follows from (3.8) that, the asymptotic distribution of B is the same as the distribution of

$$\sum_{j=2}^{k} \frac{\sum_{i=j-1}^{k} p_i}{\theta_j(1-\theta_j)} \{[E_{v_j}(\mathbf{Y}_j \mid \mathcal{I}_2)_1 - \bar{Y}_j]^2 + [E_{v_j}(\mathbf{Y}_j \mid \mathcal{I}_2)_2 - \bar{Y}_j]^2\}, \tag{3.10}$$

where $\boldsymbol{v}_j = (v_{j1}, v_{j2}) = (\sum_{i=j-1}^{k} p_{-i}, \sum_{i=j-1}^{k} p_i)$ and $\bar{Y}_j = (v_{j1}Y_{j1} + v_{j2}Y_{j2})/(v_{j1} + v_{j2})$, $2 \le j \le k$. If we denote each quantity in (3.10) under the summation by B_j for $2 \le j \le k$, then the distribution of B_j is given by

$$P(B_j \ge c) = \frac{1}{2}P(\chi_0^2 \ge c) + \frac{1}{2}P(\chi_1^2 \ge c)$$

where $\chi_0^2 \equiv 0$. It is clear from (3.9) that $\hat{\theta}_{-j}$, $\hat{\theta}_j$ are asymptotically independent, for $2 \le j \le k$. It can also be verified that $\hat{\theta}_j$ and $\hat{\theta}_{-j}$ are asymptotically independent for different j, $1 \le j \le k$. Thus it follows that the asymptotic distribution of T_2 is a convolution of a chi-square random variable with 1 degree of freedom and $k-1$ independent chi-bar square random variables. The final form of the asymptotic distribution of T_2 is given in the following theorem.

THEOREM 3.2. *When H_0 is true, then for any real number c*

$$\lim_{n\to\infty} P(T_2 \ge c) = \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} \left(\frac{1}{2}\right)^{k-1} P(\chi_{\ell+1}^2 \ge c). \tag{3.11}$$

Note that the asymptotic null distribution of T_2 is free of $\boldsymbol{p}$. Also the distribution on the right side of (3.11) is the same as the one on the right side of (2.6), approximate critical values of which are given in Table 1.

4. Consistency of MLE under bias

If k is held fixed and $n \to \infty$, then it follows from (2.5) that $\hat{p}_i^* \to p_i$, $\forall i$ under (1.3). For a general univariate random variable X with CDF F, which has a probability density function f with respect to the Lebesgue measure, the positive bias given by (1.4) can be expressed as

$$\text{(4.1)} \qquad \frac{f(x)}{f(-x)} \text{ is nondecreasing in } x \text{ (a.e.)}, \qquad 0 < x < \infty.$$

If we define the maximum likelihood estimate in the generalized sense of Kiefer and Wolfowitz (1956) which allows probability to only be placed on the sample points, the problem reduces to the discrete case. In this section, we show that in the class of all univariate distributions the MLE of F under (4.1) is strongly consistent. This guarantees that the test T_1 developed in Section 2 is a consistent test.

We assume that we have a random sample of size n from F. Also, let $\hat{F}_n(\cdot)$ denote the usual empirical CDF, and $F_n^*(\cdot)$ denote the restricted MLE under (4.1). Our proof is a refinement of the one given by Dykstra *et al.* ((1995*b*), Section 3).

We fix ω (arbitrarily in a set of probability 1) such that $\hat{F}_n(x,\omega) \to F(x)$ *uniformly* in x. It will suffice to show for a fixed $\epsilon > 0$ and t, there exists $n(\epsilon,\omega)$ such that

$$|F_n^*(t,\omega) - \hat{F}_n(t,\omega)| < \epsilon$$

for $n \geq n(\epsilon,\omega)$ (we hereafter suppress ω).

Suppose that $t > 0$. We let $s_1, s_2, \ldots, s_k$ denote the positive sample points and assume that $a_1, a_2, \ldots, a_\nu$ ($a_0 \equiv 0$) denote the upper end points of the level sets of $\boldsymbol{\theta}^* = E_{\hat{v}}(\hat{\boldsymbol{\theta}} \mid \mathcal{I}_k)$, where $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{v}}$ are defined in (2.4). We assume that $a_{r-1} < t \leq a_r$. Then

$$F_n^*(t) - F_n^*(0) = \sum_{i:0<s_i\leq t} \frac{n_{-i}+n_i}{n} E_{\hat{v}}(\hat{\boldsymbol{\theta}} \mid \mathcal{I}_k)_i$$

$$\text{(4.2)} \qquad = \frac{1}{n}\sum_{j=1}^{r-1} \sum_{a_{j-1}<s_i\leq a_j} (n_{-i}+n_i)\left[\frac{\sum_{a_{j-1}<s_i\leq a_j} n_i}{\sum_{a_{j-1}<s_i\leq a_j}(n_i+n_{-i})}\right]$$

$$\text{(4.3)} \qquad + \left[\frac{\sum_{a_{r-1}<s_i\leq t}(n_{-i}+n_i)}{\sum_{a_{r-1}<s_i\leq t} n_i} \frac{\sum_{a_{r-1}<s_i<a_r} n_i}{\sum_{a_{r-1}<s_i<a_r}(n_{-i}+n_i)}\right] \cdot \frac{1}{n} \sum_{a_{r-1}<s_i\leq t} n_i$$

$$= \frac{1}{n}\sum_{j=1}^{r-1}\sum_{a_{j-1}<s_i\le a_j} n_i + A\frac{1}{n}\sum_{a_{r-1}<s_i\le t} n_i$$
$$= \hat{F}_n(a_{r-1}) - \hat{F}_n(0) + A(\hat{F}_n(t) - \hat{F}_n(a_{r-1}))$$
$$= \hat{F}_n(t) - \hat{F}_n(0) - (1-A)(\hat{F}_n(t) - \hat{F}_n(a_{r-1})) \tag{4.4}$$

where A is the entry in brackets in (4.3).

Following similar decomposition as in (4.2), it can be seen that $\hat{F}_n(0) = F_n^*(0)$. By the minimum lower sets algorithm, $0 \le A \le 1$ and thus $F_n^*(t) \le \hat{F}_n(t)$. Since $\hat{F}_n(t) \to F(t)$, there exists a positive integer n_0 such that for $\epsilon > 0$ and $n \ge n_0$,

$$|(\hat{F}_n(t) - \hat{F}_n(a_{r-1})) - (F(t) - F(a_{r-1}))| < \epsilon/2.$$

So if $F(t) - F(a_{r-1}) < \epsilon/2$, it follows from (4.4) that for $n \ge n_0$,

$$|F_n^*(t) - \hat{F}_n(t)| < \epsilon$$

since A is bounded.

Now suppose $F(t) - F(a_{r-1}) \ge \epsilon/2$. As $\hat{F}_n(t) \to F(t)$, uniformly, there exists a positive integer $n_1 \ge n_0$ such that for $n \ge n_1$,

$$\left| A - \frac{(F(t)+F(-t)-F(a_{r-1})-F(-a_{r-1}))(F(a_r)-F(a_{r-1}))}{(F(t)-F(a_{r-1}))(F(a_r)+F(-a_r)-F(a_{r-1})-F(-a_{r-1}))} \right| < \epsilon \tag{4.5}$$

since the denominator is bounded away from 0.

Now $f(x)/f(-x)$ is nondecreasing in x if and only if $f(x)/(f(x)+f(-x))$ is nondecreasing in x. If $f(t)/(f(t)+f(-t)) = c$, then

$$f(x) \ge (\le) c(f(x)+f(-x)) \quad \text{for} \quad x \ge (\le) t.$$

Integrating both sides, it follows that

$$\int_{(t,a_r]} f(x)dx > c\int_{(t,a_r]} (f(x)+f(-x))dx$$

and

$$\int_{(a_{r-1},t]} f(x)dx \le c\int_{(a_{r-1},t]} (f(x)+f(-x))dx.$$

Dividing both sides we obtain

$$\frac{F(a_r)-F(t)}{F(t)-F(a_{r-1})} \ge \frac{F(a_r)-F(t)+F(-a_r)-F(-t)}{F(t)-F(a_{r-1})+F(-t)-F(-a_{r-1})},$$

or, equivalently,

$$1 \le \frac{F(t)-F(a_{r-1})+F(-t)-F(-a_{r-1})}{F(t)-F(a_{r-1})} \cdot \frac{F(a_r)-F(a_{r-1})}{F(a_r)-F(a_{r-1})+F(-a_r)-F(-a_{r-1})}. \tag{4.6}$$

Using (4.6) and $F(t) - F(a_{r-1}) \geq \epsilon/2$, it follows from (4.5) that $A > 1 - \epsilon$, for $n \geq n_1$. Again it follows from (4.4) that

$$|F_n^*(t) - \hat{F}_n(t)| < \epsilon$$

for $n \geq n_1$. By following similar steps it follows that $|F_n^*(-t) - \hat{F}_n(-t)| < \epsilon$ for sufficiently large n. This proves that $F_n^*(\cdot)$ converges to $F(\cdot)$ uniformly (a.s.) as $n \to \infty$.

For a general univariate random variable X, the positive bias given by (1.6) can be expressed as

$$\frac{P(X > x)}{P(-X > x)} \text{ is nondecreasing in } x \text{ (a.e.)}, \qquad 0 < x < \infty. \tag{4.7}$$

However, the generalized MLE under (4.7) is not consistent except in the case of a discrete distribution. It can be seen by inspection that under (4.7) an observation at a positive x must be "shared" with $-x$ to preserve monotonicity while an observation at a negative x need not be. On the contrary, under (4.1), both positive and negative x's are "shared" to preserve monotonicity. The conclusion of inconsistency of the MLE for the restriction (4.7) is somewhat expected based on the work of Rojo and Samaniego (1991). Thus we recommend the test T_2 mainly for the discrete situation or a grouped data situation.

5. Simulated powers of tests

We perform a simulation study to compare the powers of the likelihood ratio tests of symmetry against alternatives H_1 and H_2 as derived in Sections 2 and 3 with other tests. Similar to Dykstra *et al.* (1995*a*), we consider the shifted binomial distribution given by

$$p_j = \binom{2k}{j+k} p^{j+k}(1-p)^{k-j}, \qquad j = 0, \pm 1, \ldots, \pm k,$$

with $k = 3$ for power comparisons. Thus there are 7 cells. It is easy to verify by algebra that the distribution is symmetric when $p = 0.5$ and it satisfies the alternatives H_1 and H_2 when $p > 0.5$. We take the sample size as fixed at $n = 100$ and replication size is 10,000. We take $\alpha = 0.05$.

Recall the likelihood ratio tests against alternatives H_1 and H_2 are denoted by T_1 and T_2, respectively. For the test T_1, the level probabilities under H_0 are calculated using the formula given in Robertson *et al.* (1988), and these are 0.3017, 0.5000, 0.1983. For the test T_2, the level probabilities under H_0 are 0.25, 0.50, 0.25.

For the shifted binomial distribution, the UMP test is to reject H_0 in favor of H_1 (and H_2) if $\sum_{i=1}^{n} X_i$ is too large. We will denote this test by T_3. The unrestricted likelihood ratio test for testing symmetry versus non-symmetry (also equivalent to the usual chi-square goodness-of-fit test for symmetry) will be denoted by T_4 (distributed as a chi-square with 3 degrees of freedom). We used

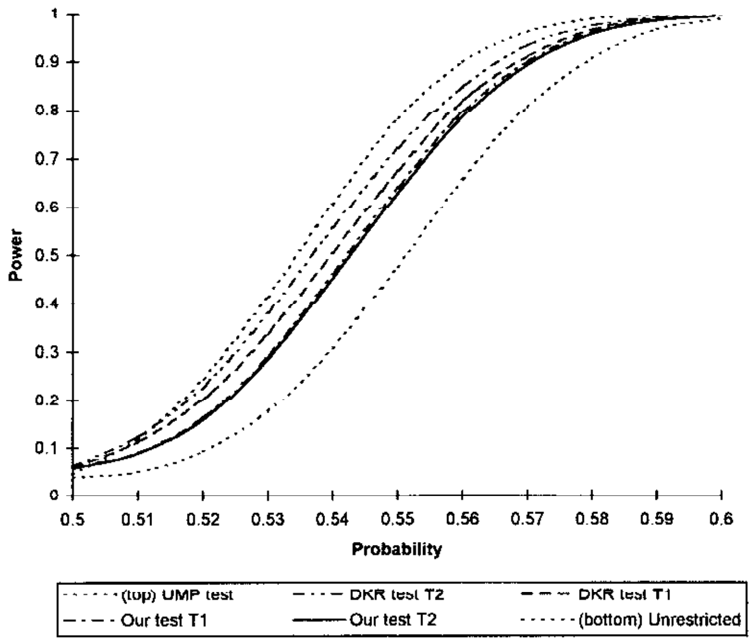

Fig. 1. Power curves for $k = 3$, $n = 100$, $\alpha = 0.05$. Legends: (top) dot = UMP test; dash-dot-dot = DKR test T_2; dash = DKR test T_1; dash-dot = Our test T_1; solid line = Our test T_2; (bottom) dot = Unrestricted test.

simulated critical values for power comparisons. Using the asymptotic 5% critical points, under the null hypothesis of symmetry the p-values of these tests are 0.0582, 0.0576, 0.0459 and 0.0396, respectively. The power functions of the proposed tests are shown in Fig. 1. Clearly the tests T_1 and T_2 perform much better than the unrestricted test T_4 and perform favorably with respect to the UMP test T_3. The test T_1 performs better than the test T_2 as the H_1 constraints are more restrictive than the H_2 constraints.

Also, we have considered the likelihood ratio tests against the alternatives (1.1) and (1.2) treated by Dykstra *et al.* (1995*a*). We will denote these tests by 'DKR test T_1' and 'DKR test T_2', respectively. Under the above context we calculated the power curves for these tests, and are shown in Fig. 1 as well. Recall that both of the DKR tests are designed to detect specific types of positive bias of the first kind, whereas both of our tests are designed to detect specific types of positive bias of the second kind. Here both of the DKR tests are found to have uniformly higher powers than both of our tests. However, this is not expected to happen in general, since neither of (1.1) or (1.2) implies (or, is implied by) any of (1.3) or (1.5).

6. Example

To illustrate the methods discussed in the earlier sections, we consider some data given in Clogg and Shockey (1988). The data, given in Table 2, consist of

Table 2. Classification of 95 nonwhites favoring President Carter in the 1980 Presidential vote based on their political views.

Political views*						
1	2	3	4	5	6	7
6	16	23	31	8	7	4

*Political views range from 1 = extremely liberal to 7 = extremely conservative.

the cross-classification of ninety five nonwhite people based on their political views who voted for President Carter in the 1980 Presidential vote. Political views range from 1 = extremely liberal to 7 = extremely conservative. We are interested in tests of symmetry about 4 against alternatives H_1 and H_2.

When testing H_0: symmetry versus the unrestricted alternative of nonsymmetry with the usual chi-square goodness of fit test, the test statistic is 11.1798 (p-value = 0.0108 from a χ_3^2 distribution).

To test H_0 : symmetry versus the alternative H_1, we first find $\bar{\boldsymbol{\theta}} = (0.2581, 0.3043, 0.4000) = \boldsymbol{\theta}^*$. Then we calculate the test statistic $T_1 = 11.5920$. To find its p-value, we estimate the v_i's under H_0, and then use the expressions from Robertson *et al.* (1988) to find the level probabilities which turn out to be 0.3185, 0.5000, 0.1815. Using these, the asymptotic p-value of the above T_1 is 0.0033.

When testing H_0 : symmetry versus the alternative H_2, then $\theta_1^0 = \theta_{-1}^0 = 0.3368$, $\theta_2^0 = \theta_{-2}^0 = 0.5156$ and $\theta_3^0 = \theta_{-3}^0 = 0.3030$. Also, $(\theta_{-1}^*, \theta_1^*) = (0.4737, 0.2000)$, $(\theta_{-2}^*, \theta_2^*) = (0.4889, 0.5789)$ and $(\theta_{-3}^*, \theta_3^*) = (0.2727, 0.3636)$. Using these values we compute the value of the test statistic $T_2 = 11.5920$. The asymptotic p-value is 0.0039. Clearly, there is strong evidence supporting both of the one-sided hypotheses H_1 and H_2 over H_0.

It is somewhat surprising that these two test statistics values came out to be the same. However, as the asymptotically least favorable distribution of T_1 is the same as the asymptotic distribution of T_2, these two p-values give an idea of how conservative the least favorable distribution of T_1 in this case is.

Also we calculated the DKR test T_1 and DKR test T_2 for this example. For DKR test T_1, the test statistic value is 0.7179 with a p-value of 0.6516 (obtained by using simulated level probabilities with the estimates under H_0). For DKR test T_2, the test statistic value is 0.0. Thus we fail to reject the null hypothesis of symmetry for both alternatives (1.1) and (1.2). So for this example, specific positive biases of the second kind are detected rather than those of the first kind.

7. Discussion

We have considered two particular types of alternatives to symmetry (positive bias of the second kind) and the corresponding likelihood ratio tests. Dykstra *et al.* (1995*a*) considered two similar alternatives to symmetry (positive bias of the first kind). It is noteworthy to compare the different behavior of our tests from theirs. First, our test T_1 and DKR test T_2 both are for alternatives based on ratios

p_i/p_{-i} (nondecreasing and greater than or equal to 1, respectively). Their test is asymptotically similar, ours is not. They get binomial weights, while ours are classic level probabilities. Their MLE's are not universally consistent, while ours are. On the other hand, our test T_2 and DKR test T_1 both are for alternatives based on ratios of tail probabilities (nondecreasing and greater than or equal to 1, respectively). In this case, the behavior just gets turned around. In particular, our test is asymptotically similar, theirs is not. We get binomial weights, they get the classic level probabilities (with the chi-square degrees of freedom reversed). Their MLE's are universally consistent, ours are not. Thus one might anticipate the presence of some kind of duality between these two types of constraints that gets the results reversed for different boundaries (constant 1 and nondecreasing). (These behaviors somewhat duplicate in the two-sample case also, see e.g. Dykstra *et al.* (1991); Robertson and Wright (1981); Dykstra *et al.* (1995*a*).) Further research in this direction is needed.

Acknowledgements

The author wishes to express his appreciation for the helpful comments and suggestions of two anonymous referees. In particular, the comments of one referee motivated the discussion in Section 7.

References

Aki, S. (1993). On nonparametric tests for symmetry in R^m, *Ann. Inst. Statist. Math.*, **45**, 787–800.

Bohrer, R. and Chow, W. (1978). Weights for one-sided multivariate inference, *Appl. Statist.*, **27**, 100–104.

Chaffin, W. W. and Rhiel, G. S. (1993). The effect of skewness and kurtosis on the one-sample t test and the impact of the knowledge of the population standard deviation, *J. Statist. Comput. Simulation*, **46**, 79–90.

Clogg, C. C. and Shockey, J. W. (1988). Multivariate analysis of discrete data, *Handbook of Multivariate Experimental Psychology* (eds. J. R. Nesselroade and R. B. Cattell), Plenum Press, New York.

Dykstra, R., Kochar, S. and Robertson, T. (1991). Statistical inference for uniform stochastic ordering in several populations, *Ann. Statist.*, **19**, 870–888.

Dykstra, R., Kochar, S. and Robertson, T. (1995*a*). Likelihood ratio tests for symmetry against one-sided alternatives, *Ann. Inst. Statist. Math.*, **47**, 719–730.

Dykstra, R., Kochar, S. and Robertson, T. (1995*b*). Inference for likelihood ratio ordering in the two-sample problem, *J. Amer. Statist. Assoc.*, **90**, 1034–1040.

Gastwirth, J. L. (1971). On the sign test for symmetry, *J. Amer. Statist. Assoc.*, **66**, 821–823.

Hettmansperger, T. (1984). *Statistical Inference Based on Ranks*, Wiley, New York.

Kiefer, J. and Wolfowitz, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many nuisance parameters, *Ann. Math. Statist.*, **27**, 887–906.

Lee, C. I. C., Robertson, T. and Wright, F. T. (1993). Bounds on distributions arising in order restricted inferences with restricted weights, *Biometrika*, **80**, 405–416.

Oh, M. S. (1994). Statistical tests concerning a set of multinomial parameters under order restrictions: approximations to null hypothesis distributions, unpublished Ph.D. dissertation, Dept. of Statistics, University of Iowa.

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

Rojo, J. and Samaniego, F. J. (1991). On nonparapetric maximum likelihood estimation of a distribution function uniformly stochastically smaller than a standard, *J. Stat. Prob. Letters.*, **11**, 267–271.

Rothman, E. D. and Woodroofe, M. (1972). A Cramer von Mises type statistic for testing symmetry, *Ann. Math. Statist.*, **43**, 2035–2038.

Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*, Wiley, New York.

Yanagimoto, T. and Sibuya, M. (1972). Test of symmetry of a one-dimensional distribution against positive biasedness, *Ann. Inst. Statist. Math.*, **24**, 423–434.