Inequalities in Statistics and Probability
IMS Lecture Notes–Monograph Series Vol. 5 (1984), 236–243

TESTS FOR AND AGAINST TRENDS AMONG POISSON INTENSITIES[1]

BY RHONDA MAGEL[2] and F. T. WRIGHT
North Dakota State University and University of Missouri-Rolla

Suppose one observes independent Poisson processes with unknown intensities λ_i, $i = 1, \ldots, k$, and that apriori it is believed that these intensities satisfy a known ordering. For preliminary analysis, it might be desirable to test for homogeneity among the intensities and, of course, one would want a test that utilizes the information in the ordering. Let t_i denote the length of time for which the ith process was observed. The case in which the t_i are equal has been studied in the literature. We develop the conditional likelihood ratio test for arbitrary t_i. This test is equivalent to the unconditional likelihood ratio test, but leads to an interesting multinomial testing situation, ie. testing for homogeneity of p_i/t_i versus a trend among the p_i/t_i, *where the* p_i are the cell probabilities. If the number of trials in the multinomial setting, or the total number of occurrences in the Poisson processes, is large, then the test statistic has an approximate chi-bar-squared distribution which has been studied in the literature. Results of a Monte Carlo study comparing this test with the maximin test developed by Lee (1980) are discussed. Similar results are also obtained for testing the null hypothesis that the intensities satisfy the prescribed ordering.

1. Introduction. Barlow, Bartholomew, Bremner and Brunk (1972) discuss the problem of estimating a finite sequence of Poisson intensities which are assumed to be nonincreasing. For instance, consider a system which is observed for t_1 units of time with X_1 failures, is then modified in an attempt to improve its performance, is observed for t_2 units of time with X_2 failures, is modified again, and this is repeated until it is observed for the kth time for t_k units of time with X_k failures. If it is believed that the modifications will not harm the system's performance, then one might wish to estimate the vector of intensities, $\lambda = (\lambda_1, \ldots, \lambda_k)$, subject to $\lambda_1 \geqslant \ldots \geqslant \lambda_k$. It would also be of interest to test for homogeneity among the intensities with the alternative $\lambda_1 \geqslant \ldots \geqslant \lambda_k$ and $\lambda_1 > \lambda_k$, or if the assumption concerning the modification were in question, one could test $\lambda_1 \geqslant \ldots \geqslant \lambda_k$ against $\lambda_i < \lambda_{i+1}$ for some i.

Suppose $X_1, \ldots, X_k$ are independent Poisson variables with means $\mu_i = \lambda_i t_i$, let $<<$ be a partial order on $\{1,2, \ldots, k\}$, let $\lambda^{(0)}$ be a fixed vector, let a be an unknown scale parameter and let H_0: $\lambda = a\lambda^{(0)}$, H_1: $\lambda_i \leqslant \lambda_j$ whenever $i << j$ and H_2: $\sim H_1$ (that is, $\lambda_i > \lambda_j$ for some $i << j$). The hypothesis H_1 stipulates that $\lambda = (\lambda_1, \ldots, \lambda_k)$ is isotonic (with respect to $<<$) and we suppose that $\lambda^{(0)}$ is isotonic. We consider the likelihood ratio test (lrt) for H_0 versus $H_1 - H_0$ and H_1 versus H_2 conditional on $\Sigma_{i=1}^{k} X_i = n$. While it will be shown that the conditional test is equivalent to the unconditional lrt, it does lead to an interesting multinomial testing situation. We also know that for $k = 2$ it is UMP unbiased. (See Ferguson (1967, p. 228)).

Robertson and Wegman (1978) consider order restricted tests for members of the exponential family, but their work requires that the sample sizes be equal. Their results can be applied in the testing situation considered here only if the t_i are all equal. Boswell (1966)

---

[1] This research was sponsored by the Office of Naval Research under ONR Contract N00014–80–C–0322.

[2] Parts of this work are taken from this author's doctoral dissertation written at the University of Missouri-Rolla.

AMS 1980 subject classification. Primary 62F03, secondary 62E20.

Key words and phrases: Order restricted inferences, trends in Poisson intensities, trends in multinomial parameters.

considers a closely related problem and develops the conditional lrt for testing that the intensity of a nonhomogeneous Poisson process is constant versus it is nondecreasing.

Lee (1980) developed a maximin test for the multinomial setting arising in the conditional framework. It could be used to provide a test of H_0 versus $H_1 - H_0$. The results of a Monte Carlo study comparing the lrt and the maximin test are given in Section 4. It is found that for small k ($k = 3,4,5$), the two perform similarly with the power of the maximin test larger for "regular" alternatives and the power of the lrt larger for "nonregular" alternatives. Furthermore, the differences in power are not too large. However, for larger k ($k = 10$) the differences are more pronounced and if nonregular alternatives cannot be ruled out the lrt should be used.

2. Estimation. The maximum likelihood estimate (mle) of λ subject to H_1 can be expressed as a projection onto a cone of isotonic functions. We introduce some notation. With $<<$ a fixed partial order on $\{1,2, \ldots , k\}$, let $\mathcal{R}_k$ denote the k-dimensional reals, let

$$C = \{x \in \mathcal{R}_k : x \text{ is isotonic with respect to } <<\},$$

let $\mathbf{w} = (w_1,w_2, \ldots , w_k)$ be a vector of positive weights, for $\mathbf{x},\mathbf{y} \in \mathcal{R}_k$ let $(x,y)_w$ denote the inner product $\Sigma_{i=1}^k w_i x_i y_i$, and for $\mathbf{y} \in \mathcal{R}_k$ let $E_\mathbf{w}(\mathbf{y}|C)$ denote the projection of $\mathbf{y}$ onto C, that is $E_\mathbf{w}(\mathbf{y}|C)$ minimizes

$$\Sigma_{i=1}^k w_i(y_i - x_i)^2 \quad \text{for } \mathbf{x} \in C.$$

Theorems 1.4 and 1.5 of Barlow, Bartholomew, Bremner and Brunk (1972) state that

$$\Sigma_{i=1}^k w_i E_w(\mathbf{y}|C)_i = \Sigma_{i=1}^k w_i y_i \tag{2.1}$$

and

$$\Sigma_{i=1}^k w_i(y_i - E_w(\mathbf{y}|C)_i) E_\mathbf{w}(\mathbf{y}|C)_i = 0 \tag{2.2}$$

The mle of λ subject to H_1 is $\bar{\lambda} = E_t(\mathbf{X}/\mathbf{t}|C)$ where $\mathbf{t} = (t_1, \ldots , t_k)$, $\mathbf{X} = (X_1, \ldots , X_k)$ and for $\mathbf{x},\mathbf{y} \in \mathcal{R}_k$, $\mathbf{x}/\mathbf{y} = (x_1/y_1, \ldots , x_k/y_k)$ and $\mathbf{xy} = (x_1y_1, \ldots , x_ky_k)$. (See Barlow, Bartholomew, Bremner and Brunk (1972, p. 44)).

Conditioning on $Y = \Sigma_{i=1}^k X_i = n$, the density of $\mathbf{X}$ is that of a multinomial with parameters n and

$$p_i = \lambda_i t_i / \Sigma_{j=1}^k \lambda_j t_j \quad \text{for } i = 1, \ldots , k. \tag{2.3}$$

So H_0: $\lambda = a\lambda^{(0)}$ is equivalent to H'_0: $p_i = p_i^{(0)} = \lambda_i^{(0)} t_i / \Sigma_{j=1}^k \lambda_j^{(0)} t_j$, $1 \leq i \leq k$,

H_1: λ is isotonic is equivalent to H'_1: $\mathbf{p}/\mathbf{t}$ is isotonic, and

H_2: λ is not isotonic is equivalent to H'_2: $\mathbf{p}/\mathbf{t}$ is not isotonic.

The mle of $\mathbf{p}$ subject to H'_1 is also of interest.

THEOREM 1. *The mle of* $\mathbf{p}$ *subject to* H'_1 *is given by* $\bar{\mathbf{p}} = \mathbf{t}E_\mathbf{t}(\mathbf{X}/\mathbf{t}|C)/n$. *These* $\bar{p}_i$ *satisfy* $\bar{p}_i \geqslant 0$ *and* $\Sigma_{i=1}^k \bar{p}_i = 1$, *and* $\bar{\mathbf{p}} \to \mathbf{p}$ *almost surely provided* $\mathbf{p}/\mathbf{t}$ *is isotonic.*

Proof. The mle of $\mathbf{p}$ under H'_1 maximizes $\Pi_{i=1}^k (p_i/t_i)^{x_i}$ subject to $\mathbf{p}/\mathbf{t}$ is isotonic, $p_i \geqslant 0$ and $\Sigma_{i=1}^k p_i = 1$. Applying the result in Barlow, Bartholomew, Bremner and Brunk (1972, p. 46), we see that $\bar{\mathbf{p}} = \mathbf{t}E_\mathbf{t}(\mathbf{X}/\mathbf{t}|C)/n$. Since the projection $E_\mathbf{w}(\cdot|C)$ is continuous for fixed C and $\mathbf{w}$, $\bar{\mathbf{p}} = \mathbf{t}E_\mathbf{t}(\mathbf{X}/(n\mathbf{t}|C) \to \mathbf{t}E_\mathbf{t}(\mathbf{p}/\mathbf{t}|C)$ almost surely. If $\mathbf{p}/\mathbf{t}$ is isotonic the right hand side (rhs) is $\mathbf{p}$ and the proof is completed. □

If $<<$ is a total order, then the pool-adjacent-violators algorithm can be used to compute the projection in the formula for $\bar{\mathbf{p}}$ and the lower sets algorithm can be used for an arbitrary partial order. (See Chapter 2 of Barlow, Bartholomew, Bremner and Brunk (1972)).

3. Tests of Hypotheses. As was mentioned in the Introduction, we consider the conditional lrts of H_0 versus H_1–H_0 and of H_1 versus H_2. However, these lead to the lrts of H'_0 versus H'_1–H'_0 and of H'_1 versus H'_2. Chacko (1966) developed the lrt and an asymptotically equivalent modified χ^2 test for H'_0 versus H'_1–H'_0 in the totally ordered case with $\mathbf{p}^{(0)}$ and $\mathbf{t}$ constant vectors. Robertson (1978) developed both lrts for partial orders with $\mathbf{t}$ a constant vector.

Denoting the conditional likelihood ratios by λ'_{01} and λ'_{12}, we see that

(3.1) $$T'_{01} = -2\,ln\,\lambda'_{01} = 2\Sigma_{i=1}^{k} X_i\{lnE_{\mathbf{t}}(\mathbf{X}/\mathbf{t}|C)_i - ln(p_i^{(0)}/t_i\} - 2n\,ln\,n$$

and

(3.2) $$T'_{12} = -2ln\lambda'_{12} = 2\Sigma_{i=1}^{k} X_i\{ln(X_i/t_i) - lnE_{\mathbf{t}}(\mathbf{X}/\mathbf{t}|C)_i\}.$$

Hence, the test statistic for H_0 versus $H_1 - H_0$ is obtained by replacing $p_i^{(0)}$ with $\lambda_i^{(0)}t_i/\Sigma_{j=1}^{k}\lambda_j^{(0)}t_j$ in (3.1) and for testing H_1 versus H_2, $T_{12} = T'_{12}$.

Remark 2. The conditional lrts, T_{01} and T_{12}, are also the unconditional lrts.

Proof. Under H_0, the mle of a is $n/\Sigma_{j=1}^{k}\lambda_j^{(0)}t_j$. Using (2.1) and straight-forward algebra and denoting the likelihood ratio for testing H_0 versus $H_1 - H_0$ (H_1 versus H_2) by λ_{01} (λ_{12}), one can show that $-2\,ln\,\lambda_{01} = T_{01}$ and $-2\,ln\,\lambda_{12} = T_{12}$. □

Next the large sample distributions of these test statistics are determined. The derivations are like those given in Robertson (1978) and in fact we will make use of several lemmas proved there.

LEMMA 3. In the multinomial setting, let Z_i be independent normal variables with mean 0 and variance $1/p_i$ and let $\bar{Z} = \Sigma_{i=1}^{k} p_i Z_i$. As $n \to \infty$, $\sqrt{n}\,(\hat{p}-p) \xrightarrow{\mathcal{D}} (p_1(Z_1-\bar{Z}), \ldots, p_k Z_k-\bar{Z}))$. Let q denote a fixed vector of probabilities with $\mathbf{q}/\mathbf{t}$ isotonic, let D denote the closed, convex cone

$$D = \{x \in \mathcal{R}_k : x_i \leq x_j \text{ if } i << j \text{ and } q_i/t_i = q_j/t_j\},$$

let $r_1 < \ldots < r_h$ denote the distinct values among t_i, $i = 1, \ldots, k$, and let $M(i) = \{j : q_j/t_j = r_i\}$ for $i = 1, \ldots, h$. Robertson (1978) observed that $E_{\mathbf{w}}(\mathbf{x}|D)$ can be computed by independently computing its values for subscripts in $M(i)$ for $i = 1, \ldots, h$. This implies that if $\mathbf{y} \in \mathcal{R}_k$ has positive entires and is constant on each $M(i)$, then

(3.3) $$E_{\mathbf{w}}(\mathbf{x}|D) = E_{\mathbf{wy}}(\mathbf{x}|D).$$

(This can be easily seen by considering the lower-sets algorithm.)

LEMMA 4. If $\mathbf{x},\mathbf{y} \in \mathcal{R}_k$ with $\mathbf{y}$ constant on each $M(i)$, then $E_{\mathbf{w}}(\mathbf{x}-\mathbf{y}|D) = E_{\mathbf{w}}(\mathbf{x}|D) - \mathbf{y}$. If in addition, $\mathbf{y}$ has nonnegative entries, then $E_{\mathbf{w}}(\mathbf{xy}|D) = \mathbf{y}E_{\mathbf{w}}(\mathbf{x}|D)$.

LEMMA 5. If $\mathbf{x} \in \mathcal{R}_k$ with $\max_{j \in M(i)} x_i < \min_{j \in M(i+1)} x_j$ for $i = 1, \ldots, h-1$, then $E_{\mathbf{w}}(\mathbf{x}|C) = E_{\mathbf{w}}(\mathbf{x}|D)$.

The cone D is determined by $<<$ and q. For a given D, define $P_{\mathbf{w}}(\ell,k)$ to be the probability of exactly ℓ distinct values in $E_{\mathbf{w}}(V|D)$ where $\mathbf{V} = (V_1, \ldots, V_k)$, with the V_i independent normal variables with mean 0 and variance $1/w_i$. *If* $<<$ is the usual total order on $1, \ldots, k$ and q_i/t_i is constant, then $D = \{x : x_1 \leq \ldots \leq x_k\}$ and the $P_w(\ell,k)$ are discussed in detail in Barlow, Bartholomew, Bremner and Brunk (1972). Approximations for nonconstant $\mathbf{w}$ are discussed in Siskind (1976) and Robertson and Wright (1983). If $<<$ is the same total order, but q_i/t_i is not constant, then the results mentioned above can be used to determine these probabilities for each $M(i)$ and the h-fold convolution gives the desired $P_w(\ell,k)$. (Cf.

Barlow, Bartholomew, Bremner and Brunk (1972)). The latter reference also discusses the the $P_w(\ell,k)$ for some partial orders. Let χ^2_ℓ denote a chi-squared variable with ℓ degrees of freedom ($\chi^2_0 \equiv 0$).

THEOREM 6. *If in the Poisson setting, the vector of intensities is a* $\lambda^{(0)}$ *and* $\mathbf{q} = \mathbf{p}^{(0)}$ (cf. (2.3)), *or in the multinomial setting, if the probability vector is* $\mathbf{q} = \mathbf{p}^{(0)}$ *and if D is determined by* $\mathbf{q}$, *then for* $t \geq 0$,

(3.4) $$\lim_{n\to\infty} P[T_{01} \geq t|Y=n] = \lim_{n\to\infty} P[T'_{01} \geq t] = \Sigma^k_{\ell=1} P_q(\ell,k) P[\chi^2_{\ell-1} \geq t].$$

If in the Poisson setting, the vector of intensities is of the form $a\nu$ with ν isotonic and $\mathbf{q}$ is determined by (2.3) with $\lambda^{(0)}$ replaced by ν, or if in the multinomial setting the vector of probabilities, $\mathbf{q}$, is such that $\mathbf{q}/\mathbf{t}$ is isotonic, and if D is determined by $\mathbf{q}$, then for $t \geq 0$,

(3.5) $$\lim_{n\to\infty} P[T_{12} \geq t|Y=n] = \lim_{n\to\infty} P[T'_{12} \geq t] = \Sigma^k_{\ell=1} P_q(\ell,k) P[\chi^2_{k-\ell} \geq t],$$

(3.6) $$\lim_{n\to\infty} P[T_{12} \geq t|Y=n] \leq \lim_{n\to\infty} P_e[T_{12} \geq t|Y=n], \text{ and}$$

(3.7) $$\lim_{n\to\infty} P[T'_{12} \geq t] \leq \lim_{n\to\infty} P_\delta[T'_{12} \geq t],$$

where in the Poisson setting $P_e[\ ]$ denotes the probability under $\lambda = (1,1, \ldots, 1)$ and in the multinomial setting, $P_\delta[\ ]$ denotes the probability under $p_i = t_i/\Sigma^k_{j=1} t_j$, $i = 1,2, \ldots, k$.

Comments. In the Poisson setting, if one wishes to test H_0 versus $H_1 - H_0$, then $\mathbf{q}$ is set equal to $\mathbf{p}^{(0)}$ determined by (2.3). The vector $\mathbf{q}$ determines D which in turn determines the $P_{\mathbf{q}}(\ell,k)$ and so large sample p-values can be calculated for the conditional test from (3.4). In testing H_1 versus H_2, (3.6 indicates that the asymptotically least favorable configuration in H_1 is $\lambda = (1,1, \ldots, 1)$ and so with $q_i = t_i/\Sigma^k_{j=1} t_j$, $D = C$ and approximate p-values can be computed from (3.5).

In the multinomial setting, in testing H'_0 versus $H'_1 - H'_0$, set $q = p^{(0)}$ and use (3.4) to determine large sample p-values. In testing H'_1 versus H'_2, the asymptotically least favorable configuration in H_1 is $q_i = t_i/\Sigma^k_{j=1} t_j$ in which case $D = C$ and approximate p-values can be computed from (3.5).

In either case, one might not want to use the asymptotically least favorable configuration when testing H_1 versus H_2 or H'_1 versus H'_2, so one could use the restricted mles of λ or $\mathbf{p}$ rather than $\mathbf{e}$ or δ.

Proof. T_{01} is defined to be T'_{01} with $p^{(0)}_i$ replaced by $\lambda^{(0)}_i t_i/\Sigma^k_{j=1} \lambda^{(0)}_j t_j$ and $T_{12} = T'_{12}$. Furthermore, conditional on $Y = n$, $(X_1, \ldots, X_k)$ is multinomial with parameters n and $p_i = \lambda_i t_i/\Sigma^k_{j=1} \lambda_j t_j$ and $H_i \equiv H'_i$, $i = 0,1,2$. So we need only consider the multinomial situation. We first consider the distribution of T'_{01} under H'_0. Setting $\hat{p}_i = X_i/n$, expressing T'_{01} as

$$2n\Sigma^k_{i=1} \hat{p}_i \{ln E_{\mathbf{t}}(\hat{\mathbf{p}}/\mathbf{t}|C)_i - ln(p^{(0)}_i/t_i)\},$$

and expanding $ln\, E_{\mathbf{t}}(\hat{\mathbf{p}}/\mathbf{t}|C)_i$ and $ln\,(p^{(0)}_i/t_i)$ about $\hat{p}_i/t_i$, we write T'_{01} as

(3.8) $$2n\Sigma^k_{i=1} t_i \{E_{\mathbf{t}}(\hat{\mathbf{p}}/\mathbf{t}|C)_i - p^{(0)}_i/t_i\} + n\Sigma^k_{i=1} \hat{p}_i \{((\hat{p}_i - p^{(0)}_i)/t_i)^2/\beta^2_i - \mathrm{E}_{\mathbf{t}}(\hat{\mathbf{p}}/\mathbf{t}|C)_i - \hat{p}_i/t_i)^2/\alpha^2_i\},$$

where $\alpha_i(\beta_i)$ is between p_i/t_i and $E_{\mathbf{t}}(\mathbf{p}/\mathbf{t}|C)_i$ $(p^{(0)}_i/t_i)$. Under H'_0, both $E_{\mathbf{t}}(\mathbf{p}/\mathbf{t}|C)_i$ and p_i/t_i converge almost surely to $p^{(0)}_i/t_i$. Recall, (2.1) implies that the first term in (3.8) vanishes.

Since $\hat{\mathbf{p}}/\mathbf{t}$ is consistent for $\mathbf{p}^{(0)}/\mathbf{t}$ under H'_0, there is for almost all ω in the underlying probability space an N, possibly depending on ω, for which $\hat{\mathbf{p}}/\mathbf{t}$ satisfies the hypothesis of Lemma 5 for $n \geq N$. Hence, for such ω and n, Lemmas 4 and 5 can be applied to the second term in (3.8) to obtain

$$\Sigma_{i=1}^{k}\hat{p}_i\{[\sqrt{n}\ (\hat{p}_i-p_i^{(0)})/t_i\beta_i]^2-[(t_iE_t(\sqrt{n}\ (\hat{\mathbf{p}}-\mathbf{p}^{(0)})/\mathbf{t}|D)_i-\sqrt{n}\ (\hat{p}_i-p_i^{(0)})/t_i\alpha_i]^2\}$$

which converges in distribution to

$$\Sigma_{i=1}^{k}(t_i^2/p_i^{(0)})\{((p_i^{(0)}/t_i)(Z_i-\bar{Z}))^2-(E_t((p^{(0)}/\mathbf{t})(\mathbf{Z}-\bar{\mathbf{Z}})|D)_i-(p_i^{(0)}/t_i)(Z_i-\bar{Z}))^2\}.$$

Applying Lemma 4 and (3.3), this can be written as

$$\Sigma_{i=1}^{k}p_i^{(0)}(\mathbf{Z}_i-\bar{\mathbf{Z}})^2-\Sigma_{i=1}^{k}p_i^{(0)}(\mathbf{Z}_i-E_{p(0)}(\mathbf{Z}|D)_i)^2 \tag{3.9}$$
$$=\Sigma_{i=1}^{k}p_i^{(0)}(E_{p(0)}(\mathbf{Z}|D)_i-\bar{\mathbf{Z}})^2 \quad + \quad 2\Sigma_{i=1}^{k}p_i^{(0)}(\mathbf{Z}_i-E_{p(0)}(\mathbf{Z}|D)_i)(E_{p(0)}(\mathbf{Z}|D)_i-\bar{\mathbf{Z}}).$$

The second term in the rhs of (3.9) can be shown to be zero using (2.1) and (2.2), and Theorem 3.1 of Barlow, Bartholomew, Bremner and Brunk (1972) shows that the first term on the rhs has the desired distribution.

Next we consider the distribution of T'_{12} with q a fixed probability vector for which $\mathbf{q}/\mathbf{t}$ is isotonic. Writing T'_{12} as $2n\Sigma_{i=1}^{k}\hat{p}_i\ ln(\hat{p}_i/t_i)-ln\ E_t(\hat{\mathbf{p}}/\mathbf{t}|C)_i$, expanding $lnE_t(\mathbf{p}/\mathbf{t}|C)_i$ about $\hat{p}_i/t_i$ and applying (2.1), T'_{12} can be written as

$$\Sigma_{i=1}^{k}\hat{p}_i(\sqrt{n}\ (E_t(\hat{\mathbf{p}}/\mathbf{t}|C)_i-\hat{p}_i/t_i))^2/\gamma_i^2$$

where γ_i is between $E_t(\hat{\mathbf{p}}/\mathbf{t}|C)_i$ and $\hat{p}_i/t_i$ and hence converges almost surely to q_i/t_i. As in the proof of the first part of the theorem, for almost all ω and n sufficiently large $\sqrt{n}\ (E_t(\hat{\mathbf{p}}/\mathbf{t}|C)-\mathbf{q}/\mathbf{t})=E_t(\sqrt{n}\ (\hat{\mathbf{p}}-\mathbf{q})/\mathbf{t}|D)$.

Hence, T'_{12} converges in distribution to

$$\Sigma_{i=1}^{k}(t_i^2/q_i)(E_t((\mathbf{q}/\mathbf{t})(\mathbf{Z}-\bar{\mathbf{Z}})|D)_i-(q_i/t_i)(Z_i-\bar{Z}))^2=\Sigma_{i=1}^{k}q_i(E_t(\mathbf{Z}|D)_i-\bar{Z}_i)^2. \tag{3.10}$$

Applying (3.3) this becomes

$$\Sigma_{i=1}^{k}q_i(E_q(Z|D)_i-Z_i)^2$$

and this has the desired distribution (cf. Theorem 2.5 of Robertson and Wegman (1978)).

We establish (3.7) to conclude the proof. The variables $U_i=\sqrt{q_i/t_i}\ Z_i$ are independent normal variables with means zero and $\text{var}(U_i)=1/t_i$. Using (3.3), the rhs of (3.10) can be written as $\Sigma_{i=1}^{k}t_i(E_t(\mathbf{U}|D)-U_i)^2$, which is the distance from $E_t(\mathbf{U}|D)$ to $\mathbf{U}$. Since $C\subset D$, by the definition of projection this is maximized for $D=C$, which occurs if $\mathbf{q}/\mathbf{t}$ is constant, ie. $q_i=t_i/\Sigma_{i=1}^{k}t_j$. □

In the multinomial setting, Lee (1980) developed a maximin test for $p_1/t_1=p_2/t_2=\ldots=p_k/t_k$ versus $p_{i+1}/t_{i+1}\geqslant dp_i/t_i$, $i=1,2,\ldots,k-1$, with $d>1$. The test statistic is $S_{01}=\Sigma_{i=1}^{k}iX_i$, which has an approximate normal distribution and under this null hypothesis its mean and variance are

$$n\Sigma_{i=1}^{k}it_i/\Sigma_{i=1}^{k}t_i \quad \text{and}\ n\{\Sigma_{i=1}^{k}i^2t_i/\Sigma_{i=1}^{k}t_i-(\Sigma_{i=1}^{k}it_i/\Sigma_{i=1}^{k}t_i)^2\},$$

respectively. The tests S_{01} and T'_{01} are compared in the next section.

4. Comparison of the lrt and Maximin Tests. The maximin test is a contrast test and Section 4.2 of Barlow, Bartholomew, Bremner and Brunk (1972) contains a discussion of the use of the likelihood ratio and contrast approaches in testing for trends among normal means. They concluded that, while the contrast test is typically much easier to use, the lrt provides the most satisfactory general way of incorporating prior information about ordering. In the case of a total order and small k, the contrast statistic provides a suitable alternative. If additional information is available about the spacing of the parameters, then a contrast test based on this additional information may be preferred.

To give some idea of the differences in the power for the two tests a Monte Carlo study was conducted. In particular, if one were testing

$$p_i = t_i/\Sigma_{j=1}^k t_j,\ 1 \leq i \leq k, \text{ versus } p_1/t_1 \leq p_2/t_2 \leq \ldots \leq p_k/t_k \text{ with } p_1/t_1 < p_k/t_k,$$

then T'_{01}, with $p_i^{(0)}$ replaced by $t_i/\Sigma_{j=1}^k t_j$, and S_{01} could be used. Because the distribution of T'_{01} under H'_1 is quite complex, Monte Carlo experiments were conducted. With $k = 3, 4, 5, 10$, $n = 25, 80$, nominal levels of .1, .05, and various choices of **t** and **p**, the powers of T'_{01} and S_{01} were approximated based on 5000 repetitions. Some of these values are given in Tables 1 and 2. To assess the accuracy of the approximations for the distributions under the null hypothesis, the estimates of the power under the null hypothesis are included. For $n = 80$ and the cases presented in these tables the largest discrepancy in the α level for $T'_{01}(S_{01})$ and a nominal level of .1 is .016 (.009) and for $\alpha = .05$ it is .005 (.007). For $n = 25$ the maximum discrepancies were larger but both approximations seem to be useful for k in this range. However, for $k = 10$ and $n = 25$ the approximation for the distribution of T'_{01} seems to give a test that is quite conservative. Its estimated α level is .065 (.031) when the nominal level was .1 (.05). The approximation for S_{01} seemed quite adequate even with $n = 25$ and $k = 10$. Its estimated α levels are .100 (.046), respectively.

It is clear from Tables 1 and 2 that neither test is uniformly better than the other. In fact, when the p_i/t_i increase regularly, such as in the cases **p/t** = (.25, .30, .45), **p/t** = (.20, .25, .35), **p/t** = (.15, .20, .30, .35), **p/t** = (.10, .15, .20, .25, .30), etc., then the maximin test outperforms T'_{01}, but for irregular increases in **p/t**, such as in the cases **p/t** = (.25, .25, .50), **p/t** = (.2, .3, .3), **p/t** = (.052, .052, .120), **p/t** = (.04, .04, .04, .07$\bar{3}$), **p/t** =(.03, .03, .07, .07), **p/t** = (.15, .2, .2, .25), etc., T'_{01} has greater power than S_{01}. For $k = 3,4,5$ the differences in power are not too large and the magnitudes are similar in both directions. So for small k one could use S_{01} if the alternative were believed to be "regular" in the sense described above or T'_{01} could be used if it is desirable to protect againt nonregular alternatives.

It is interesting to note that for $k = 3$ and 4 the above conclusions held whether the vector t was constant or not. (Several other choices of p and t, not given in Tables 1 and 2, were considered and these conclusions were substantiated in those cases, also.) For this reason only constant t's were considered for $k = 5$ and 10. Recall that for $k \geq 5$, the $P_{\mathbf{t}}(\ell,k)$ are intractable for nonconstant **t**.

Power comparisons were made for $k = 10$, but with $n = 25$ they were not very meaningful because of the conservative nature of the approximation to the null distribution of T'_{01}. For $n = 80$ both approximations were very reasonable and so power comparisons could be made in that case. Linear alternatives were considered. For $p_i = i/55$ both tests have powers that are essentially one. So $p_i = i/110 + .05$ was considered. The tests with nominal level .1 had powers .822 and .861, and the tests with nominal level .05 had powers .712 and .759. Of course, the maximin test performed better for such an alternative. The alternative $p_i = .09$, $1 \leq i \leq 9$, and $p_{10} = .19$ was considered. The estimated powers for the tests with $\alpha = .1$ are .700 and .556 and for the tests with $\alpha = .05$ they are .578 and .409. Finally the alternative $p_1 = .07$, $p_2 = \ldots = p_9 = .1$, and $p_{10} = .13$ was considered, and the approximate powers for the $\alpha = .1$ tests are .377 and .332. For the $\alpha = .05$ tests they are .252 and .199. Of course, the last two alternatives are nonregular and the lrt has the larger power. In these cases, the increase in power may be as large as 40 percent and so the lrt should definitely be considered to guard against nonregular alternatives for larger k.

TABLE 1. Estimated Powers of the Maximin and LRTs, $k=3$

	$k=3$							
	$n=25$		$n=80$		$n=25$		$n=80$	
	Nominal Level							
	.10	.05	.10	.05	.10	.05	.10	.05
	t=(1,1,1), **p/t**=(1/3,1/3,1/3)				**t**=(2,1,1), **p/t**=(.25,.25,.25)			
T'_{01}	.084	.059	.108	.047	.101	.047	.085	.052
S_{01}	.088	.056	.095	.043	.081	.049	.100	.044
	t=(1,1,1), **p/t**=(.25,.30,.45)				**t**=(2,1,1), **p/t**=(.20,.25,.35)			
T'_{01}	.389	.227	.718	.565	.459	.315	.771	.664
S_{01}	.451	.350	.812	.689	.423	.337	.802	.677
	t=(1,1,1), **p/t**=(.25,.25,.50)				**t**=(2,1,1), **p/t**=(.2,.3,.3)			
T'_{01}	.559	.445	.937	.871	.346	.220	.624	.509
S_{01}	.572	.475	.917	.843	.304	.230	.633	.473
	t=(2,3,5), **p/t**=(.10,.10,.10)				**t**=(2,3,5), **p/t**=(.05,.10,.12)			
T'_{01}	.126	.046	.093	.054	.571	.332	.891	.803
S_{01}	.095	.056	.106	.048	.504	.386	.889	.778
	t=(8,2,8), **p/t**=(1/18,1/18,1/18)				**t**=(8,2,8), **p/t**=(.04,.06,.07)			
T'_{01}	.112	.051	.084	.052	.527	.351	.832	.754
S_{01}	.080	.053	.100	.051	.452	.368	.855	.763
	t=(4,6,4), **p/t**=(1/14,1/14,1/14)				**t**=(4,6,4), **p/t**=(.052,.052,.120)			
T'_{01}	.112	.067	.101	.048	.758	.609	.981	.947
S_{01}	.117	.040	.100	.046	.723	.535	.965	.921

TABLE 2. Estimated Powers of the Maximin and LRTs, $k=4,5$

	$k=4$							
	$n=25$		$n=80$		$n=25$		$n=80$	
	Nominal Level							
	.10	.05	.10	.05	.10	.05	.10	.05
	t=(1,1,1,1),**p/t**=(.25,.25,,25,.25)				**t**=(4,6,4,6), **p/t**=(.05,.05,.05,.05)			
T'_{01}	.089	.055	.097	.051	.099	.054	.095	.055
S_{01}	.105	.050	.101	.050	.086	.042	.103	.049
	t=(1,1,1,1), **p/t**=(.15,.20,.30,.35)				**t**=(4,6,4,6), **p/t**=(.034,.042,.0495,.069)			
T'_{01}	.578	.446	.930	.860	.487	.338	.826	.729
S_{01}	.634	.482	.945	.883	.480	.342	.860	.757
	t=(1,1,1,1), **p/t**=(.10,.25,.25,.40)				**t**=(4,6,4,6), **p/t**=(.04,.04,.04,.07/3)			
T'_{01}	.752	.665	.995	.985	.501	.345	.844	.766
S_{01}	.804	.668	.993	.981	.464	.333	.834	.727
	t=(1,1,1,1), **p/t**=(.23,.23,.23,.31)				**t**=(4,6,4,6), **p/t**=(.03,.03,.07,.07)			
T'_{01}	.208	.129	.376	.256	.730	.562	.983	.955
S_{01}	.240	.143	.388	.245	.684	.544	.981	.946
	$k=5$ and **t**=(1,1,1,1)							
	p/t=(.2,.2,.2,.2,.2)				**p/t**=(.10,.15,.20,.25,.30)			
T'_{01}	.094	.050	.099	.049	.649	.489	.966	.930
S_{01}	.086	.051	.091	.050	.679	.567	.974	.946
	p/t=(.15,..15,.15,.15,.40)				**p/t**=(.15,.2,.2,.2,.25)			
T'_{01}	.711	.613	.985	.969	.264	.172	.496	.356
S_{01}	.654	.555	.955	.924	.252	.176	.488	.361

REFERENCES

BARLOW, R. E. BARTHOLOMEW, D. J., BREMNER, J. M. and BRUNK, H. D. (1972). *Statistical Inferences Under Order Restrictions*. Wiley, New York.

BOSWELL, M. T. (1966). Estimating and testing trend in a stochastic process of Poisson type. *Ann. Math. Statist. 37* 1564–1573.

CHACKO, V. J. (1966). Modified chi-square tests for ordered alternatives. *Sankyā Ser. B 28,* 185–190.

FERGUSON, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach*. Academic Press, New York.

LEE, Y. J. (1980). Tests of trend in count data: multinomial distribution case. *J. Amer. Statist. Assoc. 72* 673–675.

SISKIND, V. (1976). Approximate probability integrals and critical values for Bartholomew's test for ordered means. *Biometrika 63* 647–654.

ROBERTSON, Tim (1978). Testing for and against an order restriction on multinomial parameters. *J. Amer. Statist. Assoc. 73* 197–202.

ROBERTSON, Tim and WEGMAN, E. J. (1978). Likelihood ratio rests for order restrictions in exponential families. *Ann. Math. Statist. 6* 485–505.

ROBERTSON, Tim and WRIGHT, F. T. (1983). On approximation of the level probabilities and associated distributions in order restricted inference. *Biometrika 70* 597–606.