The Annals of Statistics
1982, Vol. 10, No. 2, 569–574

RANK TESTS GENERATED BY CONTINUOUS PIECEWISE LINEAR FUNCTIONS

BY W. J. R. EPLETT

University of Oxford

The two-sample problem of testing against location shift is fundamental to much of the theory of rank tests. Generally testing and estimation is carried out with a fixed (non-random) set of scores for the ranks. However Beran (1974), following ideas of Stein and Hájek, developed a notable class of adaptive estimators. When used in testing, these give asymptotically efficient tests, regardless of the underlying distribution. These ideas are used here to focus attention upon tests generated by continuous, piecewise linear functions (called PLRT's) which provide a practically useful class of asymptotically efficient adaptive rank tests. Under suitable conditions the rate of convergence of the consistent estimators of the score generating function is $O(N^{-1/2})$ which suggests they are quite suitable for practical application when N is large. A Riesz representation theorem for the asymptotic power of linear rank tests is obtained which amongst other things permits the derivation of optimal PLRT's under weaker conditions than are required for optimal linear rank tests. Further useful properties of PLRT's are noted.

1. Introduction. One of the most interesting possibilities for linear rank tests in the two-sample location-shift problem is that they can be guaranteed to have efficiency one asymptotically against any underlying distribution (subject to certain conditions) by choosing the score generating function defining the test statistic adaptively. The first class of tests constructed in this way is contained in Hájek and Šidák (1967), following Hájek (1962), but these are complicated and apparently the adaptive estimator constructed there converges very slowly even though it is consistent. Beran (1974) proposes a quite different adaptive estimator of the score generating function. This is a much more elegant and useable estimator; no rate of convergence result is obtained however. Since these estimators depend upon the Fourier expansion of the score generating function, their convergence properties are influenced by the slow convergence of Fourier series to functions which are not continuous (when regarded as periodic).

The present paper obtains a new adaptive estimator of the score generating function using the basic ideas of Beran's paper. This estimator is not only mean-square consistent, but a rate of convergence result may also be obtained. The resulting adaptive test belongs to the class of rank tests generated by continuous, piecewise linear functions over [0, 1] and for convenience these are called PLRT's. There are several advantages to restricting attention to tests within this class rather than using the greater generality of tests obtained from $L_2([0, 1])$. These are outlined in Section 3.

2. Optimal and adaptive PLRT's. Suppose $0 = \lambda_0 < \lambda_1 < \cdots < \lambda_r = 1$ are fixed real numbers. The score generating function of a PLRT is defined by

$$\phi(0) = 0, \qquad \phi(\lambda_i) = \alpha_i(\lambda_i - \lambda_{i-1}) + \phi(\lambda_{i-1}), \qquad 1 \le i \le r, \tag{2.1}$$

and is linear over $[\lambda_{i-1}, \lambda_i]$. We require to find the asymptotically most powerful PLRT against location shift for two samples from a distribution F with absolutely continuous

Received February 1979; revised November 1981.
AMS 1970 *subject classifications.* Primary 62G10, 62G25.
Key words and phrases. Linear rank statistic, score generating function, continuous piecewise linear function, asymptotically optimal tests, adaptive tests, consistency and rate of convergence, Riesz representation.

density f and such that $\phi_F(u) = f'(F^{-1}(u))/f(F^{-1}(u)) \in L_2([0, 1])$. This is obtained by projecting ϕ_F onto $PL(\lambda)$, the functions defined by (2.1). The following discussions are simplified by taking $\lambda_i = i/r$.

Let $\bar{r} = r + 1$ and define $\bar{A} = (\bar{a}_{ij})$ as the $\bar{r} \times \bar{r}$ symmetric tridiagonal matrix with entries given by

$$\bar{a}_{11} = (3r)^{-1} = \bar{a}_{\bar{r}\bar{r}}, \quad \bar{a}_{ii} = 2(3r)^{-1}, \quad 2 \le i \le r, \tag{2.2}$$
$$\bar{a}_{i(i+1)} = (6r)^{-1}, \quad 1 \le i \le r, \quad \bar{a}_{ij} = 0, \quad j > i + 1.$$

Define $e^T = (e_1, \cdots, e_{\bar{r}})$ by $e_i = e_{\bar{r}} = (2r)^{-1}$, $e_i = r^{-1}$, $2 \le i \le r$. Let $A = \bar{A} - ee^T$. Define

$$Q(1) = 1, \tag{2.3}$$
$$Q(i) = (6r)^{-i+1}(K_i - K_{i+1}), \quad 2 \le i \le r + 2$$

where

$$K_1 = 0, K_i = [(2 + \sqrt{3})^i - (2 - \sqrt{3})^i]/(2\sqrt{3}), \quad 2 \le i \le r + 2.$$

It is easy to verify that $Q(i)$ is the determinant of the submatrix of $\bar{A}$ obtained by taking rows $1, \cdots, i - i$ together with the corresponding columns. Furthermore

$$|\bar{A}| = (6r)^{-r-1}(K_{r+2} - 2K_{r+1} + K_r).$$

Then by applying the formula given on page 114 of Karlin (1968), one can obtain $\bar{B} = (\bar{b}_{ij}) = \bar{A}^{-1}$ as

$$\bar{b}_{ij} = (-1)^{i+j}(6r)^{i-j}Q(i)Q(r + 2 - j)|\bar{A}|^{-1}, \quad i \le j, \tag{2.4}$$

the rest of the entries being obtained by symmetry. Also $B = A^{-1} = \bar{B} - (\bar{B}e)(\bar{B}e)^T/(1 + e^T\bar{B}e)$.

Let $\mathbf{c}$ denote the $\bar{r} \times 1$ vector with entries

$$c_1 = -\int_0^{r^{-1}} f(F^{-1}(u))\,du, \quad c_{\bar{r}} = -\int_{1-r^{-1}}^1 f(F^{-1}(u))\,du \tag{2.5}$$
$$c_i = -\int_{(i-2)r^{-1}}^{(i-1)r^{-1}} f(F^{-1}(u))\,du - \int_{(i-1)r^{-1}}^{ir^{-1}} f(F^{-1}(u))\,du, \quad 2 \le i \le r.$$

Define $\mathbf{d} = B\mathbf{c}$.

LEMMA 2.1. *The asymptotically most powerful* PLRT *against location shift is obtained by taking*

$$\alpha_i = (d_{i+1} - d_i), \quad 1 \le i \le r, \tag{2.6}$$

in (2.1). *The* ARE *of this test relative to the asymptotically most powerful rank test equals* $I_F^{-1}I^{-1}(c^TBc)^2$ *where* $I_F = \int_0^1 \phi_F(u)^2\,du$ *and* $I = c^TBc$.

PROOF. The standard techniques for obtaining a projection are used. For $1 \le i \le \bar{r}$ define 'tent' functions

$$\tilde{\beta}_1(x) = \begin{cases} 1 - rx & 0 \le x \le r^{-1}, \\ 0 & \text{elsewhere;} \end{cases}$$
$$\tilde{\beta}_{\bar{r}}(x) = \begin{cases} rx - r + 1 & 1 - r^{-1} \le x \le 1, \\ 0 & \text{elsewhere} \end{cases}$$

and for $2 \le i \le r$,

$$\tilde{\beta}_i(x) = \begin{cases} (rx - i + 2) & i - 2 \le rx \le i - 1, \\ (i - rx) & i - 1 \le rx \le i, \\ 0 & \text{elsewhere.} \end{cases}$$

Hence, take $\beta_i = \tilde{\beta}_i - e_i$ so that $\int_0^1 \beta_i(x)\,dx = 0$. Then the asymptotically most powerful PLRT is generated by $f_1\beta_1 + \cdots + f_{\bar{r}}\beta_{\bar{r}}$ where $f_1, \cdots, f_{\bar{r}}$ minimize $\| \phi_F - f_1\beta_1 - \cdots - f_{\bar{r}}\beta_{\bar{r}} \|$ and standard arguments (see Luenberger, 1969) imply $f_i = d_i$, $1 \le i \le \bar{r}$, giving (2.6) and the ARE.

Lemma 2.1 is of interest in its own right and is discussed further in Section 3. Here the interest lies in using the combined sample $X_1, \cdots, X_N$ of independent random variables with distribution F to furnish a mean-square consistent estimator of ϕ_F. Write

$$T(a, b) = \int_a^b f(F^{-1}(u))\,du$$

and estimate this by

$$\hat{T}_N(a, b) = (2N\theta_N)^{-1} \textstyle\sum_{i=1}^N \{\psi((N-1)^{-1} \sum_{j\neq i} \delta(X_i - X_j + \theta_N)) - \psi((N-1)^{-1} \sum_{j\neq i} \delta(X_i - X_j - \theta_N))\} \tag{2.7}$$

where

$$\psi(u) = uI_{[a,b]}(u), \qquad \delta(u) = I_{[0,\infty)}(u)$$

($I_{[a,b]}$ denotes the indicator function for $[a, b]$) and $\theta_N = N^{-1/2}\theta$ for fixed $\theta > 0$. This is the estimator proposed by Beran (1974) and from this an estimator $\hat{\phi}_F^{(r,N)}$ is obtained using (2.5), (2.6) and (2.1) for $\lambda_i = ir^{-1}$. It remains to study the rate of convergence of $\| \hat{\phi}_F^{(r,N)} - \phi_F \|$ as $N \to \infty$ and this is made possible by the following lemma which marks an important difference between $\hat{\phi}_F^{(r,N)}$ and Beran's Fourier-based estimators.

LEMMA 2.2 *Let* $\phi^{(r)}$ *denote the minimum norm approximation to* ϕ_F *given by Lemma* 2.1. *Assume* ϕ'_F *exists and is continuous on* $[0, 1]$. *Then*

$$\| \phi^{(r)} - \phi_F \|^2 = o(r^{-1}). \tag{2.8}$$

PROOF. Use the mean-value theorem applied to the polygonal approximation of ϕ_F.

THEOREM 2.3 *Assume* (i) ϕ'_F *exists and is continuous on* $[0, 1]$, (ii) $r(N)/N \to \infty$ *as* $N \to \infty$. *Then*

$$\lim_{N\to\infty} E\{N\| \hat{\phi}^{(r(N),N)} - \phi_F \|^2\} = \tfrac{1}{2}\{\textstyle\int_0^1 u(1-u)\phi'_F(u)^2\,du - \phi_F(0)^2\}. \tag{2.9}$$

PROOF. The ideas used in this proof are rather similar to the ones used by Beran (1974) although the form of $\hat{\phi}^{(r(N),N)}$ permits some strengthening of the results for the rate of convergence. These details are omitted for lack of space and only the calculation in (2.9) is described in detail.

By the definition of $\phi^{(r)}$, $\phi_F - \phi^{(r)} \perp PL(r^{-1}, 2r^{-1}, \cdots, 1)$ and so

$$\| \hat{\phi}^{(r,N)} - \phi_F \|^2 = \| \hat{\phi}^{(r,N)} - \phi^{(r)} \|^2 + \| \phi^{(r)} - \phi_F \|^2. \tag{2.10}$$

From assumption (ii) and (2.8) it therefore remains to show that

$$\lim_{N\to\infty} NE\{\| \hat{\phi}^{(r,N)} - \phi^{(r)} \|^2\}$$

is given by (2.9). By Lemma 2.1, with $\hat{\mathbf{c}}^{(r)} = (\hat{c}_1^{(r)}, \cdots, \hat{c}_{\bar{r}}^{(r)})$ denoting the estimator of $\mathbf{c}$ obtained using (2.7), one obtains

$$\begin{aligned} E\{\| \hat{\phi}^{(r,N)} - \phi^{(r)} \|^2\} &= E\{(\hat{c}^{(r)} - c)^T B(\hat{c}^{(r)} - c)\} \\ &= \int_0^1 \int_0^1 \mathscr{C}(1 + [\bar{r}x], 1 + [\bar{r}y])\, d_{x,y}\mathscr{B}(1 + [\bar{r}x], 1 + [\bar{r}y]) \end{aligned} \tag{2.11}$$

where for $1 \le i, j \le \bar{r}$,

$$\mathscr{B}(i, j) = r^{-2} \textstyle\sum_{k=1}^i \sum_{\ell=1}^j b_{k\ell}, \quad \mathscr{C}(i, j) = r^2 E\{(\hat{c}_i - c_i)(\hat{c}_j - c_j)\}.$$

In order to prove (2.9) we shall show that

(2.12) $\lim_{r\to\infty}\mathscr{B}(1+[\bar{r}x], 1+[\bar{r}y]) = (x\wedge y - xy/3)/2 (x\wedge y = \min(x,y))$, and

(2.13) $$\lim_{N\to\infty} N\mathscr{C}(1+[\bar{r}x], 1+[\bar{r}y]) = \xi_{12}^{-}(x,y),$$

where ξ^{-} is the left-derivative of

$$\xi(x,y) = [x\wedge y - xy]\phi_F(x)\phi_F(y),$$

and

(2.14) $$N\mathscr{C}(1+[\bar{r}x], 1+[\bar{r}y]) \le K,$$

where K does not depend upon x, y or N. Although not explicitly indicated here, remember that r is a function of N. Assuming (2.12) – (2.14), the extended Helly-Bray theorem applied to (2.11) gives

$$\lim_{N\to\infty} NE\{\|\hat{\phi}^{(r,N)} - \phi^{(r)}\|^2\} = \frac{1}{2}\int_0^1\int_0^1 \xi_{12}^{-}(x,y)\, d[x\wedge y - xy/3]$$
$$= \frac{1}{2}\int_0^1\int_0^1 \xi_{12}^{-}(x,y)\, d[x\wedge y] = \frac{1}{2}\int_0^1 \xi_{12}^{-}(x,x)\, dx$$

as required.

The proofs of (2.13), (2.14) are based on Beran's paper and only (2.12) will be considered here. Define $\beta(i,j) = \sum_{k=1}^{i} \bar{b}_{kj}$ and observe that from (2.4)

$$\beta(1+[\bar{r}x], 1+[\bar{r}y]) \sim r/2\, I_{[0,x]}(y)$$

since

$$\sum_{k=1}^{1+[\bar{r}x]} (-1)^{k+1+[\bar{r}y]} Q(k)Q(\bar{r}-[\bar{r}y]) \sim (-1)^{[\bar{r}x]+[\bar{r}y]}(6r)^{-r}(2+\sqrt{3})^{\bar{r}+[\bar{r}x]-[\bar{r}y]}/(12\sqrt{3})(x\le y)$$
$$\sim (6r)^{-r}(2+\sqrt{3})^{r+1}/(12\sqrt{3}) + o[2^r(6r)^{-r-1}](x>y)$$

and

$$|A| \sim (6r)^{-r-1}(2+\sqrt{3})^{r+1}/\sqrt{3}.$$

It follows by dominated convergence that

$$r^2\bar{\mathscr{B}}(1+[\bar{r}x], 1+[\bar{r}y]) = \sum_{k=1}^{1+[\bar{r}x]}\sum_{\ell=1}^{1+[\bar{r}y]} \bar{b}_{k\ell}$$
$$= \bar{r}\int_0^y \beta(1+[\bar{r}x], 1+[\bar{r}u])\, du + o(1) \sim r^2(x\wedge y)/2.$$

Since for $1\le k, \ell\le \bar{r}$,

$$b_{k\ell} = \bar{b}_{k\ell} - (r^2)^{-1}\left(\textstyle\sum_{j=1}^{\bar{r}} \bar{b}_{kj}\right)\left(\textstyle\sum_{j=1}^{\bar{r}} \bar{b}_{\ell j}\right)/(1+(r^2)^{-1}\bar{\mathscr{B}}(\bar{r},\bar{r})),$$

(2.12) now follows.

REMARKS (i) Mean square consistency of $\hat{\phi}^{(r(N),N)}$ only requires $r, N\to\infty$.

(ii) The rate at which $r(N)\to\infty$ is only restricted by (2.8). Existence of higher derivatives permits progressively slower rates of growth. This is of some interest since r determines how complicated the estimator/test statistic is going to be (see (2.1)).

(iii) Results about the efficiency of different adaptive estimators may be obtained from Theorem 2.3. In terms of Pitman efficiency of a sequence of tests, Theorem 2.3 translates to (efficiency of adaptive test) $= 1 - O(N^{-1/2})$. Different adaptive tests (for the same F) may be compared by taking the ratio of the term in (2.9).

3. The case for PLRT's The theme of this paper is that PLRT's provide an extremely useful class of tests for location-shift (more generally regression) and also provide a natural class of R-estimators. In this section the points in their favour are summarized; some of these observations are relatively obvious but seem helpful in a practical sense.

(a) PLRT's used as adaptive tests mean that the theory of fixed score generating functions may be abandoned with some confidence in view of the $O(N^{-1/2})$ convergence result of Theorem 2.3. Adaptive testing is therefore a practical proposition.

(b) The exact null distribution of a PLRT may be calculated. This distribution may be obtained in terms of the Mann-Whitney probabilities and computed for reasonably small $r(r = 4$, for instance).

(c) Even if fixed (non-random) score generating functions are used in preference to an adaptive estimate, little is lost by taking the optimal PLRT instead of ϕ_F.

EXAMPLE. Assume F is normal, $r = 4$, $\lambda_i = ir^{-1}$. The asymptotically most powerful PLRT is defined by $\alpha_1 = \alpha_4 = 1.43$, $\alpha_3 = \alpha_2 = .58$, and the ARE (against normal scores test) is .98; when $r = 8$, this ARE is .994. Throughout the discussion it has been assumed that $\lambda_1, \cdots, \lambda_r$ are fixed. In approximating ϕ_F by (2.1) one may try to choose an optimal segmentation of [0, 1]; the dynamic programming techniques of Bellman and Roth (1969) or Hawkins (1972) provide a method for obtaining near optimal segmentations.

(d) PLRT's possess a certain amount of robustness; details omitted—see Eplett (1980) for the basic ideas.

(e) An asymptotically most powerful PLRT may exist when there is no asymptotically most powerful rank test. This observation is obtained as a corollary to Theorem 3.1 which provides a new insight into theorems of Hájek (1962) and van Eeden (1963). Assume $S_{m,n}$ is a simple linear rank statistic based on the samples $X_1, \cdots, X_m, Y_1, \cdots, Y_n$ and generated by $\phi(u) \in L_2([0, 1])$, all definitions as in Hájek and Šidák (1967). Write $T_{m,n} = m^{-1}S_{m,n}$ and assume that this is used to test $H: \theta = \theta_0$ against the sequence of alternatives that $\theta_{m,n} = \theta_0 + kN^{-1/2}$, k fixed, $N = m + n$. Assume $m/N \to \gamma$, $0 < \gamma < 1$. Under H the ranks of the combined sample are randomly distributed.

THEOREM. 3.1. *Suppose $\mathscr{S} \subset L_2([0, 1])$ is restricted so that $\phi \in \mathscr{S}$ implies $\int_0^1 \phi(u)\, du = 0$ (no generality is lost by this) and that conditions A-E of van Eeden (1963) (the standard assumptions used for computing ARE) hold for $\{T_{m,n}(\phi)\}$. In addition assume that*

(i) *if $\mu_{m,n}(\theta_{m,n}) = E(T_{m,n}(\phi) \mid \theta_{m,n})$, then $\mu_{m,n}(\theta_{m,n}) \to g(\phi)$ where $\{g(\phi): \phi \in \mathscr{S}, \|\phi\| = 1\}$ is bounded;*

(ii) *there exists at least one $\phi \in \mathscr{S}$ for which $g(\phi) > 0$. Then there exists a unique $\xi \in \mathscr{S}$ for which the ARE of the sequence $\{T_{m,n}(\phi)\}$ relative to $\{T_{m,n}(\xi)\}$ equals ρ^2 where*

$$\rho = <\phi, \xi>/\{\|\phi\| \|\xi\|\}$$

provided $g(\phi) > 0$.

PROOF. Let $\sigma^2_{m,n} = \operatorname{Var}(T_{m,n} \mid \theta_0)$, then $N\sigma^2_{m,n} \to \gamma^{-1}(1 - \gamma)\|\phi\|^2$ as $N \to \infty$. Let $e(\phi)$ denote the efficacy of the sequence $\{T_{m,n}(\phi)\}$, that is

$$e(\phi) = \lim_{N\to\infty} N^{-1/2}\mu'_{m,n}(\theta_{m,n})/\sigma_{m,n}.$$

It follows that $e(\phi) = \{\gamma/(1 - \gamma)\}^{1/2} g(\phi)/\|\phi\|$. Now g is in fact a (bounded) linear functional on $\mathscr{S}$. For consider $g(\phi_1 + \phi_2)$. The value of $g(\phi)$ is independent of the particular sequence $\{T_{m,n}(\phi)\}$ used in its definition and so given $\phi(u) \in \mathscr{S}$, the scores

$$a_N(i) = N \int_{(i-1)N^{-1}}^{iN^{-1}} \phi(u)\, du$$

can be used to define the test statistics $T_{m,n}(\phi)$. Clearly with this choice of scores $T_{m,n}(\phi_1 + \phi_2) = T_{m,n}(\phi_1) + T_{m,n}(\phi_2)$ and so $E(T_{m,n}(\phi_1 + \phi_2)) = E(T_{m,n}(\phi_1)) + E(T_{m,n}(\phi_2))$. Differentiating and taking limits yields $g(\phi_1 + \phi_2) = g(\phi_1) + g(\phi_2)$. Similarly $g(c\phi) = cg(\phi)$ for real c.

The Riesz representation theorem can now be applied to $h(\phi) = \{\gamma/(1-\gamma)\}^{1/2} g(\phi)$ to obtain $h(\phi) = <\phi, \xi>$ for some unique $\xi \in \mathscr{S}$. The theorem follows upon squaring the ratio of the efficacies to obtain the ARE. It follows that ξ generates the asymptotically most powerful test within $\mathscr{S}$.

In order to apply Theorem 3.1 to PLRT's, assume $f(F^{-1}(u)) \in L_2([0, 1])$ but $\phi_F \notin L_2([0, 1])$ so that the approach of Lemma 2.1 does not suffice. Apply the theorem, noting that of van Eeden's conditions only asymptotic normality requires verification. This may be obtained from Theorem 2.4 of Hájek (1968) since if $\phi \in PL(\lambda)$, then $\phi = \phi_1 - \phi_2$ where ϕ_1, ϕ_2 are non-decreasing absolutely continuous functions over (0, 1) obtained by taking

$$\phi_1(\lambda_i) = \textstyle\sum_{k=1}^{i} \delta(\alpha_k)\alpha_k(\lambda_k - \lambda_{k-1}), \quad 1 \leq i \leq r$$

and ϕ_1 linear over $[\lambda_{i-1}, \lambda_i]$, $1 \leq i \leq r$. Assumption (i) of Theorem 3.1 is satisfied with

$$g(\phi) = C \int_0^1 \phi'(u) f(F^{-1}(u))\, du$$

(C is a constant determined by ϕ and γ). Theorem 3.1 now implies that an asymptotically most powerful PLRT exists and there is a Riesz representation for the asymptotic power of PLRT's against this sequence of location-shift alternatives. The score generating function of this optimal test is given by (2.5), (2.6) appropriately generalized to arbitrary $\lambda_1, \cdots, \lambda_r$ if necessary.

Finally it may be worth noting that the results of Gastwirth (1965, 1966) also follow from Theorem 3.1. In these cases one takes $\mathscr{S}$ as given by step-functions which jump at fixed $\lambda_1, \cdots, \lambda_r$ or $\mathscr{S}$ as the subspace of $L_2([0, 1])$ functions which are constant over (p, 1] for some fixed $p \in (0, 1)$. The conditions required by Gastwirth may of course be weakened in the way just demonstrated for PLRT's.

REFERENCES

Bellman, R. E. and Roth, R. (1969). Curve-fitting by segmented straight lines. *J. Amer. Statist. Assoc.* **64** 1079–1084.

Beran, R. (1974). Asymptotically efficient adaptive rank estimates in location models. *Ann. Statist.* **2** 63–74.

Eplett, W. J. R. (1980). An influence curve for two-sample rank tests. *J. Roy. Stat. Soc. B* **42** 64–70.

Gastwirth, J. L. (1965). Asymptotically most powerful rank tests for the two-sample problem with censored data. *Ann. Math. Statist.* **36** 1243–1247.

Gastwirth, J. L. (1966). On robust procedures. *J. Amer. Statist. Assoc.* **61** 929–948.

Hájek, J. (1962). Asymptotically most powerful rank-order tests. *Ann. Math. Statist.* **33** 1124–1147.

Hájek, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. *Ann. Math. Statist.* **39** 325–346.

Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests.* Academic, New York.

Hawkins, D. M. (1972). On the choice of segments in piecewise approximation. *J. Inst. Math. Applic.* **9** 250–256.

Karlin, S. (1968). *Total Positivity.* Stanford University Press, California.

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

Van Eeden, C. (1963). The relation between Pitman's asymptotic relative efficiency of two tests and the correlation coefficient between their test statistics. *Ann. Math. Statist.* **34** 1442–1457.

Mathematical Institute
University of Oxford
24–29 St. Giles
Oxford OX1 3LB
England