**Poisson Approximation for Some Statistics Based on Exchangeable Trials**

A. D. Barbour; G. K. Eagleson

*Advances in Applied Probability*, Vol. 15, No. 3. (Sep., 1983), pp. 585-600.

Stable URL:
http://links.jstor.org/sici?sici=0001-8678%28198309%2915%3A3%3C585%3APAFSSB%3E2.0.CO%3B2-5

*Advances in Applied Probability* is currently published by Applied Probability Trust.

---

*Adv. Appl. Prob.* **15**, 585–600 (1983)
*Printed in N. Ireland*


# POISSON APPROXIMATION FOR SOME STATISTICS BASED ON EXCHANGEABLE TRIALS

A. D. BARBOUR,* *Gonville and Caius College, Cambridge*
G. K. EAGLESON,** *CSIRO Division of Mathematics and Statistics*

**Abstract**

Stein's (1970) method of proving limit theorems for sums of dependent random variables is used to derive Poisson approximations for a class of statistics, constructed from finitely exchangeable random variables.

Let $\{Y_j\}_{j=1}^M$ be exchangeable random elements of a space $\mathscr{H}$ and, for $I$ a $k$-subset of $\mathscr{X}$, let $X_I$ be a 0–1 function. The statistics studied here are of the form

$$W = \sum_{I \in \mathscr{N}} X_I,$$

where $\mathscr{N}$ is some collection of $k$-subsets of $\mathscr{X}$.

An estimate of the total variation distance between the distributions of $W$ and an appropriate Poisson random variable is derived and is used to give conditions sufficient for $W$ to be asymptotically Poisson. Two applications of these results are presented.

FINITELY EXCHANGEABLE RANDOM VARIABLES

## 1. Introduction

Let $\{Y_j\}_{j=1}^M$ be exchangeable random elements of a space $\mathscr{X}$, and let $\phi$ be a *symmetric* mapping from $\mathscr{X}^k$ to $\{0, 1\}$. Let $D(k; r, s)$ denote the collection of $k$-subsets of $\{r, r+1, \cdots, s\}$ and, for $I \in D(k; r, s)$, let $Y_I = \{Y_i : i \in I\}$ and $X_I = \phi(Y_{\boldsymbol{i}})$, where $\boldsymbol{i}$ is any ordering of $I$: $X_I$ is well defined because of the symmetry of $\phi$.

Many statistics can be expressed in the form

$$W = \sum_{I \in \mathscr{N}} X_I,$$

where $\mathscr{N}$ is some subset of $D \equiv D(k; 1, M)$, and we shall be concerned with studying their distribution.

Received 8 June 1982; revision received 5 January 1983.

* Present address: Institut für Angewandte Mathematik, Universität Zürich, Rämistrasse 74, CH-8001 Zürich, Switzerland.

** Postal address: CSIRO Division of Mathematics and Statistics, Bradfield Rd, West Lindfield, NSW 2070, Australia.

When the underlying $\{Y_j\}$ are independent and identically distributed and $\mathcal{N} = D$, $W$ is asymptotically distributed (as $M \to \infty$) under rather natural conditions as a Poisson random variable (Silverman and Brown (1978)). In this case, $W$ has been used to construct tests of randomness and of collinearities of points in the plane. Eagleson (1979) assumed only exchangeability, and took $\mathcal{N} = D(k; 1, r(M)) \subseteq D$; he showed that $W$ is asymptotically distributed as a Poisson random variable (as $M \to \infty$) provided that some moment conditions are satisfied and that

$$\lim_{M \to \infty} r(M)/M = 0.$$

It seems natural to expect that similar results hold when $\mathcal{N}$ is chosen to be some other relatively small subset of $D$. Such an extension is needed to provide a proof of the asymptotic distribution of Knox's statistic, used in epidemiology, where the subset $\mathcal{N}$ is chosen as those outbreaks of a disease which occur close in time (see Section 3 for more details). Unfortunately, the martingale techniques used in Eagleson (1979) do not readily give the necessary generalization.

An alternative approach to proving limit theorems for sums of dependent random variables was originated by Stein (1970) and developed in the Poisson context by Chen (1975). The basic idea is that, given a set $A \subset Z^+$, one defines a bounded function $x = x(\lambda, A): Z^+ \to R$ by

$$x(0) = 0;$$

$$x(m+1) = \lambda^{-m-1} e^{\lambda} m! \, [\mathcal{P}_\lambda(A \cap U_m) - \mathcal{P}_\lambda(A)\mathcal{P}_\lambda(U_m)], \qquad m \geqq 0,$$

where $U_m = \{0, 1, \cdots, m\}$ and where $\mathcal{P}_\lambda(B)$ denotes the probability that a Poisson random variable with mean $\lambda$ belongs to $B$. The importance of such an $x$ is that, for any non-negative integer-valued random variable $T$, it can be shown that

$$E\{\lambda x(T+1) - Tx(T)\} = P(T \in A) - \mathcal{P}_\lambda(A). \tag{1}$$

Chen used this identity to derive an estimate of the total variation distance between $T$ and $\mathcal{P}_\lambda$ when $T$ is a sum of dependent Bernoulli random variables, satisfying a mixing condition. In the present paper, the symmetry properties of the joint distribution of $\{X_I; I \in D\}$ are used to calculate some conditional expectations explicitly and thus to approximate the left-hand side of (1).

It cannot be denied that the following analysis is far from simple, but the wide applicability of the results obtained and their potential importance justifies the detailed calculations. The main results are stated in Section 2, and their proofs given in Section 3. In Section 4 two applications are described; one a justification of the use of Knox's statistic and the other a study of the asymptotic distribution of the number of 'rare' patterns occurring in a random arrangement of two types of points.

## 2. Statement of results

The principal results are expressed in terms of convergence theorems for a sequence of statistics $W_n$, $n \geqq 1$, each of which is derived in the manner described in the introduction. Thus, for example, the underlying probabilistic framework is that of a triangular array $(\{Y_{jn}\}_{j=1}^{M(n)}; n \geqq 1)$ of random variables exchangeable within rows. The value of $k$ is kept the same for all $n$, whereas all other quantities are allowed to vary: however, to simplify notation, the explicit dependence of the other quantities upon $n$ is suppressed.

With notation as before, let

$$p = EX_I, \qquad N = |\mathcal{N}| \quad \text{and} \quad \mu = EW = Np.$$

Then, for $0 \leqq j \leqq k$, let $C_j$ be the number of pairs of elements $(I, J)$ of $\mathcal{N}$ such that $|I \cap J| = j$: further, for such pairs $(I, J)$, set $q_j = EX_IX_J$ and $\eta_j = |q_j - p^2|^{\frac{1}{2}}$. Finally, for any subset $K \subseteq \{1, 2, \cdots, M\}$, let $\mathrm{In}(K)$ consist of those elements of $\mathcal{N}$ which have non-empty intersection with $K$, set $i_K = |\mathrm{In}(K)|$, and let $\mathrm{No}(K) = \mathcal{N} \setminus \mathrm{In}(K)$.

*Theorem* 1. The distribution of $W$ converges (as $n \to \infty$) in distribution to $\mathcal{P}_\lambda$, the Poisson distribution with parameter $\lambda$, if the following six conditions are satisfied:

(i) $Np \to \lambda$ (ii) $N^{-1} \max_{J \in \mathcal{N}} i_J \to 0$

(iii) $NM^{-k} \to 0$ (iv) $C_jq_j \to 0, \quad 1 \leqq j \leqq k-1$

(v) $N^2q_jM^{-j} \to 0, \quad 1 \leqq j \leqq k-1$ (vi) $N\eta_0 \to 0.$

The conditions of Theorem 1 simplify considerably in various special cases. For example, if $\mathcal{N} = D$ and each row of the $Y$'s is the same independent and identically distributed sequence, they can be reduced to those of Silverman and Brown (1978), Theorem A. In fact, Corollary 4 below exhibits an improvement in this case on the rate of convergence to the Poisson limit given in Brown and Silverman (1979). When $\mathcal{N} = D(k; 1, r(M))$, the theorem reduces to that of Eagleson (1979), and when $\mathcal{N}$ is a balanced subset and the $Y$'s are independent and identically distributed a theorem of Berman and Eagleson (1983) is recovered.

The following further cases are useful in applications.

*Corollary* 1 (subdiagonal). Let $M = n$ and $\mathcal{N} = \{\{r, r+1, \cdots, r+k-1\}\}_{r=1}^{M-k+1}$. Then $W \to^D \mathcal{P}_\lambda$ if, as $M \to \infty$, $Mp \to \lambda$, $Mq_j \to 0$, $1 \leqq j \leqq k-1$, and $M\eta_0 \to 0$.

*Proof.* Observe that here $\max_{J \in \mathcal{N}} i_J \leqq 2k-1$ and that, for $1 \leqq j \leqq k-1$, $C_j \sim 2M$. Now apply Theorem 1.

*Corollary* 2 (subdiagonal blocks). Let $M = n$, and suppose that $1 = r_0 < r_1 < \cdots < r_t \leqq M - k + 1$,

$$m = \max_{1 \leqq j \leqq t} (r_j - r_{j-1}) + k$$

$$\mathcal{N} = \bigcup_{j=1}^{t} D(k; r_{j-1}, r_j + k - 1).$$

Then $W \to^D \mathcal{P}_\lambda$ (as $M \to \infty$) if $N \to \infty$, $Np \to \lambda$, $m/M \to 0$, $m^{2k-j-1} q_j (M \wedge mt) \to 0$, $1 \leqq j \leqq k-1$, and $N\eta_0 \to 0$.

*Proof.* First, observe that

$$N \leqq (k!)^{-1} \sum_{j=1}^{t} (r_j - r_{j-1} + k)^k,$$

so that

$$N \leqq (k!)^{-1} m^{k-1} (M + kt) \leqq c_1 M m^{k-1}, \quad \text{and} \quad N \leqq (k!)^{-1} t m^k;$$

and also that $N \geqq c_2 m^k$. Then, counting coincidences,

$$C_j \leqq 2N \binom{k}{j} m^{k-j}/(k-j)! \leqq c_3 m^{2k-j-1} (M \wedge tm), \qquad 1 \leqq j \leqq k-1,$$

and, for each $J \in \mathcal{N}$, $i_J \leqq c_4 m^{k-1}$. The conditions of Theorem 1 can now be verified: note that, for (ii),

$$N^{-1} \max_{J \in \mathcal{N}} i_J \leqq c_5 (m^{-1} \wedge N^{-1} m^{k-1}) \to 0.$$

When $\lambda = Np \to \infty$, similar results can be obtained, provided $\lambda$ does not grow too quickly. For example, the following variant of Theorem 1 can be established.

*Theorem* 2. The total variation distance between the distribution of $W$ and $\mathcal{P}_\lambda$ converges to 0 if, as $n \to \infty$,

(i) $\lambda = Np \to \infty$ (ii) $p \max_{J \in \mathcal{N}} i_J \to 0$

(iii) $N\lambda^k M^{-k} \to 0$ (iv) $\lambda^{-1} q_j C_j \to 0$, $1 \leqq j \leqq k-1$

(v) $N^2 \lambda^{j-1} \eta_j^2 M^{-j} \to 0$, $1 \leqq j \leqq k-1$ (vi) $\lambda^{-\frac{1}{2}} N\eta_0 \to 0$.

*Corollary* 3 (overlapping subdiagonal blocks, $k = 2$). Let $M = n$, choose $0 \leqq b_i \leqq M - i$, $1 \leqq i \leqq M-1$, with $b_{i+1} \geqq b_i - 1$, $1 \leqq i < M-1$, and define

$$\mathcal{N} = \{(i, j) : i < j \leqq i + b_i, 1 \leqq i \leqq M-1\}.$$

Then

(i) the distribution of $W$ converges to $\mathcal{P}_\lambda$ if, as $M \to \infty$, $N \to \infty$, $Np \to \lambda$, $m = \max_i b_i = o(M)$, $q_1\{C_1 \vee N^2/M\} \to 0$ and $N\eta_0 \to 0$;

(ii) the total variation distance between the distribution of $W$ and $\mathcal{P}_\lambda$

converges to 0 if, as $M \to \infty$, $\lambda = Np \to \infty$, $mp \to 0$, $mM^{-1}\lambda^2 \to 0$, $\lambda^{-1} q_1 \{C_1 \vee \lambda N^2/M\} \to 0$ and $\lambda^{-\frac{1}{2}} N\eta_0 \to 0$.

*Proof.* Easily, $N \leqq Mm$, $N > \frac{1}{2}m^2$ and, for $J \in \mathcal{N}$, $i_J \leqq 4m$. Hence, under the conditions of Part (i) of the corollary, by an argument similar to that in the proof of Corollary 2, Conditions (ii) and (iii) of Theorem 1 are satisfied while Conditions (i), (iv), (v) and (vi) are immediate.

Part (ii) of the corollary is proved in similar fashion, by appealing to Theorem 2.

## 3. Proofs

The proofs are based on the following three lemmas. Let $S \in D(s; 1, M)$, and, for $J \in \mathcal{N}$, define

$$W_{S,J} = \sum_{I \in \mathrm{No}(S \cup J)} X_I.$$

Then, for each bounded $x : Z^+ \to R$, let

$$\|x\| = \sup_{m \in Z^+} |x(m)|,$$

and let

$$\Delta x = \sup_{m \in Z^+} |x(m+1) - x(m)|.$$

*Lemma* 1. For each bounded $x : Z^+ \to R$ and each $S \in D(s; 1, M)$,

$$\left| E\{\mu x(W+1) - Wx(W)\} - \sum_{I \in \mathrm{No}(S)} E\{(p - X_I) x(W_{S,I} + 1)\} \right|$$
$$\leqq (\Delta x \wedge 2\|x\|) \left\{ p^2 \left( Ni_S + \sum_{I \in \mathcal{N}} i_I \right) + \sum_{j=1}^{k-1} q_j C_j + Nq_0 i_S \right\} + 2p\|x\| i_S.$$

*Proof.* Consider the expression

$$E\{\mu x(W+1) - Wx(W)\} - \sum_{I \in \mathrm{No}(S)} E\{(p - X_I) x(W_{S,I} + 1)\}$$
$$= \sum_{I \in \mathrm{No}(S)} E\{p[x(W+1) - x(W_{S,I}+1)] + X_I[x(W_{S,I}+1) - x(W)]\}$$
$$+ \sum_{I \in \mathrm{In}(S)} E\{px(W+1) - X_I x(W)\}.$$

For the first term on the right-hand side, one has the estimates

$$|x(W+1) - x(W_{S,I}+1)| \leqq 2\|x\| I[x(W+1) - x(W_{S,I}+1) \neq 0] \leqq 2\|x\| \left\{ \sum_{J \in \mathrm{In}(S \cup I)} X_J \right\}$$

and

$$|x(W+1)-x(W_{S,I}+1)| \leqq \Delta x\Big\{\sum_{J\in \mathrm{In}(S\cup I)} X_J\Big\};$$

hence

$$\Big|\sum_{I\in \mathrm{No}(S)} pE\{x(W+1)-x(W_{S,I}+1)\}\Big| \leqq (2\|x\|\wedge\Delta x)\sum_{I\in\mathrm{No}(S)} i_{S\cup I}p^2$$
$$\leqq (2\|x\|\wedge\Delta x)p^2\Big\{Ni_S+\sum_{I\in\mathcal{N}} i_I\Big\}.$$

For the next term, in a similar way,

$$|X_I\{x(W_{S,I}+1)-x(W)\}| \leqq X_I(2\|x\|\wedge\Delta x)\sum_{J\in\mathrm{In}(S\cup I)\setminus I} X_J$$
$$\leqq (2\|x\|\wedge\Delta x)X_I\Big\{\sum_{J\in\mathrm{In}(I)\setminus I} X_J+\sum_{J\in\mathrm{In}(S\cup I)\setminus\mathrm{In}(I)} X_J\Big\}.$$

Adding over $I\in\mathrm{No}(S)$ and taking expectations gives

$$\Big|\sum_{I\in\mathrm{No}(S)} E\{X_I[x(W_{S,I}+1)-x(W)]\}\Big|$$
$$\leqq (2\|x\|\wedge\Delta x)\Big\{\sum_{\substack{I\neq J\in\mathcal{N}\\ I\cap J\neq\varnothing}} E(X_IX_J)+\sum_{\substack{I\in\mathrm{No}(S)\\ J\in\mathrm{In}(S\cap I)\setminus\mathrm{In}(I)}} E(X_IX_J)\Big\}$$
$$\leqq \Big\{\sum_{j=1}^{k-1} q_jC_j+Nq_0i_S\Big\}(2\|x\|\wedge\Delta x).$$

The last term is easily seen to be no larger than $2p\|x\|i_S$.

*Remark.* It is always possible to choose $S$ in such a way that $i_S\leqq ksN/M$. This is because

$$\sum_{S\in D(s;1,M)} i_S=\sum_{I\in\mathcal{N}}\Big\{\binom{M}{s}-\binom{M-k}{s}\Big\},$$

and hence, for some $S$ among the $\binom{M}{s}$ possibilities, we must have

$$i_S\leqq N\Big\{1-\binom{M-k}{s}\Big/\binom{M}{s}\Big\}\leqq N\Big\{1-\Big(1-\frac{k}{M}\Big)^s\Big\}\leqq Nks/M.$$

*Lemma* 2.

$$\Big|\sum_{I\in\mathrm{No}(S)} E\{(X_I-p)x(W_{S,I}+1)\}\Big| \leqq \|x\|(k!)^{\frac12}N\sum_{j=0}^{k}\Big\{\binom{k}{j}\Big/(k-j)!\Big\}^{\frac12}\eta_js^{-\frac12 j}.$$

*Proof.* The $\sigma$-fields

$$\mathcal{Y}_{S,I}=\sigma\{Y_j;\, j\in I\cup S\}$$

and
$$\mathscr{X}_{S,I} = \sigma\{Y_j; j \in \{1, \cdots, M\} \setminus (I \cup S)\}$$
are generated by disjoint subsets of the random variables $\{Y_j\}_{j=1}^{M}$; $X_I \in \mathscr{Y}_{S,I}$ and $W_{S,I} \in \mathscr{X}_{S,I}$. Hence

$$\begin{aligned} E\{(X_I - p)x(W_{S,I}+1)\} &= E\{E\{(X_I - p)x(W_{S,I}+1) \mid \mathscr{X}_{S,I}\}\} \\ &= E\{x(W_{S,I}+1)E\{(X_I - p) \mid \mathscr{X}_{S,I}\}\} \\ &= E\left\{x(W_{S,I}+1)\sum_{J \in D_{S \cup I}} \binom{s+k}{k}^{-1} E\{(X_J - p) \mid \mathscr{X}_{S,I}\}\right\}, \end{aligned}$$

where $D_K$ denotes the set of $k$-subsets of $K$, the last equality because the $Y$'s are exchangeable. Thus

$$\begin{aligned} |E\{(X_I - p)x(W_{S,I}+1)\}| &= \left|E\left\{x(W_{S,I}+1)\sum_{J \in D_{S \cup I}} \binom{s+k}{k}^{-1}(X_J - p)\right\}\right| \\ &\leqq \left\{\|x\| \Big/ \binom{s+k}{k}\right\} E\left[\left\{\sum_{J \in D_{S \cup I}} (X_J - p)\right\}^2\right]^{\frac{1}{2}}. \end{aligned}$$

Direct computation now shows that for each $I \in \mathrm{No}\,(S)$

$$E\left\{\sum_{J \in D_{S \cup I}} (X_J - p)\right\}^2 \leqq \binom{s+k}{k} \sum_{j=0}^{k} \binom{k}{j}\binom{s}{k-j}\eta_j^2,$$

and hence the lemma follows.

*Lemma* 3. For any $A \subset Z^+$ and $\lambda > 0$,

$$\begin{aligned} |P[W \in A] - \mathscr{P}_\lambda(A)| \leqq\ & 2(1 \wedge 1{\cdot}4\lambda^{-\frac{1}{2}})\left\{\tfrac{1}{2}\varepsilon + pi_S + \sum_{j=0}^{k} N\eta_j s^{-j/2}\psi_{kj}\right\} \\ & + (1 \wedge \lambda^{-1})\left\{p^2 \sum_{I \in \mathscr{N}} i_I + \sum_{j=1}^{k-1} q_j C_j + N(p^2 + q_0)i_S\right\}, \end{aligned}$$

where $\varepsilon = |\lambda - Np|$ and $\psi_{kj} = \frac{1}{2}\binom{k}{j}(j!)^{\frac{1}{2}}$.

*Proof.* Given $A \subset Z^+$, define $x = x(\lambda, A)$, $\mathscr{P}_\lambda(\cdot)$ and $U_m$ as in the introduction. Recall that

$$E[\lambda x(W+1) - Wx(W)] = P[W \in A] - \mathscr{P}_\lambda(A).$$

Now
$$|E\{(\lambda - Np)x(W+1)\}| \leqq \varepsilon \|x\|,$$

and, applying Lemmas 1 and 2,

$$\begin{aligned} |E\{Npx(W+1) - Wx(W)\}| \leqq\ & (2\|x\| \wedge \Delta x)\left\{p^2\left(Ni_S + \sum_{I \in \mathscr{N}} i_I\right) + \sum_{j=1}^{k-1} q_j C_j + Nq_0 i_S\right\} \\ & + 2\|x\|\left\{pi_S + \tfrac{1}{2}(k!)^{\frac{1}{2}} \sum_{j=0}^{k} \left\{\binom{k}{j} \Big/ (k-j)!\right\}^{\frac{1}{2}} N\eta_j s^{-j/2}\right\}. \end{aligned}$$

The lemma now follows from the estimates for $\|x\|$ and $\Delta x$ which are proved in the appendix.

*Remark.* Suppose that, as in Brown and Silverman (1979), $\mathcal{N} = D$ and the $Y_j$'s are independent and identically distributed. Then the random variables $\{Y_j\}_{j=1}^{M}$ are exchangeable amongst the larger collection $\{Y_j\}_{j=1}^{\infty}$, and so we may reformulate their problem by taking $M = \infty$ and $\mathcal{N} = D(k; 1, n)$. Now $S$ may be chosen as large as we like from $\{n+1, n+2, \cdots\}$ while keeping $i_S = 0$, and, because of independence of the $Y_j$'s, $\eta_0 = 0$. Hence the estimate in Lemma 3 is no larger than

$$(1 \wedge 1{\cdot}4\lambda^{-\frac{1}{2}})\varepsilon + (1 \wedge \lambda^{-1})\left\{Np^2 \,.\, Nk^2/n + \sum_{j=1}^{k-1} q_j N\binom{k}{j}\frac{n^{k-j}}{(k-j)!}\right\}$$
$$\leqq (1 \wedge \lambda^{-1})\left\{\lambda^2 k^2/n + q_{k-1} N n^{k-1} \sum_{j=1}^{k-1}\binom{k}{j}\frac{n^{1-j}}{(k-j)!}\right\},$$

by Lemma 1 of Silverman and Brown (1978), if $\lambda = Np$. Setting $\rho = n^{2k}(q_{k-1} - p^2)$, this yields a bound on the total variation distance between $W$ and $\mathcal{P}_\lambda$ as follows.

*Corollary* 4. In the above circumstances, the total variation distance between $W$ and $\mathcal{P}_\lambda$ cannot exceed

$$(1 \wedge \lambda^{-1})\{n^{-1}\lambda^2(k^2 + b_n) + \rho n^{-1} c_n / k!\},$$

where

$$c_n = \sum_{j=1}^{k-1}\binom{k}{j}\frac{n^{1-j}}{(k-j)!} \to \frac{k}{(k-1)!},$$

$$b_n = \left\{n^k \Big/ \binom{n}{k}\right\} c_n \to k^2,$$

as $n \to \infty$.

This improves on the result of Brown and Silverman (1979) because $(\rho n^{-1})^{\frac{1}{2}}$ is replaced by $\rho n^{-1}$ in the convergence rate, because of the extra factor $1 \wedge \lambda^{-1}$, useful for large $\lambda$, and because all the constants are explicit. Of course, if better estimates of $q_j$, $1 \leqq j \leqq k-2$, are known, there is a corresponding improvement in the total variation bound.

*Proof of Theorem* 1. Choose $S$ so that $i_S \leqq ksN/M$, and let $s \to \infty$ in such a way that $Ns^{-k} \to 0$ and $N^2 q_j s^{-j} \to 0$, $1 \leqq j \leqq k-1$, whilst still ensuring that $s/M \to 0$; this is possible by (iii) and (v). Taking the estimate in Lemma 3 term

by term, it follows that:

$$\varepsilon \to 0 \text{ by (i);}$$

$$pi_S \leqq ksNp/M \to 0 \text{ by (i) and the choice of } s;$$

$$\sum_{j=0}^{k} \psi_{kj} N \eta_j s^{-\frac{1}{2}j} \to 0 \text{ by (i), (vi) and the choice of } s;$$

$$p^2 \sum_{i \in \mathcal{N}} i_I \to 0 \text{ by (i) and (ii);}$$

$$\sum_{j=1}^{k-1} q_j C_j \to 0 \text{ by (iv);}$$

and

$$N(p^2+q_0)i_S \leqq ksM^{-1}(N^2p^2+N^2q_0) \to 0 \text{ by (i), (vi) and the choice of } s.$$

Hence the distribution of $W$ converges to $\mathcal{P}_\lambda$.

*Remarks.* Let $M=n$, $n \geqq k$, and take $\mathcal{N}=\{(1, 2, \cdots, k-1, r)\}_{r=k}^{M}$. Then, for each $I \in \mathcal{N}$, $i_I = N$, showing that Condition (ii) need not always be satisfied. It is clear that none of the other conditions is automatic. Nor is (iv) necessarily implied by (v). For if $M=n$, $n \geqq k$, and, for any $j$, $0 \leqq j < k$, $\mathcal{N} = \{(1, 2, \cdots, j, j+r+1, \cdots, k+r)\}_{r=0}^{M-k}$, then

$$C_{j'}=0, \qquad j'<j, \qquad C_j \sim n^2, \quad \text{and} \quad C_{j'} \sim 2n, \qquad j'>j.$$

*Proof of Theorem* 2. Argue as in the proof of Theorem 1, using the $\lambda$-dependent estimates in Lemma 3, and now letting $s \to \infty$ in such a way that $Ns^{-k} \to 0$ and $\lambda^{-1}N^2\eta_j^2 s^{-j} \to 0$, $1 \leqq j \leqq k-1$, while still ensuring that $\lambda s M^{-1} \to 0$.

## 4. Applications

(i) *Knox's statistic.* In a study of childhood leukaemia, Knox (1964) suggested that a suitable test statistic for contagion would be the number of reported outbreaks of leukaemia which are close in time and close in space. He conjectured that, under the hypothesis of no association between time and position of occurrence, this statistic could be regarded as a Poisson random variable. In support of this conjecture, Barton and David (1966) gave a graph-theoretic proof that, under the above hypothesis, the mean and variance of this statistic are equal, and that a Poisson limit holds under certain assumptions.

Let $\boldsymbol{Z}_1, \cdots, \boldsymbol{Z}_n$ be $n$ points on the plane and $0 \leqq T_1 \leqq T_2 \leqq \cdots \leqq T_n$ be $n$ time points. Knox's statistic is

$$K_n = \sum_{i<j} I(|\boldsymbol{Z}_i - \boldsymbol{Z}_j| < d_n) I(|T_i - T_j| < t_n),$$

where the measures of closeness, $d_n$ and $t_n$, are chosen to make the mean of $K_n$ small. We are interested in the distribution of $K_n$ when the $\boldsymbol{Z}$'s and $T$'s are fixed but associated at random.

Suppose that $\pi$ is a random permutation of $(1, \cdots, n)$ and consider

$$K_n = \sum_{i<j} I(|\boldsymbol{Z}_{\pi(i)} - \boldsymbol{Z}_{\pi(j)}| < d_n) I(|T_i - T_j| < t_n)$$
$$= \sum_{(i,j)\in\mathcal{N}} \phi(\boldsymbol{Z}_{\pi(i)}, \boldsymbol{Z}_{\pi(j)}),$$

where $\mathcal{N} = \{(i, j) : |T_i - T_j| < t_n\}$. As the $\{\boldsymbol{Z}_{\pi(i)}\}_{i=1}^n$ are exchangeable, $K_n$ is an example of the sort of statistic studied in the previous section. In fact there is rather more structure in $K_n$, due to the fact that time is measured on a linear scale. If $|T_i - T_j| < t_n$ for some $j > i$ then both $|T_i - T_{j-k}| < t_n$ and $|T_{i+k} - T_j| < t_n$ for $k = 1, \cdots, j-i-1$. Thus $K_n$ is a statistic of the form considered in Corollary 3 and hence, under the conditions listed there, is asymptotically distributed as a Poisson random variable.

Given a set of observations, one cannot tell whether it is one of a possible sequence of such observations satisfying the conditions of Corollary 3. However, Lemma 3 can be used to bound the total variation distance of the distribution of $K_n$ from that of a Poisson random variable. Unfortunately, like most general bounds, Lemma 3 is not useful in a practical situation. Data similar to Knox's were simulated and the criteria $d_n$ and $t_n$ chosen to ensure that a $2 \times 2$ contingency table of close/not close distances against close/not close times had the same marginals as Knox's table. More precisely, 96 points independently and uniformly distributed on the unit square and 96 time points independently and uniformly distributed on the unit interval were chosen. The criteria $d_n$ and $t_n$ were chosen so that 25 of the interpoint distances were less than $d_n$ and 152 of the time differences were less than $t_n$. The statistic $K_n$ was then calculated from these data, as was the bound in Lemma 3 for the special case of $s_n$ being taken as large as possible while leaving $i_S = 0$. In the particular simulation, $K_n$ was 0 and the bound 0·71.

While the bound calculated from Lemma 3 was bad, 10000 repetitions of the above simulation gave an empirical distribution of $K_n$ which, when fitted to the Poisson distribution with mean $25 \times 152/4{,}560 = 5/6$, produced a total variation distance of 0·016 and a Pearson $\chi^2$ of 13·404 with 7 degrees of freedom.

In order to obtain a general result for the distribution of $K_n$, a model for generating the $\boldsymbol{Z}$'s and $T$'s has to be assumed. If these random variables are all independent, the $\boldsymbol{Z}$'s and $T$'s being associated at random, then, when a sufficient number of points are observed, most realisations give a $K_n$ which is close to Poisson in a sense which is made precise by the following corollary.

Suppose that $\boldsymbol{Z}_1, \cdots, \boldsymbol{Z}_n$ are independent, identically distributed points on the plane with a common density which is bounded and continuous. Suppose also that $T_1, \cdots, T_n$ are independent and identically distributed points on the positive half-line. For fixed $n$, choose random variables $\delta_n$ and $\tau_n$ such that

$$\frac{1}{\binom{n}{2}} \sum_{i<j} I(|\boldsymbol{Z}_i - \boldsymbol{Z}_j| \leqq \delta_n) \lessapprox \lambda/n$$

and

$$\sum_{i<j} I(|T_i - T_j| \leqq \tau_n) \approx n.$$

Define

$$K_n = \sum_{i<j} I(|T_i - T_j| < \tau_n) I(|\boldsymbol{Z}_{\pi(i)} - \boldsymbol{Z}_{\pi(j)}| < \delta_n),$$

where $\pi$ is a random permutation of $(1, \cdots, n)$.

*Corollary* 5. With the statistic $K_n$ defined as above, for all $\varepsilon > 0$ there exists an $n_0$ such that for all $n \geqq n_0$

$$P\left\{ \sup_{A \subset Z^+} |P(K_n \in A \mid \boldsymbol{Z}_1, \cdots, \boldsymbol{Z}_n; T_1, \cdots, T_n) - \mathscr{P}_\lambda(A)| < \varepsilon \right\} > 1 - \varepsilon$$

for all $A \subset Z^+$.

*Proof.* The corollary will be proved if it can be shown that the conditions of Corollary 3 hold 'in probability'. We have

$$N = n \to \infty;$$

$$m \leqq \sqrt{(2n)} = o(n) \text{ always};$$

$$Np \to \lambda, \text{ by the choice of } \delta_n \text{ and } \tau_n;$$

$$C_1 \leqq Cn^{\frac{3}{2}} \text{ always.}$$

Furthermore,

$$\begin{aligned} q_1 &\leqq Cn^{-3} \sum_i \sum_{j<k} I(|\boldsymbol{Z}_i - \boldsymbol{Z}_j| \leqq \delta_n) I(|\boldsymbol{Z}_i - \boldsymbol{Z}_k| \leqq \delta_n) \\ &\leqq Cn^{-3} \sum\sum I(|\boldsymbol{Z}_i - \boldsymbol{Z}_j| \leqq \delta_n) \{ I(|\boldsymbol{Z}_i - \boldsymbol{Z}_k| \leqq \varepsilon_n) + I(\delta_n \geqq \varepsilon_n) \} \\ &\leqq Cn^{-3} \sum\sum \{ I(|\boldsymbol{Z}_i - \boldsymbol{Z}_j| \leqq \varepsilon_n) + I(\delta_n \geqq \varepsilon_n) \} I(|\boldsymbol{Z}_i - \boldsymbol{Z}_k| \leqq \varepsilon_n) \\ &\quad + C\lambda n^{-1} I(\delta_n \geqq \varepsilon_n). \end{aligned}$$

Now on the set $[\delta_n \geqq \varepsilon_n]$,

$$\frac{1}{\binom{n}{2}} \sum_{i<k} I(|\boldsymbol{Z}_i - \boldsymbol{Z}_k| \leqq \varepsilon_n) \leqq \lambda/n,$$

and so

$$q_1 \leqq Cn^{-3} \sum\sum I(|\boldsymbol{Z}_i - \boldsymbol{Z}_j| \leqq \varepsilon_n) I(|\boldsymbol{Z}_i - \boldsymbol{Z}_k| \leqq \varepsilon_n) + 2c\lambda n^{-1} I(\delta_n \geqq \varepsilon_n).$$

Hence

$$Eq_1(\boldsymbol{Z}_1, \cdots, \boldsymbol{Z}_n; T_1, \cdots, T_n) \leqq c_1 n^{-1} P(\delta_n \geqq \varepsilon_n) + c_2 P(|\boldsymbol{Z}_1 - \boldsymbol{Z}_2| \leqq \varepsilon_n, |\boldsymbol{Z}_1 - \boldsymbol{Z}_3| \leqq \varepsilon_n).$$

Since the $\boldsymbol{Z}$'s have a bounded density,

$$P(|\boldsymbol{Z}_1 - \boldsymbol{Z}_2| \leqq \varepsilon_n, |\boldsymbol{Z}_1 - \boldsymbol{Z}_3| \leqq \varepsilon_n) = E(P(|\boldsymbol{Z}_1 - \boldsymbol{Z}_2| \leqq \varepsilon_n \mid \boldsymbol{Z}_1)^2) \leqq 4\varepsilon_n^2 L^2,$$

where $L$ is a bound on the density of the $\boldsymbol{Z}$'s. Choose $\varepsilon_n \sim c/n$ so that $P[|\boldsymbol{Z}_i - \boldsymbol{Z}_j| \leqq \varepsilon_n] \geqq 2\lambda/n$, possible because the $\boldsymbol{Z}$'s have a continuous density.

Then note that

$$I(\delta_n \geqq \varepsilon_n) = I\left[\binom{n}{2}^{-1} \sum_{i<j} I(|\boldsymbol{Z}_i - \boldsymbol{Z}_j| \leqq \varepsilon_n) \leqq \lambda/n\right],$$

so that if

$$U = \binom{n}{2}^{-1} \sum_{i<j} I(|\boldsymbol{Z}_i - \boldsymbol{Z}_j| \leqq \varepsilon_n),$$

then

$$P(\delta_n \geqq \varepsilon_n) = P(U \leqq \lambda/n) \leqq n^2 \lambda^{-2} \operatorname{Var} U.$$

It is easy to see that

$$\begin{aligned} \operatorname{Var} U &= O(n^{-2}) P(|\boldsymbol{Z}_i - \boldsymbol{Z}_j| \leqq \varepsilon_n) \\ &\quad + O(n^{-1}) P(|\boldsymbol{Z}_1 - \boldsymbol{Z}_2| \leqq \varepsilon_n, |\boldsymbol{Z}_1 - \boldsymbol{Z}_3| \leqq \varepsilon_n) \\ &= O(n^{-2}) P(|\boldsymbol{Z}_i - \boldsymbol{Z}_j| \leqq \varepsilon_n) \\ &\quad + O(n^{-1}\varepsilon_n^2). \end{aligned}$$

Thus

$$P(\delta_n \geqq \varepsilon_n) = O(n^{-1}),$$

so that, with this choice of $\varepsilon_n$,

$$Eq_1 \leqq c_3 n^{-2}.$$

It follows that

$$C_1 q_1 \to^p 0 \quad \text{as} \quad n \to \infty.$$

Finally, consider

$$N^2|q_0-p^2| = n^2\left\{\binom{n-2}{2}\right\}^{-1}\Bigg|\sum_{\substack{(i,j)\cap(k,l)=\varnothing \\ i<j,k<l}} I(|\boldsymbol{Z}_i-\boldsymbol{Z}_j|\leqq\delta_n)I(|\boldsymbol{Z}_k-\boldsymbol{Z}_l|\leqq\delta_n)$$

$$-\left(\sum_{i<j} I(|\boldsymbol{Z}_i-\boldsymbol{Z}_j|\leqq\delta_n)\right)^2\Bigg| + O(n^{-1})$$

$$\leqq \frac{c_4}{n^2}\left\{\sum_{i<j} I(|\boldsymbol{Z}_i-\boldsymbol{Z}_j|\leqq\delta_n)\right.$$

$$\left.+\sum_i\sum_{j<k} I(|\boldsymbol{Z}_i-\boldsymbol{Z}_j|\leqq\delta_n)I(|\boldsymbol{Z}_i-\boldsymbol{Z}_k|\leqq\delta_n)\right\} + O(n^{-1})$$

$$= O(n^{-1}) + c_5 n q_1.$$

It follows from the above calculations that

$$EN^2|q_0-p^2| \to 0$$

and hence that

$$N^2|q_0-p^2| \to^p 0$$

also.

(ii) *Rare patterns.* Suppose that one has $n$ sites on a line, $m$ of them coloured or distinguished in some way. Assume, as a null hypothesis, that the coloured sites have been chosen at random. To test an alternative of contagion, one might count the number of times $k$ coloured sites are contiguous, rejecting the null hypothesis if the count were too large. In order to fix the significance points of the test, the distribution under the null hypothesis of the number of occurrences of $k$ continguous coloured sites needs to be found. If $k$ is chosen such that the occurrence of $k$ coloured sites together is suitably rare, one would expect the number of such occurrences to be close to Poisson. A similar theorem is proved for a Bernoulli model in Brown (1981), Corollary 2.

Let $Y_1, \cdots, Y_n$ be random variables, indicating the type of site:

$$Y_i = \left.\begin{matrix}1\\0\end{matrix}\right\} \text{ if site is } \begin{cases}\text{coloured}\\ \text{otherwise.}\end{cases}$$

The $Y_1, \cdots, Y_n$ are exchangeable. Consider $k$ to be fixed and set

$$E_n = \sum_{i=1}^{n-k+1} I(Y_i Y_{i+1} \cdots Y_{i+k-1} = 1),$$

the number of occurrences of $k$ coloured sites together. Under the conditions of Corollary 1, $E_n$ will be asymptotically equivalent to the $W$ defined there, so one needs only to check those conditions. Here $M = n$ and

$$p = (m)_k/(n)_k$$

so that

$$np = n(m)_k/(n)_k.$$

If $np \to \lambda$, $m = O(n^{1-k^{-1}})$ and

$$nq_j = n(m)_{2k-j}/(n)_{2k-j} = O(n^{j/k-1}) \to 0 \quad \text{if} \quad 1 \leqq j \leqq k-1.$$

Finally,

$$n^2 |q_0 - p^2| = n^2 \left| \frac{(m)_{2k}}{(n)_{2k}} - \left[\frac{(m)_k}{(n)_k}\right]^2 \right|$$
$$= O(n^{k^{-1}-1}) \to 0 \quad \text{if} \quad k \geqq 2.$$

Thus we have proved the following.

*Corollary* 6. Let $Y_{n1}, \cdots, Y_{nn}$ be exchangeable 0–1 random variables and set

$$E_n = \sum_{i=1}^{n-k+1} I(Y_{ni} Y_{n,i+1} \cdots Y_{n,i+k-1} = 1).$$

Then $E_n$ converges in distribution to a Poisson law as $n \to \infty$ if $m = \sum_{i=1}^n Y_{ni}$ is such that

$$\lim_{n\to\infty} n(m)_k/(n)_k = \lambda.$$

## 5. Appendix

Bounds are required for the test function, defined in the introduction, namely

$$x(0) = 0$$
$$x(m+1) = \lambda^{-m-1} e^{\lambda} m! \, [\mathcal{P}_\lambda(A \cap U_m) - \mathcal{P}_\lambda(A)\mathcal{P}_\lambda(U_m)], \qquad m \geqq 0,$$

where $A \subset Z^+$ and where $\mathcal{P}_\lambda(\cdot)$ and $U_m$ have already been defined. It is easy to show that $x$ is bounded and Chen (1975), Lemmas 3.3 and 3.4, shows that

$$\|x\| \leqq 4(\lambda^{-\frac{1}{2}} \wedge 1) \quad \text{and} \quad \Delta x \leqq 6(\lambda^{-\frac{1}{2}} \wedge 1).$$

In order to have reasonable constants in Lemma 3 and to improve the results for large $\lambda$, we here derive better bounds for $\|x\|$ and $\Delta x$.

*Lemma* 4. For $x$ defined as above,
(i) $\|x\| \leqq 1 \wedge 1{\cdot}4\lambda^{-\frac{1}{2}}$;
(ii) $\Delta x \leqq \lambda^{-1}(1 - e^{-\lambda}) \leqq (1 \wedge \lambda^{-1})$.

*Proof.* From the definition of $x$,

$$x(m+1) = \lambda^{-m-1} m! \, e^{\lambda} \{\mathcal{P}_\lambda(A \cap U_m)\mathcal{P}_\lambda(U_m^-) - \mathcal{P}_\lambda(A \cap U_m^-)\mathcal{P}_\lambda(U_m)\},$$

where $B^-$ denotes the complement of $B$. Hence, since $0 \leqq P(A \cap B)P(B^-) \leqq P(B)P(B^-)$, for any $A \subset Z^+$,

$$|x(m+1)| \leqq m!\, e^{\lambda}\lambda^{-m-1}\mathscr{P}_{\lambda}(U_m)\mathscr{P}_{\lambda}(U_m^-). \tag{2}$$

For $m \leqq \lambda$,

$$|x(m+1)| \leqq m!\, e^{\lambda}\lambda^{-m-1}\mathscr{P}_{\lambda}(U_m) = \frac{1}{\lambda}\sum_{j=0}^{m} \lambda^{-j} m!/(m-j)! \tag{3}$$

$$\leqq \frac{1}{\lambda}\sum_{j=0}^{m} (m/\lambda)^j \leqq (\lambda - m)^{-1};$$

similarly, for $m \geqq \lambda - 3$,

$$|x(m+1)| \leqq m!\, e^{\lambda}\lambda^{-m-1}\mathscr{P}_{\lambda}(U_m^-) = \sum_{j=0}^{\infty} \lambda^j m!/(m+1+j)!$$

$$\leqq \frac{1}{m+1}\left\{1 + \frac{\lambda}{m+2}\sum_{j=0}^{\infty} (\lambda/(m+3))^j\right\} \tag{4}$$

$$= \frac{(m+2)(m+3)+\lambda}{(m+1)(m+2)(m+3-\lambda)}.$$

Hence, from (3), $|x(m+1)| \leqq 1$ whenever $m \leqq \lambda - 1$; from (4), $|x(m+1)| \leqq 1$ whenever $m \geqq \lambda - 1$ if $m \geqq 2$, and whenever $\lambda \leqq 12/7$ if $m = 1$. It therefore remains to consider $m = 0$ with $0 \leqq \lambda \leqq 1$ and $m = 1$ with $12/7 \leqq \lambda \leqq 2$. However, it follows from (2) that, for $m = 0$, $|x(1)| \leqq \lambda^{-1}(1 - e^{-\lambda}) \leqq 1$ for all $\lambda$: for $m = 1$,

$$|x(2)| \leqq \{\lambda^{-2}(1+\lambda)\}\{1 - e^{-\lambda}(1+\lambda)\},$$

and, by observing that the two terms in braces are monotonic in $\lambda$, it follows that, for $12/7 \leqq \lambda \leqq 2$, $|x(2)| \leqq (133/144)\{1 - 3e^{-2}\} < 1$. Thus $\|x\| \leqq 1$.

Returning to (2), it is always true that $\mathscr{P}_{\lambda}(U_m)\mathscr{P}_{\lambda}(U_m^-) \leqq \frac{1}{4}$. Using Stirling's approximation that $m! \leqq \sqrt{(2\pi)} m^{m+\frac{1}{2}} \exp(-m + 1/12m)$, $m \geqq 1$, it follows that

$$|x(m+1)| \leqq \tfrac{1}{4}\sqrt{(2\pi)}\lambda^{-\frac{1}{2}}(m/\lambda)^{m+\frac{1}{2}} \exp\left(\lambda - m + \frac{1}{12m}\right)$$

$$\leqq \tfrac{1}{4}\sqrt{(2\pi)}\lambda^{-\frac{1}{2}} \exp\left\{\lambda^{-1}(m-\lambda)(m-\lambda+\tfrac{1}{2}) + \frac{1}{12m}\right\}, \qquad m \geqq 1. \tag{5}$$

As before, $|x(1)| \leqq \lambda^{-1}(1 - e^{-\lambda}) \leqq \lambda^{-1}$. Comparing (5) with (3) and (4), and taking the smaller, yields the estimate $\|x\| \leqq c\lambda^{-\frac{1}{2}}$, for $c$ with a value no greater than 1·4. This completes the proof of Part (i).

For Part (ii), let $x_j$ denote $x_{\lambda,\{j\}}$, and note that $x_{\lambda,A} = \sum_{j \in A} x_j$. From its definition,

$$x_j(m+1) = \begin{cases} (\lambda^{j-1}/j!)\{m!\, \lambda^{-m}\mathscr{P}_{\lambda}(U_m^-)\} & m \geqq j \\ -(\lambda^{j-1}/j!)\{m!\, \lambda^{-m}\mathscr{P}_{\lambda}(U_m)\} & m < j, \end{cases}$$

and, from the series expansion of the Poisson probabilities, it is easily seen that $x_j$ is positive and decreasing in $m \geqq j+1$ and is negative and decreasing in $m \leqq j$. Hence the only positive value taken by $x_j(m+1)-x_j(m)$ is

$$x_j(j+1)-x_j(j) = \frac{e^{-\lambda}}{\lambda}\left\{\sum_{r=j+1}^{\infty}(\lambda^r/r!)+\sum_{r=1}^{j}(\lambda^r/r!)\frac{r}{j}\right\}$$
$$\leqq \frac{e^{-\lambda}}{\lambda}\{e^{\lambda}-1\} = \lambda^{-1}\{1-e^{-\lambda}\}.$$

Thus, for any $A \subset Z^+$,

$$x_{\lambda,A}(m+1)-x_{\lambda,A}(m) = I[m \in A]\{x_m(m+1)-x_m(m)\}$$
$$+\sum_{\substack{j\in A\\ j\neq m}}\{x_j(m+1)-x_j(m)\},$$

in which expression only the first term is positive, and so

$$\sup_{m\in Z^+}\max_{A\subset Z^+}\{x_{\lambda,A}(m+1)-x_{\lambda,A}(m)\} = \sup_{m\in Z^+}\{x_m(m+1)-x_m(m)\} \leqq \lambda^{-1}(1-e^{-\lambda}).$$

Since also $x_{\lambda,A} = -x_{\lambda,A^-}$, Part (ii) follows from the above estimate.

## Acknowledgements

The authors are grateful for the help of J. Dawson in performing the simulations mentioned in Section 4, for fruitful discussions with T. C. Brown and D. J. Aldous, and for a referee's pertinent comments on the presentation of the paper.

## References

BARTON, D. E. AND DAVID, F. N. (1966) The random intersection of two graphs. In *Research Papers in Statistics*, ed. F. N. David, Wiley, New York, 445–459.

BERMAN, M. AND EAGLESON, G. K. (1983) A Poisson limit theorem for incomplete symmetric statistics. *J. Appl. Prob.* **20,** 47–60.

BROWN, T. C. (1981) Compensators and Cox convergence. *Math. Proc. Camb. Phil. Soc.* **90,** 305–319.

BROWN, T. C. AND SILVERMAN, B. S. (1979) Rates of Poisson convergence for $U$-statistics. *J. Appl. Prob.* **16,** 428–432.

CHEN, L. H. Y. (1975) Poisson approximation for dependent trials. *Ann. Prob.* **3,** 534–545.

EAGLESON, G. K. (1979) A Poisson limit theorem for weakly exchangeable arrays. *J. Appl. Prob.* **16,** 794–802.

KNOX, G. (1964) Epidemiology of childhood leukaemia in Northumberland and Durham. *Brit. J. Prev. Soc. Med.* **18,** 17–24.

SILVERMAN, B. S. AND BROWN, T. C. (1978) Short distances, flat triangles and Poisson limits. *J. Appl. Prob.* **15,** 815–825.

STEIN, C. (1970) A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. 6th Berkeley Symp. Math. Statist. Prob.* **2,** 583–602.