The Annals of Probability
1984, Vol. 12, No. 3, 805–818

RUNS IN m-DEPENDENT SEQUENCES

BY SVANTE JANSON

Uppsala University

Consider a stationary m-dependent sequence of random indicator variables. If $m > 1$, assume further that any two nonzero values are separated by at least $m - 1$ zeros.

This paper studies the sequence of the lengths of the successive intervals between the nonzero values of the original sequence, and it is shown that, provided a technical condition holds, these lengths converge in distribution (and their moments converge exponentially fast) in all cases but one.

1. Introduction. Let m be a positive integer, fixed throughout this paper, and consider a stationary sequence of random variables $\{I_i\}_0^\infty$ with the following properties (I will denote a generic element of $\{I_i\}$).

I_i are indicator variables, i.e. $I_i = 0$ or 1. To avoid trivial complications we assume that $0 < P(I = 1) < 1$.

$\{I_i\}$ is m-dependent, i.e. $\{I_i\}_0^n$ and I_{n+m+1} are independent for every n.

$\{I_i\}$ is m-separated, i.e. $I_n I_{n+k} = 0$ if $k = 1, \cdots, m - 1$.

Note that the last condition is void when $m = 1$. We define, with m as above,

$$S_n = \textstyle\sum_m^n I_i, \quad n = 0, 1, \cdots, \tag{1.1}$$

$$N_k = \min\{n: S_n = k\}, \quad k = 0, 1, \cdots, \tag{1.2}$$

(the corresponding renewal process) and

$$L_k = N_k - N_{k-1}, \quad k = 1, 2, \cdots. \tag{1.3}$$

Note that, by this definition, $S_n = 0$ for $n \le m - 1$. Thus $N_0 = 0$ and $L_1 = N_1 \ge m$. Further, the assumption that $\{I_i\}$ be m-separated is equivalent to $L_k \ge m$ for $k \ge 2$. Thus $L_k \ge m$ for every k.

The purpose of this paper is to study the distributions of L_k and, in particular, to prove convergence theorems.

To obtain complete results we will impose one further condition.

(*) There exists a sequence $\{\xi_i\}$ of i.i.d. random variables and a measurable function α such that $I_i = \alpha(\xi_{i-m}, \cdots, \xi_i)$.

Obviously, any sequence $\{I_i\}$ satisfying (*) is m-dependent. It seems to be unknown whether the converse holds, i.e. whether every m-dependent stationary sequence may be thus represented. Hence it is conceivable that this condition is redundant.

Received June 1983.

AMS 1980 *subject classifications.* Primary 60K99; secondary 60G99, 60K05, 60F05, 60C05.

Key words and phrases. Runs, m-dependent sequences, random permutations.

We will prove that if (*) holds, then L_k converges in distribution, except in one exceptional case.

THEOREM. *Suppose that* $\{I_i\}_0^\infty$ *is an m-dependent, m-separated stationary sequence of indicator variables such that* (*) *holds. Then, unless* $\{I_i\}_0^\infty$ *has the distribution given in Example* 1 *of Section* 7,

$$L_k \to_d L_\infty \quad as \quad k \to \infty, \tag{1.4}$$

where the distribution of L_∞ *equals the conditional distribution of* L_1 *given that* $I_0 = 1$. *Furthermore, in this case there exists* $R > 1$ *such that*

$$EL_k^\ell = EL_\infty^\ell + O(R^{-k}) \quad for \quad \ell = 1, 2, \cdots. \tag{1.5}$$

The theorem is proved in the following sections together with various formulae for moments and generating functions. Examples, including applications to runs in very long permutations, are given in Section 7.

2. Preliminary lemmas.

LEMMA 1. *There exists* $C < \infty$, *such that for all k and* $\ell \geq 1$,

$$EL_k^\ell \leq \ell! C^\ell. \tag{2.1}$$

PROOF. Fix k and n. Since $I(N_{k-1} = n)$ and $I_{n+m+1}, I_{n+2(m+1)}, \cdots$ are independent,

$$P(L_k > j(m+1) \mid N_{k-1} = n) \leq P(I_{n+m+1} = 0, \cdots, I_{n+j(m+1)} = 0 \mid N_{k-1} = n)$$
$$= P(I = 0)^j.$$

Hence

$$P(L_k > j(m+1)) \leq P(I = 0)^j \tag{2.2}$$

and (2.1) follows by elementary computations.

We define

$$\mu_k = EL_k \tag{2.3}$$

and introduce the probability generating functions

$$g_k(z) = E(z^{L_k}), \quad |z| \leq 1. \tag{2.4}$$

LEMMA 2. *Let* $\tilde{g}$ *denote the probability generating function of the conditional distribution of* L_1 *given that* $I_0 = 1$. *Then*

$$\tilde{g}(z) = 1 - (1 - z)(EI)^{-1} z^{-m} g_1(z) \tag{2.5}$$

and

$$E(L_1 \mid I_0 = 1) = \tilde{g}'(1) = 1/EI. \tag{2.6}$$

PROOF. Since $\{I_i\}$ is stationary and m-separated, for every $n \geq m$,

$$
\begin{aligned}
&P(I_0 = 1)\cdot P(L_1 = n \mid I_0 = 1)\\
&\qquad = P(I_0 = 1, I_m = 0, \cdots, I_{n-1} = 0, I_n = 1)\\
&\qquad = P(I_0 = 1, I_1 = 0, \cdots, I_{n-1} = 0, I_n = 1)\\
&\qquad = P(I_1 = 0, \cdots, I_{n-1} = 0, I_n = 1)\\
&\qquad\quad - P(I_0 = 0, I_1 = 0, \cdots, I_{n-1} = 0, I_n = 1)\\
&\qquad = P(I_m = 0, \cdots, I_{n+m-2} = 0, I_{n+m-1} = 1)\\
&\qquad\quad - P(I_m = 0, \cdots, I_{n+m-1} = 0, I_{n+m} = 1)\\
&\qquad = P(L_1 = n + m - 1) - P(L_1 = n + m).
\end{aligned}
\tag{2.7}
$$

Hence, since $P(L_1 = n) = P(I_n = 1)$ for $m \leq n < 2m$,

$$
\begin{aligned}
P(I = 1)\tilde{g}(z) &= \textstyle\sum_m^\infty P(I = 1)P(L_1 = n \mid I_0 = 1)z^n\\
&= \textstyle\sum_m^\infty (P(L_1 = n + m - 1) - P(L_1 = n + m))z^n\\
&= \textstyle\sum_1^\infty (P(L_1 = n + m - 1) - P(L_1 = n + m))z^n\\
&= z^{1-m}g_1(z) - z^{-m}g_1(z) + P(I_m = 1).
\end{aligned}
$$

(2.5) is an immediate consequence. (2.6) follows by differentiation.

3. The basic lemma and its consequences.

LEMMA 3. *Let f be an arbitrary function. If $f \geq 0$, or if $E\mid f(L_1)\mid < \infty$, then for $k = 1, 2, \cdots$,*

$$\mu_k\cdot Ef(L_1) = E\left(\textstyle\sum_m^{L_k-1} f(j) + \sum_0^{m-1} f(L_{k+1} + j)\right). \tag{3.1}$$

PROOF. Define, for $i, n \geq 0$, $I_i^{(n)} = I_{i+n}$. Since $\{I_i\}$ is stationary, $\{I_i^{(n)}\}_{i=0}^\infty$ is a sequence with the same distribution as $\{I_i\}_0^\infty$. By (1.2) and (1.3),

$$L_1 = N_1 = \min\{n\colon S_n = 1\} = \min\{i \geq m\colon I_i = 1\} \tag{3.2}$$

and we imitate this and define

$$L_1^{(n)} = \min\{i \geq m\colon I_i^{(n)} = 1\} = \min\{j \geq m + n\colon I_j = 1\} - n. \tag{3.3}$$

Since $I_j = 1$ if and only if j equals some N_k, we see that if $N_{k-1} \leq n \leq N_k - m$, then $L_1^{(n)} = N_k - n$, and if $N_k - m < n \leq N_k$, then $L_1^{(n)} = N_{k+1} - n$.

Consequently, if we fix $k \geq 1$ and define

$$Z = \textstyle\sum_{n=1}^\infty I(S_{n-1} = k - 1)f(L_1^{(n)}), \tag{3.4}$$

then, since $S_{n-1} = k - 1$ if and only if $N_{k-1} \le n - 1 < N_k$,

$$
\begin{aligned}
Z &= \textstyle\sum_{N_{k-1}+1}^{N_k} f(L_1^{(n)}) = \sum_{N_{k-1}+1}^{N_k - m} f(N_k - n) + \sum_{N_k - m+1}^{N_k} f(N_{k+1} - n) \\
&= \textstyle\sum_{m}^{N_k - N_{k-1}-1} f(j) + \sum_0^{m-1} f(N_{k+1} - N_k + j) \qquad (3.5)\\
&= \textstyle\sum_{m}^{L_k - 1} f(j) + \sum_0^{m-1} f(L_{k+1} + j).
\end{aligned}
$$

However, since $L_1^{(n)}$ depends on $\{I_i\}_{n+m}^{\infty}$ only, $I(S_{n-1} = k - 1)$ and $f(L_1^{(n)})$ are independent. Furthermore, $L_1^{(n)}$ and L_1 are equidistributed, whence $Ef(L_1^{(n)}) = Ef(L_1)$. Thus,

$$
\begin{aligned}
EZ &= \textstyle\sum_{n=1}^{\infty} P(S_{n-1} = k - 1) Ef(L_1^{(n)}) = Ef(L_1) \sum_{n=1}^{\infty} P(S_{n-1} = k - 1) \qquad (3.6)\\
&= Ef(L_1) E \textstyle\sum_{n=1}^{\infty} I(S_{n-1} = k - 1) = Ef(L_1) EL_k.
\end{aligned}
$$

Combining (3.6) and (3.5), we obtain (3.1).

In the first application of Lemma 3 we choose $f(j) = j$. By (3.1),

$$
\begin{aligned}
\mu_k \cdot \mu_1 &= E(\textstyle\sum_m^{L_k - 1} j + \sum_0^{m-1} (L_{k+1} + j)) = E(\sum_0^{L_k - 1} j + mL_{k+1}) \qquad (3.7)\\
&= E(L_k(L_k - 1)/2) + m\mu_{k+1}.
\end{aligned}
$$

This relation enables us to compute the second moment of L_k if we know the first moments μ_1, μ_k, μ_{k+1}. Conversely, we may express μ_{k+1} in moments of L_1 and L_k.

More generally, we let $\ell = 1, 2, \cdots$, and choose $f(j) = \binom{j}{\ell}$. Since

$$
\begin{aligned}
&\textstyle\sum_m^{L_k - 1} \binom{j}{\ell} + \sum_0^{m-1} \binom{L_{k+1} + j}{\ell} \\
&= \binom{L_k}{\ell + 1} - \binom{m}{\ell + 1} + \binom{L_{k+1} + m}{\ell + 1} - \binom{L_{k+1}}{\ell + 1} \\
&= \binom{L_k}{\ell + 1} - \binom{m}{\ell + 1} + \textstyle\sum_{j=0}^{\ell+1} \binom{L_{k+1}}{j} \binom{m}{\ell + 1 - j} - \binom{L_{k+1}}{\ell + 1} \\
&= \binom{L_k}{\ell + 1} + \textstyle\sum_{j=1}^{\ell} \binom{L_{k+1}}{j} \binom{m}{\ell + 1 - j},
\end{aligned}
$$

(3.1) then yields

$$
\mu_k E \binom{L_1}{\ell} = E \binom{L_k}{\ell + 1} + \textstyle\sum_{j=1}^{\ell} \binom{m}{\ell + 1 - j} E \binom{L_{k+1}}{j}, \qquad (3.8)
$$

which we rearrange as

$$
E \binom{L_k}{\ell + 1} = \mu_k E \binom{L_1}{\ell} - \textstyle\sum_{j=1}^{\ell} \binom{m}{\ell + 1 - j} E \binom{L_{k+1}}{j}. \qquad (3.9)
$$

Using (3.8), we may recursively express moments EL_k^{ℓ} in terms of moments of L_1.

Finally, we choose $f(j) = z^j$, where $|z| < 1$. Then

$$\sum_m^{L_k-1} f(j) + \sum_0^{m-1} f(L_{k+1} + j) = \frac{z^m - z^{L_k}}{1 - z} + z^{L_{k+1}} \cdot \frac{1 - z^m}{1 - z}$$

and (3.1) yields, for $k = 1, 2, \cdots$

$$(3.10)\qquad \begin{aligned} \mu_k g_1(z) &= \frac{1}{1 - z}\,(z^m - g_k(z) + (1 - z^m) g_{k+1}(z)) \\ &= \frac{1}{1 - z}\,(1 - g_k(z) - (1 - z^m)(1 - g_{k+1}(z))). \end{aligned}$$

Alternatively, (3.8) may be obtained by differentiating (3.10), or (3.10) may be obtained by (3.8) and summation of power series.

Another form of (3.10) is

$$(3.11)\qquad g_{k+1}(z) = (1 - z^m)^{-1}(g_k(z) - z^m + (1 - z)\mu_k g_1(z)), \quad k = 1, 2, \cdots.$$

This recursion formula yields $g_2, g_3, \cdots$, provided g_1 is known.

4. Convergence. We introduce the generating function

$$(4.1)\qquad U(z) = \sum_1^\infty \mu_k z^k, \quad |z| < 1.$$

(By Lemma 1, this power series converges.)

Multiplying (3.10) by $(1 - z)(1 - z^m)^k$ and summing, we obtain, if $|z| < 1$ and $|1 - z^m| < 1$,

$$(4.2)\qquad \begin{aligned} &(1 - z) U(1 - z^m) g_1(z) \\ &\quad = \sum_1^\infty (1 - z)\mu_k (1 - z^m)^k g_1(z) \\ &\quad = \sum_1^\infty ((1 - z^m)^k (1 - g_k(z)) - (1 - z^m)^{k+1}(1 - g_{k+1}(z))) \\ &\quad = (1 - z^m)(1 - g_1(z)) \end{aligned}$$

and thus

$$(4.3)\qquad \frac{U(1 - z^m)}{1 - z^m} = \frac{1}{1 - z}\left(\frac{1}{g_1(z)} - 1\right).$$

Denote the right-hand side of (4.3) by $h(z)$. Thus h is a meromorphic function in the unit disc, and if $\omega^m = 1$ then, by (4.3), $h(\omega z) = h(z)$ e.g. for $0 < z < 1$, and hence for any z. Consequently, $h(z^{1/m})$ is a single-valued meromorphic function in the unit disc and, if $|z|, |1 - z| < 1$, $U(1 - z) = (1 - z)h(z^{1/m})$ and

$$(4.4)\qquad U(z) = zh((1 - z)^{1/m}).$$

When $m = 1$, this simplifies to

$$(4.5)\qquad U(z) = 1/g_1(1 - z) - 1.$$

PROOF OF THE THEOREM. We invoke the condition (*) through two lemmas whose proofs are postponed to Section 6.

LEMMA 4. *If (*) holds, then g_1 may be extended to a meromorphic function in the entire complex plane.*

In the following, g_1 denotes this (unique) extension.

LEMMA 5. *If (*) holds, then $g_1(z) \neq 0$ for every $z \neq 0$ with $|1 - z^m| = 1$, unless $m = 1$ and $\{I_i\}_0^\infty$ has the distribution given in Example 1 in Section 7.*

By Lemma 4, $h(z)$ defined above is a meromorphic function in the complex plane with poles at the zeroes of $g_1(z)$. Hence (4.4) defines $U(z)$ as a meromorphic function in the complex plane with the set of poles $\{1 - z^m : g_1(z) = 0\}$. Thus, an equivalent formulation of Lemma 5 is as follows.

LEMMA 5′. *If (*) holds, then $U(z)$ has no poles on $\{z: |z| = 1$ and $z \neq 1\}$, except in the exceptional case of Example 1.*

(Recall that $U(z)$ is analytic for $|z| < 1$.)
Since $g_1(z) = P(L_1 = m)z^m + \cdots$, and $P(L_1 = m) = P(I_m = 1)$, (4.3) yields

$$\lim_{z\to 0} z^m U(1 - z^m) = \lim_{z\to 0} \frac{1 - z^m}{1 - z}\left(\frac{z^m}{g_1(z)} - z^m\right) = \lim_{z\to 0} \frac{z^m}{g_1(z)} = \frac{1}{P(I = 1)}$$

whence

$$\lim_{z\to 1}(1 - z)U(z) = 1/P(I = 1) = 1/EI. \tag{4.6}$$

Thus $U(z)$ has a simple pole at $z = 1$ with residue $-1/EI$.

Let R be any positive number and let $\{z_i\}_1^N$ be the set of poles of U in $\{z: |z| \leq R\}$. The principal part of U at the pole z_i is a polynomial, $\sum_1^{d_i} c_{ij}(z - z_i)^{-j}$, in $(z - z_i)^{-1}$ of degree d_i, the multiplicity of the pole. Hence its Taylor coefficients are $\{p_i(k)z_i^{-k}\}_0^\infty$, where p_i is a polynomial of degree $d_i - 1$. If we subtract these principal parts from U, the remainder is analytic in $\{z: |z| \leq R\}$ whence its Taylor coefficients are $O(R^{-k})$. Consequently,

$$\mu_k = \sum p_i(k)z_i^{-k} + O(R^{-k}). \tag{4.7}$$

By Lemma 5′ we may choose $R > 1$ such that the only pole in the disc of radius R is 1. Hence (4.7) yields

$$\mu_k = \mu_\infty + O(R^{-k}), \tag{4.8}$$

where $\mu_\infty = 1/EI$ because of (4.6). By (3.9), (4.8) and induction on ℓ, $E\binom{L_k}{\ell}$ converges exponentially fast as $k \to \infty$ for every ℓ. Hence all moments of L_k converge. The method of moments, which is applicable because of Lemma 1, yields the existence of some L_∞ such that (1.4) and (1.5) holds.

Let $g_\infty(z)$ denote $E(z^{L_\infty})$. Since $g_k(z) \to g_\infty(z)$ for $|z| < 1$, (3.10) yields

$$\mu_\infty g_1(z) = \frac{1}{1-z}(z^m - g_\infty(z) + (1 - z^m)g_\infty(z)) = \frac{z^m}{1-z}(1 - g_\infty(z)),$$

whence

(4.9) $$g_\infty(z) = 1 - (1-z)(EI)^{-1}z^{-m}g_1(z).$$

A comparison with Lemma 2 completes the proof of the theorem.

5. Miscellaneous remarks.

1. It is (as remarked in the introduction) not known whether there exists any sequence $\{I_i\}$ that does not satisfy (*). However, if such a sequence exists, then the proof above shows that the conclusion of the theorem holds, provided g_1 is meromorphic in a sufficiently large region of the complex plane and $g_1(z) \neq 0$ when $|1 - z^m| = 1$, $z \neq 0$.

2. The largest allowed R in (1.5) is $\min\{|1 - z^m| : g_1(z) = 0, z \neq 0\}$, provided the corresponding zeroes of g_1 are simple. When the zeroes are multiple, any smaller R will do.

3. More detailed information is obtained by (4.7) for larger R. In particular, note that if there is a unique element with minimal modulus of $\{z \neq 1: z \text{ is a pole of } U\}$ and that element is positive, then $\{\mu_k\}$ is ultimately monotone, while $\{\mu_k\}$ oscillates if the element with minimal modulus is negative or if the elements with minimal modulus are two complex conjugates.

4. We may introduce generating functions $U_\ell(z) = \sum_1^\infty E\binom{L_k}{\ell}z^k$, obtain a recursion formula from (3.9) and conclude that each U_ℓ is a meromorphic function with (at most) the same poles as U. More detailed information on higher moments of L_k is obtained as above.

5. The definition of L_k depends only on $\{I_i\}_m^\infty$. Further, we may more generally assume that $\{I_i\}_{m+1}^\infty$ is stationary, while I_m may have a different distribution. If we let $\{\tilde{I}_i\}_m^\infty$ denote such a sequence with $\{\tilde{I}_i\}_{m+1}^\infty$ and $\{I_i\}_{m+1}^\infty$ equidistributed, and let $\tilde{L}_k$ etc. have the obvious meanings, a simple modification of Lemma 3 (with $\tilde{\mu}_k$, L_1, $\tilde{L}_k$ and $\tilde{L}_{k+1}$) holds, and the same proof as above shows that $\tilde{L}_k \to L_\infty$. The most important case is when $\{\tilde{I}_i\}$ has the conditional distribution of $\{I_i\}$ given that $I_0 = 0$. In that case $\tilde{g}_1 = g_\infty$ by Lemma 2. A modification of (3.10) shows that this is stationary, i.e. all $\tilde{L}_k$ are identically distributed.

6. Another generalization of Lemma 3 is

(5.1) $$\mu_k Ef(L_1, L_2) = E\left(\sum_m^{L_k - 1} f(j, L_{k+1}) + \sum_0^{m-1} f(L_{k+1} + j, L_{k+2})\right)$$

and corresponding formulae for functions of more than two variables. The proof is similar to the one given, using (in obvious notation) $L_1^{(n)}, L_2^{(n)}, \cdots$. This yields a recursion formula for mixed moments $E\binom{L_k}{\ell_1}\binom{L_{k+1}}{\ell_2}\ldots$ generalizing (3.9). It follows that if μ_k converges, then all mixed moments converge, whence the joint distribution of $\{L_{k+n}\}_{n=0}^{\infty}$ converges as $k \to \infty$.

We note the particular case $f(i, j) = j$ of (5.1)

$$\mu_k\mu_2 = EL_kL_{k+1} - m\mu_{k+1} + m\mu_{k+2}, \tag{5.2}$$

and its generalization obtained with $f(L_1, \cdots, L_{\ell+1}) = L_{\ell+1}$,

$$\mu_k\mu_{\ell+1} = EL_kL_{k+\ell} - m\mu_{k+\ell} + m\mu_{k+\ell+1}, \quad \ell = 1, 2, \cdots. \tag{5.3}$$

Furthermore, generating functions such as $Ez_1^{L_1}z_2^{L_2}$ may be expressed in g_1 using (5.1) appropriately. Moments and generating functions of N_k may be obtained by these methods.

7. That $EN_k = k/EI + O(1)$ follows also from more general m-dependent renewal theory, cf. Janson (1983), Theorem 3.1. However, Example 1 is an example where EL_k, and hence $EN_k - k/EI$, does not converge. Thus, the lattice case of Blackwell's renewal theorem does not extend to m-dependent variables. (Berbee (1979), Corollary 6.3.3 shows that the nonlattice case extends to even more general situations.)

6. The consequences of (*).

We denote the m-tuple $(\xi_{jm}, \cdots, \xi_{jm+m-1})$ by X_j. Thus $X_0, X_1, \cdots$ is a sequence of i.i.d. random variables and we let ν denote their common distribution (ν thus is the product of m copies of the distribution of ξ_i).

By (*), there exist functions α_i, $i = 0, \cdots, m - 1$ such that

$$I_{jm+i} = \alpha_i(X_{j-1}, X_j), \quad i = 0, \cdots, m-1, \quad j = 1, 2, \cdots. \tag{6.1}$$

We define, for $k = 0, \cdots, m - 1$,

$$\beta_k(x, y) = \textstyle\prod_0^k (1 - \alpha_i(x, y)) = 1 - \sum_0^k \alpha_i(x, y), \tag{6.2}$$

where the last equality holds since $\{I_i\}$ is m-separated and thus $\alpha_i\alpha_j = 0$ when $i \neq j$. Let $\beta = \beta_{m-1}$. Thus

$$\beta(X_{j-1}, X_j) = 1 \Leftrightarrow I_{jm}, \cdots, I_{jm+m-1} = 0. \tag{6.3}$$

PROOF OF LEMMA 4. If $j \geq 1$ and $0 \leq k \leq m - 1$,

$$\begin{aligned} L_1 = jm + k &\Leftrightarrow I_m = 0, \cdots, I_{jm+k-1} = 0 \quad \text{and} \quad I_{jm+k} = 1 \\ &\Leftrightarrow I_m = 0, \cdots, I_{jm-1} = 0 \quad \text{and} \quad I_{jm+k} = 1 \\ &\Leftrightarrow \beta(X_0, X_1) = 1, \cdots, \beta(X_{j-2}, X_{j-1}) = 1 \\ &\qquad \text{and} \quad \alpha_k(X_{j-1}, X_j) = 1. \end{aligned}$$

Hence

$$P(L_1 = jm + k) = E\beta(X_0, X_1)\cdot \ldots \cdot\beta(X_{j-2}, X_{j-1})\alpha_k(X_{j-1}, X_j). \tag{6.4}$$

Let T denote the integral operator with kernel β on $L^2(d\nu)$, i.e.

$$Tf(x) = \int \beta(x, y)f(y)\, d\nu(y) = E\beta(x, X_j)f(X_j),$$

and put $\tilde{\alpha}_k(x) = E\alpha_k(x, X_j)$. Then (6.4) may be written

$$P(L_1 = jm + k) = ET^{j-1}\tilde{\alpha}_k(X_0) = \langle T^{j-1}\tilde{\alpha}_k, 1\rangle. \tag{6.5}$$

Consequently, if $|z| < 1$,

$$\begin{aligned} g_1(z) &= \textstyle\sum_{n=m}^{\infty} P(L_1 = n)z^n = \sum_{k=0}^{m-1} \sum_{j=1}^{\infty} z^{jm+k}P(L_1 = jm + k) \\ &= \textstyle\sum_{k=0}^{m-1} z^k \sum_{j=1}^{\infty} z^{jm}\langle T^{j-1}\tilde{\alpha}_k, 1\rangle = \sum_{k=0}^{m-1} z^{k+m}\langle(1 - zT)^{-1}\tilde{\alpha}_k, 1\rangle. \end{aligned} \tag{6.6}$$

Since T, being a Hilbert-Schmidt operator, is compact, its resolvent $(\lambda - T)^{-1}$ is meromorphic for $\lambda \neq 0$, cf. Dunford and Schwartz (1958), Theorem VII.4.5, and the right-hand side of (6.6) defines a meromorphic function in the complex plane.

PROOF OF LEMMA 5. We will prove the equivalent Lemma 5′. Since $S_n = 0$ for $0 \le n < L_1$, $S_n = 1$ for $L_1 \le n < L_1 + L_2$, etc., it follows that

$$\textstyle\sum_0^{\infty} z^{S_n+1} = \sum_1^{\infty} L_k z^k, \quad |z| < 1. \tag{6.7}$$

Hence, if $|z| < 1$,

$$U(z) = E\textstyle\sum_1^{\infty} L_k z^k = Ez\sum_0^{\infty} z^{S_n} = z\sum_0^{\infty} Ez^{S_n}. \tag{6.8}$$

Consequently, if $|\zeta| = 1$,

$$\zeta \text{ is a pole of } U(z) \Leftrightarrow |\textstyle\sum_0^{\infty} Ez^{S_n}| \to \infty \quad \text{as} \quad z \to \zeta, \quad |z| < 1. \tag{6.9}$$

By (6.1) and (6.2), if $j \ge 1$ and $0 \le k \le m - 1$,

$$\begin{aligned} S_{jm+k} &= \textstyle\sum_1^{j-1} \sum_{im}^{im+m-1} I_n + \sum_{jm}^{jm+k} I_n \\ &= \textstyle\sum_1^{j-1} (1 - \beta(X_{i-1}, X_i)) + 1 - \beta_k(X_{j-1}, X_j). \end{aligned} \tag{6.10}$$

Since β assumes the values 0 and 1, $z^{1-\beta} = z + (1 - z)\beta$ and hence

$$z^{S_{jm+k}} = \textstyle\prod_1^{j-1} (z + (1 - z)\beta(X_{i-1}, X_i))\cdot(z + (1 - z)\beta_k(X_{j-1}, X_j)). \tag{6.11}$$

Consequently, if T_z denotes the integral operator with kernel $K_z(x, y) = z + (1 - z)\beta(x, y)$ and $\tilde{\beta}_k(x) = E\beta_k(x, X_j)$,

$$Ez^{S_{jm+k}} = \langle T_z^{j-1}(z + (1 - z)\tilde{\beta}_k), 1\rangle \tag{6.12}$$

and, since $|z + (1 - z)\tilde{\beta}_k| \le \max(|z|, 1)$,

$$|Ez^{S_{jm+k}}| \le \| T_z^{j-1} \| \, \| z + (1 - z)\tilde{\beta}_k \| \le \| T_z^{j-1} \|, \quad |z| \le 1. \tag{6.13}$$

Now, suppose that $|\zeta| = 1$ and that the spectral radius of T_ζ is strictly less than 1. Since $z \to T_z$ is continuous, it follows that there exist $\varepsilon > 0$, $R < 1$ and $C < \infty$ such that

$$\| T_z^j \| \le CR^j \quad \text{if} \quad |z - \zeta| < \varepsilon \quad \text{and} \quad j = 0, 1, \cdots. \tag{6.14}$$

Hence, using (6.13), if $|z - \zeta| < \varepsilon$ and $|z| \le 1$, then

$$|\textstyle\sum_0^\infty Ez^{S_n}| \le \sum_0^\infty |Ez^{S_n}| \le \sum_0^\infty CR^{n/m-2} < \infty. \tag{6.15}$$

Consequently, by (6.9), ζ is a regular point of U.

For the rest of the proof we suppose, on the contrary, that $\zeta \neq 1$ is a pole of U with $|\zeta| = 1$. By the argument above, T_ζ has spectral radius at least 1. Thus, there exists an eigenvalue λ with $|\lambda| \ge 1$ and an eigenfunction $\varphi \in L^2(d\nu)$ with $\|\varphi\| = 1$ such that $T_\zeta\varphi = \lambda\varphi$. Choose an orthonormal basis $\{\varphi_i\}_1^\infty$ in $L^2(d\nu)$ with $\varphi_1 = \varphi$. For the Hilbert-Schmidt norm of T_ζ we have the two expressions

$$\|T_\zeta\|_{HS}^2 = \int\int |K_\zeta(x, y)|^2\, d\nu(x)\, d\nu(y) = 1 \tag{6.16}$$

and

$$\|T_\zeta\|_{HS}^2 = \textstyle\sum_{i=1}^\infty \|T_\zeta\varphi_i\|^2 = |\lambda|^2 + \sum_2^\infty \|T_\zeta\varphi_i\|^2. \tag{6.17}$$

Hence, $|\lambda| = 1$ and $T_\zeta\varphi_i = 0$ when $i \neq 1$. Consequently, $T_\zeta\psi = \lambda\varphi\langle\psi, \varphi\rangle$ for ψ belonging to the basis, and thus for every $\psi \in L^2(d\nu)$. Thus, $T_\zeta\psi(x) = \lambda\varphi(x)\int \psi(y)\overline{\varphi(y)}\, d\nu(y) = \int \lambda\varphi(x)\overline{\varphi(y)}\psi(y)\, d\nu(y)$, whence

$$K_\zeta(x, y) = \lambda\varphi(x)\overline{\varphi(y)} \quad \text{a.s.} \tag{6.18}$$

By definition, K_ζ assumes only the values 1 and ζ. Thus, the product of the two independent random variables $\varphi(X_1)$ and $\lambda\overline{\varphi(X_2)}$ assumes only two different values, both nonzero, whence $\varphi(X_1)$ is a discrete random variable assuming at most two different values.

There are two cases:

(i) $\varphi(X_1)$ is a.s. constant. By (6.18), $K_\zeta(x, y)$ is a.s. constant, i.e. $\beta(x, y)$ is a.s. constant. We have two subcases: $\beta = 0$ and $\beta = 1$. If $\beta = 1$, then by (6.3), $I_{jm} = 0$ a.s., which is a contradiction. If $\beta = 0$, (6.3) yields $\sum_{jm}^{jm+m-1} I_i = 1$ a.s. Since $\{I_i\}$ is stationary, also $\sum_{jm+1}^{jm+m} I_i = 1$ a.s. Consequently, $I_{jm} = I_{jm+m}$, whence $I_m = I_{2m} = I_{3m}$, which is a contradiction because I_m and I_{3m} are independent. Thus, both subcases lead to contradictions and we turn to the second case.

(ii) φ assumes two values γ_1 and γ_2 with positive probabilities. By (6.18), $|\varphi(X_1)||\varphi(X_2)| = |K_\zeta(X_1, X_2)| = 1$ a.s., whence $|\gamma_1| = |\gamma_2| = 1$. Replacing φ by $\bar{\gamma}_1\varphi$ we may assume that $\gamma_1 = 1$. (6.18) yields

$$\lambda = \lambda\cdot 1\cdot 1 = 1 \quad \text{or} \quad \zeta, \tag{6.19a}$$

$$\lambda\gamma_2 = \lambda\cdot\gamma_2\cdot 1 = 1 \quad \text{or} \quad \zeta, \tag{6.19b}$$

$$\lambda\bar{\gamma}_2 = \lambda\cdot 1\cdot\bar{\gamma}_2 = 1 \quad \text{or} \quad \zeta. \tag{6.19c}$$

Thus, at least two of the three values λ, $\lambda\gamma_2$ and $\lambda\bar{\gamma}_2$ coincide. Since $\lambda \neq 0$ and $\gamma_2 \neq \gamma_1 = 1$, the only possibility is $\gamma_2 = \bar{\gamma}_2$, whence $\gamma_2 = -1$. (Thus $\varphi(x) = \pm 1$.) Hence both λ and $-\lambda$ equal 1 or ζ, and we conclude that $\zeta = -1$ and $\lambda = \pm 1$. It follows from (6.18) and the definition of K_ζ that

$$\text{If} \quad \lambda = 1, \quad \beta(x, y) = I(\varphi(x) = \varphi(y)) \tag{6.20$'$}$$

(6.20″) If $\lambda = -1$, $\beta(x, y) = I(\varphi(x) \neq \varphi(y))$.

Recall that $X_0 = (\xi_0, \cdots, \xi_{m-1})$ and put $\varphi_0(\xi_0) = E(\varphi(X_0) \mid \xi_0)$ and $E = \{\xi_0: -1 < \varphi_0(\xi_0) < 1\}$. Suppose that $P(E) > 0$. If $\xi_0, \cdots, \xi_m$ are given with $\xi_m \in E$, we may choose $\xi_{m+1}, \cdots, \xi_{2m-1}$ such that $\varphi(\xi_m, \cdots, \xi_{2m-1})$ is any of the two possible values ± 1. In particular, we may by (6.20) choose them such that $\beta((\xi_0, \cdots, \xi_{m-1}), (\xi_m, \cdots, \xi_{2m-1})) = 1$. By (6.3), $\alpha(\xi_0, \cdots, \xi_m) = I_{jm} = 0$. Hence:

(6.21) If $\xi_m \in E$, then $\alpha(\xi_0, \cdots, \xi_m) = 0$.

Now, let $X_1 = (\xi_m, \cdots, \xi_{2m-1})$ be such that each $\xi_{m+k} \in E$, $0 \leq k \leq m-1$. For any X_0, (6.21) implies that
$I_{jm} = \alpha(\xi_0, \cdots, \xi_m) = 0$, $I_{jm+1} = \alpha(\xi_1, \cdots, \xi_{m+1}) = 0, \cdots, I_{jm+m-1} = 0$. Thus, by (6.3), $\beta(X_0, X_1) = 1$ for any X_0. This contradicts (6.20) and we are forced to conclude that $P(E) = 0$, i.e. $\varphi_0(\zeta_0) = \pm 1$ a.s. This proves that $\varphi(X_0) = \varphi_0(\zeta_0)$ a.s., i.e. $\varphi(\xi_0, \cdots, \xi_{m-1})$ depends only on the first coordinate.

A mirror image of the argument above shows that $\varphi(\xi_0, \cdots, \xi_{m-1})$ depends only on the last coordinate ξ_{m-1}. If $m \geq 2$, this yields a contradiction. Thus $m = 1$, which implies that $X_j = \xi_j$ and, by (6.3), $I_j = 1 - \beta(\xi_{j-1}, \xi_j)$. By (6.20), either $I_j = I(\varphi(\xi_{j-1}) \neq \varphi(\xi_j))$ or $I_j = I(\varphi(\xi_{j-1}) = \varphi(\xi_j))$. Replacing ξ_j by $\frac{1}{2}(1 + \varphi(\xi_j))$, we see that these are exactly Examples 1 and 2 in the next section. We will show, by explicit computations, that, in fact, -1 is a regular point of U in Example 2. Thus, the only remaining possibility is Example 1, and the proof is completed.

7. Examples. All examples that follow satisfy (*). We begin with the exceptional case, in which the non-zero I_i occur in pairs.

1. Let $\{\xi_i\}$ be i.i.d. Bernoulli distributed variables with $P(\xi_i = 1) = p$ and $P(\xi_i = 0) = 1 - p = q$, where $0 < p < 1$, and let $I_i = I(\xi_{i-1} \neq \xi_i)$. Thus $m = 1$. If $p = \frac{1}{2}$, $\{I_i\}$ has the same distribution as $\{\xi_i\}$ and L_k are i.i.d. geometrical random variables. We exclude this case in the sequel.

$$P(L_1 = n) = P(\xi_0 = \xi_1 = \cdots = \xi_{n-1} \neq \xi_n) = p^n q + q^n p,$$

whence

$$g_1(z) = \frac{pqz}{1 - pz} + \frac{pqz}{1 - qz} = pqz \frac{2 - z}{(1 - pz)(1 - qz)}.$$

Thus $g_1(2) = 0$. By (4.5),

$$U(z) = \frac{1}{g_1(1 - z)} - 1 = \frac{z((1/pq) - 2) + 2z^2}{1 - z^2}. \tag{7.1}$$

Hence, $\mu_{2k-1} = 1/pq - 2$ and $\mu_{2k} = 2$, $k = 1, 2, \cdots$. Repeated use of (3.11) shows that

$$g_2(z) = z \frac{1 - 2pq - (1 - 3pq)z}{(1 - pz)(1 - qz)} \tag{7.2}$$

and $g_3(z) = g_1(z)$, $g_4(z) = g_2(z)$, etc. Hence, $\{L_{2k-1}\}_1^\infty$ and $\{L_{2k}\}_1^\infty$ are two sets of identically distributed random variables, but the two distributions differ. Consequently, L_k does not converge.

Another way to see this is to note that the conditional distribution of L_{k+1} given ξ_{L_k} is geometric with parameter $P(\xi = \xi_{L_k})$. However, S_n is even $\Leftrightarrow \xi_n = \xi_0$; hence $P(\xi_{L_k} = 1) = p$ when k is even but q when k is odd. Note also that $E(-1)^{S_n} = P(\xi_n = \xi_0) - P(\xi_n \neq \xi_0) = (p-q)^2 > 0$, i.e. S_n is even more often than odd.

2. Let $\{\xi_i\}$ be as in Example 1 but let instead $I_i = I(\xi_{i-1} = \xi_i)$. Again $m = 1$. Then $P(L_1 = n) = P(\zeta_0 \neq \cdots \neq \zeta_{n-1} = \zeta_n)$, whence

$$P(L_1 = 2k) = p^{k+1}q^k + p^k q^{k+1} = (pq)^k,$$

$$P(L_1 = 2k-1) = p^{k+1}q^{k-1} + p^{k-1}q^{k+1} = (pq)^{k-1}(1-2pq),$$

and

$$g_1(z) = \frac{z(1-2pq) + pqz^2}{1 - pqz^2}. \tag{7.3}$$

Thus $g_1(z) = 0 \Leftrightarrow z = 0$ or $z = 2 - 1/pq < -2$ (for $p \neq ½$). Hence U is regular on $\{z: |z| = 1, z \neq 1\}$ and $L_k \to_d L_\infty$. By (4.5),

$$U(z) = \frac{z}{1-2pq}\left(\frac{1}{1-z} + \frac{(1-4pq)pq}{1 - pq - pqz}\right), \tag{7.4}$$

whence

$$\mu_k = \frac{1}{1-2pq} + \frac{1-4pq}{1-2pq}\left(\frac{1}{pq} - 1\right)^{-k}. \tag{7.5}$$

3. Let $\{\xi_i\}$ be i.i.d. $U(0, 1)$ random variables and let $I_i = I(\xi_{i-1} > \xi_i)$. Thus $m = 1$. L_k may be interpreted as the length of the kth run in a random very long permutation.

This has been studied by several authors, cf. for instance Barton and Mallows (1965), giving inter alia the recursion formula (3.11), and Pittel (1980, 1981) and Esseen (1982). The last two references prove convergence results by different methods.

Our theorem immediately yields convergence in distribution and of moments. Explicit results are obtained from the easily verified relations, cf. the references given above,

$$P(L_1 = n) = P(\xi_0 < \cdots < \xi_{n-1} > \xi_n) = n/(n+1)!, \tag{7.6}$$

$$g_1(z) = \frac{(z-1)e^z + 1}{z}, \tag{7.7}$$

and, by (4.5),

$$U(z) = \frac{z(1-z)}{e^{z-1} - z} - z. \tag{7.8}$$

The poles of $U(z)$, apart from $z = 1$, closest to the origin are $z \approx 3.09 \pm 7.46i$, with $|z| \approx 8.08$. Hence, using $EI = 1/2$, $\mu_k = 2 + o(8^{-k})$, and, since the poles are complex, the means μ_k oscillate about 2. See Knuth (1973), page 39–46 for further details.

For higher moments we e.g. obtain from (3.7) and (5.2)

$$EL_k^2 = 2(\mu_k\mu_1 - \mu_{k+1}) + \mu_k = 4\mu_1 - 2 + o(8^{-k}), \tag{7.9}$$

$$EL_kL_{k+1} = \mu_k\mu_2 + \mu_{k+1} - \mu_{k+2} = 2\mu_2 + o(8^{-k}), \tag{7.10}$$

whence, as $k \to \infty$,

$$\operatorname{Var} L_k \to 4\mu_1 - 6 = 4e - 10 \approx 0.873 \tag{7.11}$$

$$\operatorname{Cov}(L_k, L_{k+1}) \to 2\mu_2 - 4 = 2e^2 - 4e - 4 \approx -0.095. \tag{7.12}$$

4. Let $\{\xi_i\}$ be as in Example 3 and let $I_i = I(\xi_{i-2} < \xi_{i-1} > \xi_i)$. Thus $m = 2$. $\{L_k\}$ is the process of wavelengths between successive peaks in a random permutation. Esseen (1982, 1983) has proved convergence theorems using Markov chain methods. Our theorem furnishes an alternative proof.

In this case, see Esseen (1983) for details,

$$g_1(z) = \frac{1 + z - (1 - z)e^{2z}}{2z}. \tag{7.13}$$

It follows that $g_1(z) = 0 \Leftrightarrow \tanh z = z$, whence the set of poles of U is $\{1 - z^2: g_1(z) = 0\} = \{1 + x^2: \tan x = x\}$. Thus, the poles are positive and the smallest one, after 1, is ≈ 21.19. Furthermore, the residues are negative whence μ_k decreases monotonically to $1/EI = 3$. By (4.4),

$$U(z) = \frac{z}{\sqrt{1 - z}\coth\sqrt{1 - z} - 1}. \tag{7.14}$$

5. Let $m \geq 1$, let $\{\xi_i\}$ be as in Example 3 and let

$$I_i = I(\xi_{i-m} < \cdots < \xi_{i-1} > \xi_i).$$

(Thus Examples 3 and 4 are the cases $m = 1$ and 2.) The variables L_k are the distances between the ends of the increasing runs of length at least m in a random permutation. The theorem shows that $L_k \to_d L_\infty$ and

$$EL_k \to EL_\infty = 1/EI = (m + 1)!/m = (m + 1)\cdot(m - 1)!$$

6. Let $\{\xi_i\}$ be i.i.d. and let $I_i = I(\xi_{i-1} + \xi_i \leq A)$, where A is a fixed number. Thus $m = 1$. Again, the theorem shows that $\{L_k\}$ converges. As an application, consider a Poisson process with constant intensity λ on $\{t: t \geq 0\}$ and let $\{\xi_i\}$ be the intervals between the points of the process. Then ξ_i are independent and exponentially distributed, and $I_i = 1$ as soon as three points are clustered within an interval of length A. Thus L_k counts the number of observed points between such clusters. We obtain $EL_k \to EI^{-1} = (1 - e^{-\lambda A} - \lambda Ae^{-\lambda A})^{-1}$.

REFERENCES

BARTON, D. E. and MALLOWS, C. L. (1965). Some aspects of the random sequence. *Ann. Math. Statist.* **36** 236–260.

BERBEE, H. C. P. (1979). Random walks with stationary increments and renewal theory. *Mathematical Centre Tract* 112. Amsterdam.

DUNFORD, N. and SCHWARTZ, J. (1958). *Linear Operators I.* Interscience, New York.

ESSEEN, C.-G. (1982). On the application of the theory of probability to two combinatorial problems involving permutations. *Proceedings of the Seventh Conference on Probability Theory.* Braşov.

ESSEEN, C.-G. (1983). Unpublished lectures.

JANSON, S. (1983). Renewal theory for m-dependent variables. *Ann. Probab.* **11** 558–568.

KNUTH, D. E. (1973). *The Art of Computer Programming, Vol.* 3, *Sorting and Searching.* Addison-Wesley, Reading.

PITTEL, B. G. (1980). A process of runs and its convergence to the Brownian motion. *Stochastic Process. Appl.* **10** 33–48.

PITTEL, B. G. (1981). Limiting behaviour of a process of runs. *Ann. Probab.* **9** 119–129.

UPPSALA UNIVERSITY
DEPARTMENT OF MATHEMATICS
THUNBERGSVÄGEN 3
S-752 38 UPPSALA, SWEDEN