MATH. SCAND. 3 (1955), 127–149

TRANSFORMATIONS OF STATIONARY RANDOM SEQUENCES

EDWARD L. KAPLAN

Introduction. This paper is concerned with stationary random processes whose elements or realizations consist of sequences of points or events on a line. The theory is simplified by defining the sequences as infinite in both directions, although in practice only a finite portion of the sequence can be observed. The Poisson process is the best-known process of this type, and has applications in which the random events may be the placing of calls by telephone subscribers, failures of vacuum tubes, radio-active disintegrations, or many other possibilities. Possibly the next simplest example is the renewal process, which for example describes the sequence of failures of electric lamps in a given socket, when the illumination is provided continuously and the lifetimes of the lamps are independently distributed with the same (arbitrary) distribution. The same kind of process occurs in the theory of queues. Examples of still more complicated processes are given by the zeros or maxima of random Gaussian noise functions; the arrival of scheduled airplanes at an airport, when they are intended to arrive at uniform intervals of time; and the events of meteorology, geology, biology, etc.

All the above examples involve positions in time. But one may also be concerned with the spacial positions of objects such as organisms, colloidal particles, vehicles, or wavecrests. Suppose now that these objects are in motion, subject to statistical laws, and that their positions are observed not continuously but only at selected instants, as by successive photographic exposures. If the objects do not differ sufficiently to permit them to be distinguished by their appearance, one is confronted with the question of how well one can preserve the identities of the objects from one exposure to another on the basis of the observed positions alone. This is the motivation for the study of permutations of the points of the sequence in Sections 3 and 4 below; the subject is considered in much greater detail in the author's doctoral dissertation [6].

Received May 10, 1955.

Some of the results of this paper are taken from the author's doctoral dissertation, Princeton, 1950, which was supported by the office of Naval Research.

Let $\{u_i\}$, $i = \ldots, -2, -1, 0, 1, 2, \ldots$, be an infinite collection of real-valued random variables with the following properties:

(1 a) Stationarity: The transformation $\bar{u}_i = u_{i+1}$ is measure-preserving;
(1 b) Non-negativity: $\Pr\{u_i < 0\} = 0$;
(1 c) Non-degeneracy: $\Pr\{u_i = 0, \text{ all } i\} = 0$.

In practical applications, the stationarity is likely to be a simplifying approximation which is valid for a finite period of time, or a finite number of u_i.

In this study the u_i are interpreted as the successive interval-lengths between consecutive points v_{i-1}, v_i on a line, hence the conditions (1 b) and (1 c). Thus for $n > 0$

(1 d) $$v_n = v_0 + \sum_{i=1}^{n} u_i \qquad \text{and} \qquad v_{-n} = v_0 - \sum_{i=0}^{n-1} u_{-i} .$$

The term "sequence" refers primarily to $\{v_i\}$, but one cannot very well avoid using it also for $\{u_i\}$. Except in Section 2, v_0 is set equal to zero. Except where the Poisson process is considered, it is nowhere required in this paper that the u_i be mutually independent, although most of the existing literature is restricted to this case (the renewal process).

In Section 1 the principal result is that if $v_0 = 0$, the expected number of points v_i contained in a finite interval does not exceed that for an interval twice as long, centered on the origin. A by-product of the proof is the fact that if x and y are independent random variables having the same arbitrary distribution, then

$$\Pr\{c' \leqq x - y \leqq c + c'\} \leqq \Pr\{-c \leqq x - y \leqq c\}$$

for all c and c'.

Section 2 considers the transformation of the ensemble of sequences $\{u_i\}$ or $\{v_i \mid v_0 = 0\}$ into an ensemble $\{v_i\}$ stationary with respect to a *continuous*-parameter group of transformations, and vice versa. The correspondence is not perfect; certain ensembles of each of the two types cannot be so transformed. The transformation involves the introduction of v_0 as a random variable in addition to and dependent on the u_i, and generalizes a theorem proved by J. L. Doob [2] for renewal processes. The idea is extended further in Theorem 2.C. to construct interesting ensembles of both types from much simpler ensembles of patterns of unlabeled points.

In Sections 3 and 4 a second sequence of random variables x_i is adjoined. For many of the theorems they are only required to be jointly stationary with the u_i; that is, the transformation

$$\{\bar{u}_i, \bar{x}_i\} = \{u_{i+1}, x_{i+1}\}$$

is measure-preserving. The x_i are interpreted as displacements applied to the v_i, so that the given ensemble is mapped into an ensemble of sequences $\{V_i\}$ with $V_i = v_i + x_i$. It is shown that the V_i have zero probability of having a cluster point. Thus they can be relabeled as V_j' with $V_j' \leqq V_{j+1}'$. It now follows easily that the ensemble $\{V_j'\}$ has the same general properties that have been assumed for $\{v_i\}$.

The relabeling just mentioned defines an infinite permutation $i = \pi(j)$, which is a random 1–1 function from integers to integers. Section 4 is devoted to the investigation of integer-valued random variables (numbers of inversions) defined solely in terms of the π's. The existence and finiteness of Ex_i is found to be sufficient for the finiteness of the number of inversions associated with each point. The number of inversions may be of interest for applications, because one of the most obvious ways of identifying the images of moving objects (in a linear array) in one exposure with the images of the same objects in another exposure is simply to assume that no inversions (changes of order) have occurred. This is equivalent to assuming that $\pi(j) = j + \pi_0$, where the integer π_0 remains to be determined.

1. Stationary random sequences. If U denotes a sample sequence $\{u_i\}$ from the process, an important function of U is

$$\lambda(U) = \lim_{n\to\infty} \sum_{i=1}^{n} u_i/n = \lim_{n\to\infty} \sum_{i=1}^{n} u_{-i}/n \,, \tag{1 e}$$

which is the average distance between consecutive points. By the ergodic theorem (Hopf [5, p. 49 ff.]) and (1 b), $\lambda(U)$ is defined almost everywhere if the value $+\infty$ is included. It is also part of this theorem that the expectation $E\lambda(U) = Eu_i$.

THEOREM 1.A. *The probability is zero that* $\lambda(U) = 0$, *and hence, so is the probability that the sequence* $\{v_i\}$ *have a cluster point* (*other than at infinity*).

PROOF. For a sequence of intervals u_i leading to a cluster point one would have

$$\lambda(U) = \lim_{n\to\infty} \sum_{i=1}^{n} u_i/n = 0 \,,$$

since $\Sigma_1^n u_i$ would be bounded. Let Z be the set of sequences U for which $\lambda(U) = 0$, and E_Z the conditional expectation within Z, supposed for the moment to have positive probability. Then one has $E_Z\lambda(U) = 0$, but since

Z is an invariant set, $E_Z\lambda(U)=E_Z u_i$ for all i. Thus, since $u_i \geqq 0$, all the u_i vanish almost certainly in Z, which must therefore have zero probability by (1 c).

While the number of points v_i contained in a finite interval is thus almost always finite, the applications of 2.C. below will show that the expected number of such points may be infinite. Some sufficient conditions for finiteness of the expected number, and related results, will now be given.

THEOREM 1.B. *The expected number of points v_i contained in a finite fixed interval does not exceed that for an interval twice as long, centered on $v_0=0$. (The latter interval is to be closed if the first interval is closed; otherwise it may be open.) Thus the expected number $\varphi(v)$ in $(0, v)$, if finite, is at most of order v for v large; if in addition $E(1/\lambda(U))>0$, it is exactly of order v.*

PROOF. Let $(c', c+c')$ be the given interval; we may assume $c>0$ and $c'>0$, since the expected number in $(-c-c', -c')$ is the same. The proof depends on the obvious fact that if A denotes the set of points from the sequence v_i falling in $(c', c+c')$, then all these points fall in $(-c, c)$ whenever the origin of abscissas is shifted to any one of these same points A. For the sake of definiteness, all intervals will be regarded as closed.

Let μ_i, $i=\ldots, -2, -1, 0, 1, 2, \ldots$, denote the number of points falling in $(c', c+c')$ when the origin is shifted from $v_0=0$ to v_i, and let μ_i^0 denote the same for $(-c, c)$. We have to show that $E\mu_0^0 \geqq E\mu_0$, which is the same as

$$E \lim \sum_{i=-1}^{-n} \mu_i^0/n \geqq E \lim \sum_{i=-1}^{-n} \mu_i/n .$$

So it suffices to show that

$$\lim \sum_{i=-1}^{-n} \mu_i^0/n \geqq \lim \sum_{i=-1}^{-n} \mu_i/n$$

for almost all sequences.

Now let $M(t_1, t_2)$ denote the number of different ordered pairs (v, v') of points of the sequence in (t_1, t_2) such that $v'-v$ falls in the interval $(c', c+c')$, and $M_0(t_1, t_2)$ the number such that $v'-v$ falls in $(-c, c)$. Pairs for which $v \geqq v'$ must be included in $M_0(t_1, t_2)$, and multiple points treated as if distinct. Then we have

$$M(v_{-n}, v_{-1})/n \leqq \sum_{i=-1}^{-n} \mu_i/n \leqq M(v_{-n}, c+c')/n .$$

When $n \to \infty$, the two extremes approach equality since $M(v_{-n}, c+c') - M(v_{-n}, v_{-1})$ remains constant (and finite) as soon as $v_{-1} - v_{-n} \geqq c+c'$. Hence

$$\lim \sum_{i=-1}^{-n} \mu_i/n = \lim M(v_{-n}, v_{-1})/n ,$$

and similarly

$$\lim \sum_{i=-1}^{-n} \mu_i^0/n = \lim M_0(v_{-n}, v_{-1})/n .$$

So it will suffice to show that

$$M_0(v_{-n}, v_{-1}) \geqq M(v_{-n}, v_{-1})$$

for all n and all sequences v_i.

The set of point-pairs (v, v') whose (finite) cardinality defines $M(v_{-n}, v_{-1})$ can be represented as a set C of distinct positions in an $n \times n$ lattice or matrix, in which the row identifies the first point v and the column the second point v'. Similarly we may define a set C_0, symmetric about the principal diagonal of the matrix, whose cardinality is $M_0(v_{-n}, v_{-1})$. Thus we have to show that the cardinality of C_0 is as great or greater than that of C.

By the remark made at the start of the proof, the set C_0 contains the two sets C' and C'' derived from C by the following rules:

$$(v, v') \in C \quad \text{implies} \quad (v, v) \in C' \text{ and } (v', v') \in C'' .$$

$$(v', v) \text{ and } (v'', v) \in C \quad \text{implies} \quad (v', v'') \text{ and } (v'', v') \in C' .$$

$$(v, v') \text{ and } (v, v'') \in C \quad \text{implies} \quad (v', v'') \text{ and } (v'', v') \in C'' .$$

For example, let us take $n=4$, with $-v_{-i} = 3, 5, 6, 9$ for $i=1$ to 4, and $(c', c+c') = (2, 3)$ and $(-c, c) = (-1, 1)$. Then the arrays in question take the forms

	3	5	6	9
3	0	x	x	
5		0	0	
6		0	0	x
9				0

$0 \in C_0$, $x \in C$.

	3	5	6	9
3	1			
5		2	2	
6		2	1, 2	
9				2

$1 \in C'$, $2 \in C''$.

The arrays C_0 (zeros) and C(x's) are non-overlapping because this is true of the intervals $(-1, 1)$ and $(2, 3)$, but the arrays C' (ones) and

C'' (twos) do overlap in one position. However, this will be immaterial if we can show that the sum of the cardinalities of C' and C'' equals or exceeds *twice* the cardinality of C. To do this, choose any element h of C. Let the row and column of h contain m' and m'' elements of C respectively. Then if this row and column are deleted from C, its cardinality is decreased by $m'+m''-1$. The cardinalities of C' and C'' are decreased by at least $2(m'-1)+1$ and $2(m''-1)+1$ respectively, thus $2(m'+m''-1)$ for the sum. By repeating this process, all cardinalities can be reduced to zero. This proves the first part of the theorem.

For the second part of the theorem we note that

$$\lim_{n\to\infty} v_n/n = \lambda(U) \qquad \text{is equivalent to} \qquad \lim_{v\to\infty} N^\star(v)/v = 1/\lambda(U)\,,$$

where the random variable $N^\star(v)$ is the number of points falling in the open interval $(0, v)$. Then $\varphi(v) = EN^\star(v)$, and applying Fatou's lemma gives

$$\liminf_{v\to\infty} \varphi(v)/v \geqq E\bigl(1/\lambda(U)\bigr)$$

as desired. The condition $E\bigl(1/\lambda(U)\bigr) > 0$ is of course equivalent to having a positive probability that $\lambda(U)$ be finite.

The bound given by the theorem may actually be attained. As an example, let the u_i have the pattern

$$\ldots,\ 0,\ 4,\ 1,\ 4,\ 0,\ 4,\ 1,\ 4,\ \ldots,$$

so that the v_i consist of double points at intervals of 9 units, with pairs of distinct points 1 unit apart, centered in the intervening spaces. Then whatever point of the sequence is taken as the origin, the closed interval $(4, 5)$ always contains just two points; the interval $(-1, 1)$, but no subinterval thereof, has the same property. To be sure, if the u_i are mutually independent, the interval $(-c, c)$ can obviously be replaced by the closed interval $(0, c)$. (Translate the origin to the first point of the sequence contained in $(c', c+c')$.) Also, if $c' > 0$, one can show that the arrays C' and C'' cannot be identical, and hence the arrays C and C_0 cannot have equal cardinalities, although the ratio of these cardinalities may approach unity as $n \to \infty$.

THEOREM 1.C. *If x and y are independent random variables having the same arbitrary distribution, then*

$$\Pr\{c' \leqq x-y \leqq c+c') \leqq \Pr\{-c \leqq x-y \leqq c\}\,.$$

In the course of proving the preceding theorem, this result has in effect been established for the case in which the common distribution consists of discrete probabilities $1/n$ assigned to n numbers, some of which may be equal. Distributions of this sort with $n \to \infty$ give arbitrarily close approximations to the true probabilities in the general case.

THEOREM 1.D. *The expected number of points of the sequence contained in any finite interval is finite if any one of the following conditions is satisfied:*

(a) *the expected number is finite for some interval about the origin;*

(b) $Ev_m^{-2} < \infty$, *where* $v_m = u_1 + u_2 + \ldots + u_m$, *for some* m *(a special case is* $v_m \geqq a > 0$ *with probability one).*

(c) u_i *is independent of* $u_{i\pm m}, u_{i\pm(m+1)}, \ldots$, *for some* m.

Detailed proofs will be omitted. Condition (a) follows from 1.B., condition (b) from the relation

$$\Pr\{u_1 + \ldots + u_n < c\} \leqq n \Pr\{u_1 < c/n\},$$

and (c) from the relation

$$\Pr\{u_1 + \ldots + u_n < c\} \leqq \Pr\{u_i < c, \quad \text{for} \quad 1 \leqq i \leqq n\}.$$

2. Change of parameter. Here the question concerns the continuity or the discreteness of the parameter of the group of transformations that are required to be measure-preserving. In case v_0 is set equal to zero, only the transformations

$$T_{(n)}: \quad \bar{v}_i = v_{i+n} - v_n$$

can be considered, and the parameter is discrete; but by admitting v_0 as a random variable satisfying $0 \leqq v_0 < u_0$ one can admit the transformations

$$T_c: \quad \bar{v}_i = v_{i+m} - c,$$

where the parameter c is any real number, and m the least integer such that $v_m \geqq c$.

The intuitively obvious way to construct the process with continuous parameter from that with discrete parameter is to use the latter in specifying the relative positions of the points, and then to choose the origin (or equivalently, the value of v_0) at random, with uniform probability, in an interval whose length approaches infinity. A modification of this procedure is given in the following theorem, which generalizes Theorem 4 of Doob [2]. The interval-lengths u_i associated with the discrete-parameter process are replaced by u_i^* when the parameter is continuous, to reflect the change in their distributions. The letter I denotes finite sets of integers (subscripts) that include the integer zero. The common expectation of the u_i is denoted by Eu.

THEOREM 2.A. *Given the stationary discrete-parameter process $\{u_i\}$, with $Eu < \infty$, let the distribution of the variables $v_0, u_0^\star, u_1^\star, \ldots$ be determined by*

$$\Pr\{0 \leqq v_0 < u_0^\star\} = 1 ,$$

$$\Pr\{v_0 \leqq b_0;\ u_i^\star \geqq a_i,\ i \in I\} = \Pr\{u_i \geqq a_i,\ i \in I\}\, b_0 / Eu$$
$$(b_0 \leqq a_0;\ \text{all } I) .$$

Then the new process $\{v_0, u_i^\star\}$ is stationary with respect to the continuous-parameter group of transformations

$$T_c:\ \bar{v}_i = v_{i+m} - c ,$$

where m is the least integer such that $v_m \geqq c$, and $v_i - v_{i-1} = u_i^\star$.

PROOF. In a more suggestive notation, if the variables u_i, $i \in I$, have the distribution $dL_I(u_I)$, u_I being a vector, then the new variables are to have the distribution

$$dL_I(u_I^\star)\, dv_0 / Eu ,$$

subject to the condition $0 \leqq v_0 < u_0^\star$. The various distributions obtained by varying I are obviously mutually consistent and so define a stochastic process by the theorem first proved by Kolmogoroff. The $u_i^\star$ alone have the distribution

$$u_0^\star\, dL_I(u_I^\star) / Eu ,$$

which differs from that of the u_i. The variable v_0 has the marginal distribution

$$[1 - L(v_0)]\, dv_0 / Eu ,$$

where L is the common (cumulative) distribution function of any one of the u_i. The Poisson process gives a convenient illustration of these facts. Here v_0 happens to have the same distribution

$$e^{-t/Eu}\, dt / Eu$$

as every $u_i^\star$ $(i \neq 0)$ and every u_i, while $u_0^\star$ has the distribution

$$e^{-t/Eu}\, t\, dt / (Eu)^2 ,$$

so that $Eu_0^\star = 2Eu$. This familiar "paradox" reflects the fact that when the origin of abscissas is located at random on the line, it is more likely to fall in a long interval $(u_0^\star)$ than in a short one. It also reflects the fact that the behavior of the rest of the process is unaffected when conditioned by the occurrence of an event (i.e., a point) in an infinitesimal interval about the origin.

It remains to demonstrate the stationarity. The transformation T_c is equivalent to

$$\bar{u}_i^\star = u_{i+m}^\star ,$$

$$\bar{v}_0 = \begin{cases} v_0 - c & \text{if} \quad m = 0 , \\ v_0 + u_1^\star + u_2^\star + \ldots + u_m^\star - c & \text{if} \quad c > 0,\ m > 0 , \\ v_0 - u_0^\star - u_{-1}^\star - \ldots - u_{m+1}^\star - c & \text{if} \quad c < 0,\ m < 0 . \end{cases}$$

Any set M_m of sequences for which a fixed m applies has its measure preserved, since by hypothesis a shift in the index i ($\bar{u}_i^\star = u_{i+m}^\star$) leaves unchanged the distribution of the u_i, which appears as the factor $dL_I(u_I^\star)$, and the factor dv_0 is unchanged by the translation of v_0. But any measurable set M can be expressed as a countable union of sets M_m (with m assuming all integral values) where the M_m are disjoint and have disjoint images. Hence M has its measure preserved.

Evidently the expected number of points of the continuous-parameter process falling in a finite interval is equal to the length of the interval divided by Eu, when the multiplicities of the points are taken into account.

A converse of 2.A. is

THEOREM 2.B. *Any stationary process with a continuous parameter t whose elements can be defined as sequences of random variables representing sequences of points of finite expected density on the t-axis can be obtained from a stationary process with discrete parameter by the construction of* 2.A.

PROOF. By the (expected) density of the points we mean the ratio of the expected number of points in an interval to the length of the interval; it is independent of the interval for a stationary continuous-parameter process. The finiteness of this density implies that sequences having cluster points are of zero probability, and so the given continuous-parameter process can be represented by the collection $\{v_0, u_i^\star\}$ of random variables used in 2.A.

Consider a set defined by

$$0 \leqq v_0' < v_0 \leqq v_0'', \qquad u_0' < u_0^\star \leqq u_0''$$

and similar intervals for a finite number of other $u_i^\star$. If $v_0'' \leqq u_0'$ the rectangular character of the set is unaffected by the restriction $v_0 < u_0^\star$. Then if c is such that $0 \leqq v_0' + c$ and $v_0'' + c \leqq u_0'$, the transformation $t = \bar{t} - c$ carries the given set into another rectangle having the same limits on the $\bar{u}_i^\star$ but with $v_0' + c < \bar{v}_0 \leqq v_0'' + c$. These sets have the same measure by hypothesis, which can be true in general only if the variables have distributions of the form

$$\sigma\, dL_I^\star(u_I^\star)\, dv_0 \qquad (0 \leqq v_0 < u_0^\star)$$

as in 2.A. The expected density of stacks of points (points counted without regard for their multiplicities) is

$$\lim_{\varepsilon \to 0} \Pr\{v_0 < \varepsilon\}/\varepsilon\,,$$

which is the integral of the expression $\sigma\, dL_I^\star(u_I^\star)$ over the region $u_0^\star > 0$, all the other $u_i^\star$ being unrestricted. The hypothesis implies that this density is positive and finite, and so we may make the integral of $dL_I^\star(u_I)$ over $u_0 > 0$ equal unity and let σ equal the stack density.

In case there is indeed a positive probability of having multiple points in the sequence, the distributions $L^\star$ just obtained are not quite those desired, since they make $\Pr\{u_0 = 0\} = 0$ and hence are not stationary in the discrete parameter. Before making the necessary adjustments, we first show that these distributions are stationary in a restricted sense, namely, when attention is confined to sets and transformations such that all the variables playing the role of u_0 are required to be positive. Consider a set A defined by

$$0 < u_0' < u_0 \leqq u_0'', \qquad u_1 = u_2 = \ldots = u_{n-1} = 0, \qquad 0 < u_n' < u_n \leqq u_n'',$$

and arbitrary inequalities on a finite number of other u_i $(n \geqq 1)$. Let η denote the smaller of u_0' and u_n'. Define a corresponding set $A^\star$ in the space of the continuous-parameter process by replacing the u_i by $u_i^\star$ in the definition of A and adding the condition $0 \leqq v_0 < \eta$. Then the transformation $\bar{t} = t - \eta$ applied to $A^\star$ is equivalent to

$$\bar{u}_i^\star = u_{i+n}^\star,$$

$$\bar{v}_0 = v_0 + u_n^\star - \eta\,.$$

Dividing out the factor η/Eu (relating to v_0) in the distribution, we see that the transformation $\bar{u}_i = u_{i+n}$ leaves the measure of A unchanged.

If one were content with this incomplete stationarity in the discrete parameter, one could assign probabilities arbitrarily within the region $u_0 = 0$, since the factor dv_0 and the condition $0 \leqq v_0 < u_0^\star$ reduce the probability of this region to zero in the continuous-parameter process. However, the requirement of complete stationarity serves to determine these probabilities uniquely in terms of those (the $L^\star$) conditioned by $u_0 > 0$.

Suppose for the moment that $p_1 = \Pr\{u_0 = 0\}$ is known. Let R be a finite-dimensional rectangle, and let T_n be the transformation

$$T_n: \quad (u_i) \to (\bar{u}_i) \qquad \text{with} \quad \bar{u}_i = u_{i+n} .$$

For any n such that $u_n > 0$ throughout R, the measure of $T_n R$ must be $1 - p_1$ times the measure assigned to it by the $L^\star$, which has been shown to be independent of n. For other values of n, the measure of $T_n R$ is hereby defined to have this same value. If there is no n such that $u_n > 0$ throughout R, then R may be partitioned into a finite union of rectangles for each of which such an n exists, plus a set of the form $u_i = 0$ for $i \in I_0$. For simplicity, let the class I_0 be enlarged if necessary so that it is composed of m consecutive integers. Then if

$$p_m = \Pr\{u_i = 0, \; i = 1, \ldots, m\} ,$$

one must have

$$p_m = p_\infty + \sum_{j=0}^{\infty} \Pr\{u_{-j} > 0; \;\; u_i = 0 \text{ for } i = -j+1 \text{ to } m\} .$$

The stationarity of the original continuous-parameter process requires that $p_\infty = \Pr\{u_i = 0 \text{ for all } i\}$ be zero, in agreement with (1 c). The other probabilities have already been defined, so that one has

$$\text{(2 a)} \qquad p_m = (1 - p_1) \sum_{j=m}^{\infty} \Pr\{u_i = 0 \text{ for } i = 1 \text{ to } j \mid u_0 > 0\} .$$

Putting $m = 0$ or 1 (and $p_0 = 1$) in (2 a) determines p_1:

$$\text{(2 b)} \quad 1/(1 - p_1) = 1 + \Pr\{u_1 = 0 \mid u_0 > 0\} + \Pr\{u_1 = u_2 = 0 \mid u_0 > 0\} + \ldots .$$

It is evidently necessary that this series converge, so that $p_1 \neq 1$. This it does because its product with the positive constant σ (the stack density) gives the expected density of points (taking their multiplicities into account), which is finite by hypothesis. Thus the factor $1/Eu$ of 2.A. equals $\sigma/(1 - p_1)$. It is easily seen that the new distributions defined above are mutually consistent as well as stationary. This completes the proof of 2.B.

In 2.A. a stationary process with continuous parameter was obtained, roughly speaking, by taking a space of sequences of points whose relative positions only were specified, and locating the origin of abscissas at random with respect to the points of each sequence. The identity of u_0 or any other u_i was supposed to be known in advance. If this is not true, an analogous procedure can be applied at an earlier stage. That is, suppose one is given merely a space of sequences of numbers, such as

$$\ldots, 1, \pi, 0, 1, \pi, 0, 1, \pi, 0, \ldots ,$$

without any information as to which number of the sequence is to be designated as u_0. Then by assigning the index zero with equal probabilities to each of N consecutive numbers in the sequence, and letting $N \to \infty$, one can obtain a stationary process with discrete parameter, provided the original space of unlabeled sequences has a suitable probability measure. The latter may be far simpler than any other description of the new process, hence the practical importance of the result. The following applies this idea and gives a method for specifying the original sequences.

THEOREM 2.C. *Let a stationary process of sequences*

$$\{\mu_j\}, \qquad j = \ldots, -1, 0, 1, 2, \ldots ,$$

of positive integers be given such that $E\mu_j = \bar{\mu}$ *is finite. Let the arbitrary probability distributions* $H_\mu(w_1, \ldots, w_\mu)$ *be defined for all positive integers* μ *that occur with positive probability in the sequences* $\{\mu_j\}$. *The* H_μ *need not be mutually consistent. Then the space of sequences* $\{u_i\} =$

$$\ldots, w_1^{(-1)}, w_2^{(-1)}, \ldots, w_{\mu_{-1}}^{(-1)}, w_1^{(0)}, w_2^{(0)}, \ldots, w_{\mu_0}^{(0)}, w_1^{(1)}, w_2^{(1)}, \ldots, w_{\mu_1}^{(1)}, \ldots$$

is stationary in the index i, provided that for given values of the μ_j, *any set of variables*

$$w^{(j)} = \left(w_1^{(j)}, \ldots, w_{\mu_j}^{(j)}\right)$$

is distributed according to H_{μ_j} *independently of all other such sets* $w^{(k)}$ *with* $k \neq j$, *and the index zero (identifying* u_0*) is located at random as described below. (The resulting distributions are written down explicitly in* (2 c) *below.)*

In the present applications the $w_1, \ldots, w_\mu$ will be non-negative, but this is not necessary for the theorem.

PROOF OF THEOREM 2.C. As in 2.A., let I be a variable whose range is all the finite sets of integers that include zero. The new process is defined by identifying u_0 with the number $w_m^{(0)}$, where m is a new random positive integer. If $p_I(\mu_I)$, where μ_I is a vector, is the (discrete) probability in the original process of the simultaneous realization of the values μ_j with $j \in I$, then their joint probability with m in the new process is defined to be $p_I(\mu_I)/\bar{\mu}$ for each integer m such that $1 \leqq m \leqq \mu_0$, and zero otherwise. As with the u_i in 2.A., this involves a change in the probability measure of the μ_j, which now have the marginal probability $\mu_0 p_I(u_I)/\bar{\mu}$, while that of m is

$$\sum_{\mu_0=m}^{\infty} p(\mu_0)/\bar{\mu} \, .$$

The distribution of the w's or u_i as conditioned by $\{m, \mu_j\}$ is given by the H_μ, with the understanding that any set of variables $w_1^{(j)}, \ldots, w_\mu^{(j)}$ is independent of all other $w_\nu^{(k)}$ with $k \neq j$. This defines the distribution function for any finite selection of variables of the form

$$\{m, \mu_i, u_i;\ \ r \leqq i \leqq s\}.$$

The distribution function of u_r to u_s inclusive thus has the form

$$\text{(2c)}\quad L_{s-r+1}(u_r, u_{r+1}, \ldots, u_s) = \bar{\mu}^{-1} \sum_{(\mu_j)} \sum_{q=1}^{s-r+1} \sum_{\varrho=0}^{\infty} \sum_{\sigma=0}^{\infty} p_q H_{\varrho+\mu_1'} H_{\mu_2} \ldots H_{\mu_q'+\sigma}$$

where

$$\begin{aligned} p_q &= p_q(\varrho+\mu_1', \mu_2, \ldots, \mu_{q-1}, \mu_q'+\sigma), \\ H_{\varrho+\mu_1'} &= H_{\varrho+\mu_1'}(\infty, \ldots, \infty, u_r, \ldots, u_{r+\mu_1'-1}), \\ H_{\mu_2} &= H_{\mu_2}(u_{r+\mu_1'}, \ldots, u_{r+\mu_1'+\mu_2-1}), \\ &\ldots \\ H_{\mu_q'+\sigma} &= H_{\mu_q'+\sigma}(u_{s-\mu_q'+1}, \ldots, u_s, \infty, \ldots, \infty), \end{aligned}$$

and where ϱ and σ are the numbers of symbols ∞ occurring as arguments of the H's at the beginning and the end, respectively, and $p_q(\ldots)$ is the probability that q consecutive μ_j in the original process have the values specified. The outer summation is taken over all the ordered partitions of $s-r+1$ into a sum

$$\mu_1' + \mu_2 + \ldots + \mu_{q-1} + \mu_q'$$

of q positive integers; "ordered" means for example that $1+2$ and $2+1$ are regarded as distinct partitions of 3. If $q=1$, then the typical term is

$$\bar{\mu}^{-1}\, p(\varrho + s - r + 1 + \sigma)\, H_{\varrho+s-r+1+\sigma}(\infty, \ldots, \infty, u_r, \ldots, u_s, \infty, \ldots, \infty).$$

It is obvious that, as implied by the notation, the function L_{s-r+1} does not depend on r and s individually, but only on their difference. These distributions will therefore be mutually consistent and stationary if

$$L_{s-r+1}(\infty, u_{r+1}, \ldots, u_s) = L_{s-r}(u_{r+1}, \ldots, u_s) = L_{s-r+1}(u_{r+1}, \ldots, u_s, \infty).$$

The equality between the first two members is typical. Putting $u_r = \infty$ in L_{s-r+1} evidently converts all the terms of L_{s-r+1} having $\mu_1' \geqq 2$ into those terms of L_{s-r} having $\mu_1' \geqq 1$, $\varrho \geqq 1$, and the same values of $\varrho + \mu_1'$ and the other variables, only the $\varrho \geqq 1$ being a significant restriction. Those terms of L_{s-r+1} having $\mu_1' = 1$ drop the unit factor $H_{\varrho+\mu_1'}(\infty, \ldots, \infty)$ when $u_r = \infty$, and pass into the remaining terms of L_{s-r}, which have $\varrho = 0$, a value $\bar{q}$ equal to the old q diminished by one, and values of

$$\varrho + \mu_1', \mu_2, \ldots, \mu_{\bar{q}-1}, \mu_{\bar{q}}' + \sigma$$

equal to the old values of

$$\mu_2, \mu_3, \ldots, \mu_{q-1}, \mu_q' + \sigma,$$

respectively. This is consistent in view of the equality

$$\sum_{\varrho=0}^{\infty} p_q(\varrho+1, \mu_2, \ldots, \mu_{q-1}, \mu_q'+\sigma) = p_{q-1}(\mu_2, \ldots, \mu_{q-1}, \mu_q'+\sigma).$$

The last step in the process, showing that $L_1(\infty)=1$, gives

$$\frac{1}{\bar{\mu}} \sum_{\varrho=0}^{\infty} \sum_{\sigma=0}^{\infty} p_1(\varrho+\sigma+1) = \frac{1}{\bar{\mu}} \sum_{k=1}^{\infty} k\, p_1(k) = 1.$$

This completes the proof of 2.C.

Theorem 2.C. was formulated in the course of showing that the expected number of points falling in a finite interval may be infinite when the discrete-parameter formulation is employed, although this will cease to be so if the process is converted by 2.A. to one having a continuous parameter.

To do this one assumes that the variance of the μ_i in the original measure is infinite, so that in the new measure the mean of μ_0 (which acquires the marginal probability $\mu_0 p(\mu_0)/\bar{\mu}$) is infinite. The μ_i may be assumed to be originally mutually independent. The simplest set of H_μ are the degenerate distributions which make $w_1=1$ and $w_i=0$ $(i>1)$ with probability one. The sequences are then obtained by putting μ_i coincident points at the place of abscissa i, for $i=\ldots, -1, 0, 1, 2, \ldots$. Since a large value of μ_i casts a larger net and is more likely to capture the reference index zero than are small values, the expected number of points at $v=i=0$ is infinite, although it is finite $(=\bar{\mu})$ for any other value of i, and $\lambda(U)=1/\bar{\mu}$.

The restriction in 2.C. that $E\mu_j$ be finite is necessary for the theorem but not for the construction of interesting examples. If $E\mu_j=\infty$ and μ_j coincident points are placed at abscissa j, $j=\ldots, -1, 0, 1, 2, \ldots$, a continuous-parameter process with infinite point-density can be constructed by a random placement of the origin, but no corresponding stationary discrete-parameter process exists, because $\lambda(U)\equiv 0$. If $E\mu_j=\infty$ and μ_j coincident points are placed at 0 for $j=0$, at

$$\sum_{i=1}^{j} \mu_i \quad \text{for} \quad j>0$$

and at

$$-\sum_{i=0}^{|j-1|} \mu_{-i} \quad \text{for} \quad j<0,$$

the point-density may be said to be finite, but neither the discrete nor the continuous-parameter stationary process can be constructed; if so, one could use the ergodic theorem to show (falsely) that in the discrete case $\Pr\{u_i=0\}=1$, and in the continuous case, that the density of points or stacks was zero. In all these examples, the coincidence of the μ_j points is easily seen to be inessential.

3. Permutation of the points of the sequence. We now suppose that each point v_i of the typical sample sequence is subjected to a random displacement x_i, so that the new abscissa V_i is v_i+x_i. The minimum assumption is that the stationarity condition (1 a) continue to hold for the combined process $\{u_i, x_i\}$, so that the transformation

$$(u_i, x_i) \rightarrow (u_{i+1}, x_{i+1})$$

is measure-preserving. The common distribution function of the x_i is denoted by $F(x)$.

Although $v_i \leqq v_{i+1}$ for all i, this is not in general true for the V_i, hence the term "permuted" sequences. However, if the V_i have no (finite) cluster points, they may be renumbered and denoted by V_j', so that $V_j' \leqq V_{j+1}'$. The first results of this section are that the space of sequences $\{u_j'\}=\{V_j'-V_{j-1}'\}$ exists and has the same properties (1 a–c) as the space of the $\{u_i\}$ originally given.

THEOREM 3.A. *The probability is zero that the permuted sequence have a cluster point.*

PROOF. By the stationarity, the probability that the permuted sequence have a cluster point in the half-open interval (v_n, v_{n+1}) is independent of n. (As n varies from $-\infty$ to ∞, all possible cluster points are accounted for.) If this probability was positive, the expected number of points displaced into such an interval would be infinite. But this expected number is

$$\sum_{i=-\infty}^{\infty} \Pr\{v_n \leqq v_i+x_i < v_{n+1}\},$$

which by stationarity is

$$\sum_{i=-\infty}^{\infty} \Pr\{v_{n-i} \leqq x_0 < v_{n+1-i}\} = 1,$$

since $v_0=0$ and the events $v_{n-i} \leqq x_0 < v_{n+1-i}$ are mutually exclusive and exhaustive. This completes the proof.

The following partial proof of the above theorem, valid only when

Ex_i exists (and is finite), may help to show its significance. Consider the stationary variables

$$\bar{u}_i = u_i + x_i - x_{i-1},$$

which are the distances after displacement between points which were consecutive in the original sequence. Evidently $E\bar{u}_i = Eu_i > 0$, the value $+\infty$ being included. Then by the ergodic theorem,

$$\sum_1^\infty \bar{u}_i = \infty,$$

which is incompatible with the existence of a (finite) cluster point.

An immediate consequence of 3.A. is

THEOREM 3.B. *The set of permuted sequences is again a stationary process, satisfying* (1 a–c).

PROOF. Let the permuted sequence be denoted by

$$\cdots \leqq V_{-2}' \leqq V_{-1}' \leqq V_0' \leqq V_1' \leqq V_2' \leqq \cdots,$$

which is possible and satisfies (1 b–c) by 3.A., and let $u_j' = V_j' - V_{j-1}'$ ($V_0' = V_0 = x_0$). It remains to show that the transformation $u_j' \to u_{j+1}'$ is measure-preserving. This it is by the argument used in 2.A.: It is equivalent to a combination of the measure-preserving transformations $u_i \to u_{i+m}$, where m is such that if V_0' is obtained by displacing v_0, then V_1' is obtained by displacing v_m.

It is easily seen that in general the independence of the x_i does not imply that the reverse displacements x_j' (the $-x_i$ taken in the order of the V_j') are either mutually independent or independent of the $V_{j+1}' - V_j'$, although they are stationary by the same argument as above. These conclusions do hold, however, when the original sequences are Poisson, by the following theorems.

THEOREM 3.C. *The space of permuted sequences obtained from a Poisson process by subjecting each point to a displacement distributed independently of all else and having a common distribution* $F(x)$, *is again a Poisson process.*

THEOREM 3.D. *If the transformation* $\{v_i\} \to \{V_j'\}$ *of* 3.C. *is reversed, the displacements (taken in their new order) are likewise independent of one another and of the sequence* $\{V_j'\}$; *their distribution is of course* $1 - F(-x)$.

The proof of 3.C. is given in Doob [3, pp. 404–407]. It is easily modified to give 3.D. as well, by classifying the points not only according to the interval I_j into which they are displaced (by the forward transformation), but also according to the interval X_i in which the values of their displacements fall (i and j here have new uses). Then in place of $F(I_j - \xi)$ one has $F[(I_j - \xi) \cap X_i]$ with

$$\int_{-\infty}^{\infty} F[(I_j - \xi) \cap X_i] \, d\xi = (b_j - a_j) F(X_i) ,$$

where I_j and X_i are the intervals (a_j, b_j) and (c_i, d_i), and $F(X_i) = F(d_i) - F(c_i)$. The points of $\{V_j'\}$ which have undergone displacements $x \in X_i$ are found to constitute a Poisson process of density $F(X_i)$ times the density of the v_i, and independent of all other such processes corresponding to intervals X_k that are disjoint with X_i.

In the notation introduced at the start of this section, 3.D. states the independence of the X_j' and the sequence $\{V_j'\}$, which is true reciprocity; obviously the x_i and $\{V_i\} = \{v_i + x_i\}$ are not independent. In the following corollaries, the displacements are assumed to be independent as above.

COROLLARY 3.D′. *Given a Poisson sequence and a single permutation thereof, one cannot ascertain which of the two sequences is the original except perhaps by making use of the sense of the displacements.* (*The latter fails if* $F(x)$ *is symmetric about zero*).

COROLLARY 3.D″. *Two permuted Poisson sequences arising from the same original sequence are related as an original and one permuted sequence such that the displacements have the symmetrical distribution*

$$G(x) = F(x) * [1 - F(-x)] = \int_{-\infty}^{\infty} F(x+t) \, dF(t) .$$

4. The number of inversions. Some simple combinatorial ideas will now be introduced in connection with the permutations defined in the preceding section. Let the displaced points $V_i = v_i + x_i$ be considered in the order in which they occur on the line, from left to right. (To avoid ambiguities, it will be agreed that when two or more points coincide after displacement, they shall be regarded as having the same ordinal relations as before displacement; i.e. there is no inversion among them.) In general the subscripts i will then not occur in their natural order, and we may write $i = \pi(j)$, where $\pi(0)$ equals some arbitrary integer h which may or may not depend on the sequence considered, and $\pi(j)$ denotes the subscript i of the j-th point V_i to the right of V_h (if $j > 0$) or the $(-j)$-th point to the left of V_h (if $j < 0$). Thus one may write $V_j' = V_{\pi(j)}$, in the notation of the preceding section. The choice of h

is essentially a matter of notation. The simplest choice $h=0$ will be adopted, although a different determination is suggested in connection with equation (4 a) and theorem 4.F. below. The present section is concerned only with the properties of the random function $\pi(j)$.

In terms of $\pi(j)$, other integer-valued functions may be defined: $k_+(j_0)$ denotes the number of integers j such that $j>j_0$ but $\pi(j)<\pi(j_0)$. Thus $k_+(j_0)$ also denotes the number of points (called right crossovers) which were to the left of $v_{i_0}=v_{\pi(j_0)}$ in the original sequence, but after displacement found themselves to the right of the corresponding point $V_{i_0}=v_{i_0}+x_{i_0}$. Similarly $k_-(j_0)$ denotes the number of left crossovers, or values of j such that $j<j_0$ but $\pi(j)>\pi(j_0)$. Finally, $K(j)$ may be defined as the number of inversions that straddle a cut between the points $V_{\pi(j)}=V_j'$ and V_{j+1}'; more precisely, it is the number of different pairs of integers j_1, j_2 such that $j_1 \leqq j, j+1 \leqq j_2$, but $\pi(j_1)>\pi(j_2)$. Of the integer-valued variables used in this section, $k_+(j)$, $k_-(j)$, and $K(j)$ alone are necessarily non-negative.

The total number of inversions in which the point $V_{\pi(j)}=V_j'$ is involved is then $k_+(j)+k_-(j)$; but the difference $k(j)=k_+(j)-k_-(j)$ of these functions seems to be of greater interest. (In this paragraph it is assumed that $k_+(j)$ and $k_-(j)$ (and hence $K(j)$, as shown in 4.A. below) are finite-valued.) For example, one can show

$$\pi(j) = j+k(j)-m\,, \tag{4 a}$$

where m is an integer (possibly zero or negative) that depends on the permutation but not on the index j. Evidently $m=k(0)-\pi(0)=k(0)-h$, so that the choice $h=k(0)$ would eliminate the constant m from the equation. Also one has

$$k(j_0) = K(j_0) - K(j_0-1), \qquad \text{or} \qquad K(j_0) = \text{const.} + \sum_{n=0}^{j_0} k(n)\,, \tag{4 b}$$

since moving the "cut" from left to right across the point V_{j_0}' causes $k_-(j_0)$ inversions (involving V_{j_0}' and V_j' with $j<j_0$) no longer to straddle the cut, while $k_+(j_0)$ inversions (involving V_{j_0}' and V_j' with $j>j_0$) now straddle the cut which did not straddle it before. These relations are illustrated by the example at the end of the paper, in which $m=-2$.

The assumed stationarity has not been used thus far in this section, but it is required hereafter. Then any one of the variables $k(j)$, such as $k(0)$, has the same properties as any other, and similarly for the $k_+(j)$, $k_-(j)$, and $K(j)$.

THEOREM 4.A. *The variables $k_+(j)$ and $k_-(j)$ have equal means* ($+\infty$ *included*). *Thus* $Ek(j)=0$ *or is undefined* (*and not* $+\infty$ *or* $-\infty$).

PROOF. If the $k(j)$ are finite-valued,

$$\sup_{j \leqq j_0} \pi(j) \qquad \text{and} \qquad \inf_{j \leqq j_0+1} \pi(j)$$

are finite. $K(j_0)$ cannot exceed the square of half the difference between these two bounds, so it is finite. By (4 b) and the ergodic theorem, a non-zero $Ek(j)$ would imply

$$\lim_{j \to \infty} K(j) = -\infty \qquad \text{or} \qquad \lim_{j \to -\infty} K(j) = -\infty .$$

It remains to reduce to absurdity the assumption that $k_-(j)$, say, is infinite with positive probability, while $Ek_+(j)$ is finite. Let $k_\pm(j; m,n)$ be defined like $k_\pm(j)$, but with consideration limited to the finite portion $\pi(m), \pi(m+1), \ldots, \pi(n)$ of the permutation. Then

$$\sum_{j=-n}^{n} k_-(j; -n,n) = \sum_{j=-n}^{n} k_+(j; -n,n) ,$$

since each member gives the total number of inversions among the $2n+1$ points considered. Since $k_+(j; -n,n) \leqq k_+(j)$ and $Ek_+(j)$ is finite,

$$\limsup_{n \to \infty} \frac{1}{2n} \sum_{j=-n}^{n} k_+(j; -n,n) < \infty .$$

The desired contradiction will be obtained by showing that

$$\lim_{n \to \infty} \frac{1}{2n} \sum_{j=-n}^{n} k_-(j; -n,n) = \infty$$

when $k_-(j) = \infty$.

Let A be an arbitrary (large) constant, and define $\chi_N(j)$ as unity when $k_-(j; j-N, j) \geqq A$ and zero otherwise. Then the $\{\chi_N(j)\}$ (with fixed N) are stationary sequences,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} \chi_N(j) = f(N)$$

exists with probability one, and

$$\lim_{N \to \infty} \chi_N(j) = 1 \qquad \text{for every} \quad j .$$

Also $Ef(N) = E\chi_N(j)$, and letting $N \to \infty$ shows that $E \lim f(N) = 1$ and hence $\lim f(N) = 1$. Thus for almost every sequence or permutation, there exist an n_0 and an N_0 such that

$$\frac{1}{n} \sum_{j=1}^{n} \chi_N(j) > \tfrac{1}{2} \qquad \text{whenever} \qquad n > n_0 \ \text{ and } \ N > N_0 .$$

This implies that

$$\frac{1}{2n}\sum_{j=-n}^{n} k_-(j; -n,n) > \tfrac{1}{4}A \qquad \text{whenever} \quad n > \max(n_0, N_0)\,.$$

THEOREM 4.B. *Any variable $k_-(j_0)$ is zero with positive probability, or else infinite with unit probability; and similarly for $k_+(j_0)$.*

PROOF. If $\Pr\{k_-(j_0)>0\}=1$, then there is also unit probability that $k_-(j)>0$ for all j. Thus any point $\pi(j_0)$ has at least one $\pi(j_1)>\pi(j_0)$ with $j_1<j_0$, while $\pi(j_1)$ has at least one $\pi(j_2)>\pi(j_1)$ with $j_2<j_1$, and so on. All these $\pi(j_\gamma)$ $(\gamma>0)$ are left crossovers with respect to $\pi(j_0)$, so that $k_-(j_0)=\infty$.

THEOREM 4.C. *The variables $k_-(j)$ are almost always finite if*

$$\int_{-\infty}^{0} x\,dF(x)$$

converges, and the $k_+(j)$ are finite if

$$\int_{0}^{\infty} x\,dF(x)$$

converges.

PROOF. The variable $k_-(0)$ is representative. By writing $x_i=x_i^+-x_i^-$, where $x_i^\pm \geqq 0$ and at least one of x_i^+, x_i^- is zero for every i, the displacements may be applied in two stages, first $-x_i^-$ and then x_i^+. It is easily shown that the final $k_-(0)$ cannot exceed the sum of those at the two stages. The $k_-(0)$ resulting from the $-x_i^-$ is finite by a proof exactly like the second (partial) proof of 3.A., since Ex_i^- exists; and by 3.B. one has again a space of stationary sequences. The $k_-(0)$ resulting from the x_i^+ will be finite if $x_0^+<v_i+x_i^+$ for all but a finite number of positive integers i, with probability one. But by the ergodic theorem the statement is true for $x_0^+<v_i$, and so a fortiori for $x_0^+<v_i+x_i^+$, since $x_i^+\geqq 0$.

A partial converse is given by

THEOREM 4.D. *The variables $k_-(j)$ are almost always infinite if*

$$\int_{-\infty}^{0} x\,dF(x)$$

diverges, and the $k_+(j)$ are infinite if

$$\int_0^\infty x\,dF(x)$$

diverges, provided that $\lambda(U)$ is almost always finite and each x_i is independent of all else.

PROOF. By hypothesis, for any $\varepsilon > 0$, there exists a c such that

$$\Pr\{\lambda(U) < c\} > 1-\varepsilon ,$$

and hence

$$\lim_{N\to\infty} \Pr\{v_i \leqq ic, \text{ all } i \geqq N\} \geqq 1-\varepsilon .$$

Hence with probability at least $1-\varepsilon$,

$$\lim_{N\to\infty} \sum_{i=N}^{\infty} F(-v_i+x) \geqq \lim_{N\to\infty} \frac{1}{c} \int_{(N+1)c}^{\infty} F(-v+x)\,dv .$$

The latter is infinite since $\int_{-\infty}^{0} x dF(x)$ diverges, so

$$\prod_{i=N}^{\infty} [1-F(-v_i+x)] = 0;$$

since ε is arbitrary, this holds with probability one. The expectation of this vanishing expression is the probability that all left crossovers ($k_-(0)$ in number) are confined to the first $N-1$ points originally to the right of $v_0=0$.

The case in which Ex_i does not exist but all x_i are equal for any one sequence is a counterexample to dropping the independence of the x_i.

THEOREM 4.E. *If each x_i is independent of all else, then for any fixed realization of $\{v_i\}$ with v_n and $-v_{-n} \to \infty$ monotonely as $n \to \infty$, the probability that $k_-(0)$ is finite is either zero or one, and similarly for $k_+(0)$. If both $k_-(0)$ and $k_+(0)$ are finite, then they vanish simultaneously with positive probability.*

PROOF. The probability that $k_-(0)$ is finite is

$$\lim_{N\to\infty} \prod_{i=N}^{\infty} [1-F(-v_i+x)] ,$$

where x is the displacement applied to the point $v_0=0$, and $1-F(-v_i+x)$ is the probability that the point v_i is not displaced to the left of x. Evidently the limit is either 0 or 1 independently of x. Now if both $k_-(0)$ and $k_+(0)$ are finite one has

$$\lim_{N\to\infty} \prod_{i=N}^{\infty} F(|v_{-i}|+x)\,[1-F(-v_i+x)] = 1 .$$

It follows that

$$\prod_{i=1}^{\infty} F(|v_{-i}|+x)\,[1-F(-v_i+x)] > 0\,, \tag{4 c}$$

since an individual factor can vanish only on a set of measure zero. (The remark in the parenthesis at the beginning of this section brings ambiguous cases into conformity with the statement.) (4 c) is the probability that $k_-(0)$ and $k_+(0)$ vanish simultaneously.

THEOREM 4.F. *If each x_i is independent of all else and the variables $k_-(j)$ and $k_+(j)$ are almost always finite, then the infinite permutation is almost certainly expressible as the product of a translation $j=\bar{j}+m$ together with permutations involving finite disjoint sets of consecutive integers (the indices of the points).*

PROOF. It follows from 4.E. and the mutual independence of the x_i that for almost every permuted sequence, there is an infinite number of values of j such that $k_+(j)=k_-(j)=0$. In each such case one evidently has also $K(j-1)=K(j)=0$. Now the translation is used, if it is needed, to eliminate the constant m in equation (4 a). It then remains to show that if j_1 and j_2 are two integers such that $j_1<j_2$ and $K(j_1)=K(j_2)=0$, then the integers

$$\pi(j_1+1),\ \pi(j_1+2),\ \ldots,\ \pi(j_2)$$

are a permutation of

$$j_1+1, j_1+2, \ldots, j_2\,.$$

By their construction, they must be a permutation of

$$j_1+h+1, \ldots, j_2+h\,,$$

where h is to be determined. Let $\pi(j_0)=j_2+h$ be the greatest of these integers. Then it has j_2-j_0 smaller integers on its right, and no larger integers on its left. Thus $k(j_0)=j_2-j_0$, and substitution in (4 a) with $m=0$ gives $h=0$ as desired. This completes the proof.

If one only requires that the x_i be independent of one another (for every fixed sequence $\{u_i\}$), or that the $\{x_i\}$ be independent of the $\{u_i\}$, one can easily arrange to obtain always a permutation such as

j =	...	0,	1,	2,	3,	4,	5,	6,	7,	8,	9,	...
$\pi(j)$ =	...	6,	1,	2,	9,	4,	5,	12,	7,	8,	15,	...
$k_+(j)$ =				0	4	0	0	4	0	0	4	
$k_-(j)$ =				2	0	2	2	0	2	2	0	
$k(j)$ =				−2	4	−2	−2	4	−2	−2	4	
$K(j)$ =				2	6	4	2	6	4	2	6	

for which $k(j)$ and $K(j)$ never vanish.

REFERENCES

1. H. Cramér, *Mathematical methods of statistics*, Uppsala, 1945, Princeton, 1946.
2. J. L. Doob, *Renewal theory from the point of view of the theory of probability*, Trans. Amer. Math. Soc. 63 (1948), 422–438.
3. J. L. Doob, *Stochastic processes*, New York, 1953.
4. W. Feller, *An introduction to probability theory and its applications* I, New York, 1950.
5. E. Hopf, *Ergodentheorie* (Ergebnisse der Mathematik 5), Berlin, 1937.
6. E. L. Kaplan, *Infinite permutations of stationary random sequences*, Doctoral dissertation, Princeton, November 1950.

BELL TELEPHONE LABORATORIES, INC., MURRAY HILL, N.J., U.S.A.