ON FUNCTIONS OF SEQUENCES OF INDEPENDENT CHANCE VECTORS WITH APPLICATIONS TO THE PROBLEM OF THE "RANDOM WALK" IN k DIMENSIONS

BY D. BLACKWELL AND M. A. GIRSHICK

Howard University and U. S. Department of Agriculture

1. Summary. Consider a sequence $\{x_i\}$ of independent chance vectors in k dimensions with identical distributions, and a sequence of mutually exclusive events $S_1, S_2, \cdots$, such that S_i depends only on the first i vectors and $\Sigma P(S_i) = 1$. Let φ_i be a real or complex function of the first i vectors in the sequence satisfying conditions: (1) $E(\varphi_i) = 0$ and (2) $E(\varphi_j \mid X_1, \cdots, X_i) = \varphi_i$ for $j \geq i$. Let $\varphi = \varphi_i$ and $n = i$ when S_i occurs. A general theorem is proved which gives the conditions φ_i must satisfy such that $E\varphi = 0$. This theorem generalizes some of the important results obtained by Wald for $k = 1$. A method is also given for obtaining the distribution of φ and n in the problem of the "random walk" in k dimensions for the case in which the components of the vector take on a finite number of integral values.

2. A basic theorem.

2.1 Let $\{X_i\} = \{(X_{1i}, X_{2i}, \cdots, X_{ki})\}$ be a sequence of independent k-dimensional chance variables with identical distributions. Let $S_1, S_2, S_3, \cdots$, be mutually exclusive events such that (1) S_i depends only on $X_1, X_2, \cdots, X_i$, and (2) $\Sigma P(S_i) = 1$. Let $\varphi_i(X_1, X_2, \cdots, X_i)$ be a sequence of real or complex variables satisfying the following two conditions:

Condition 1: $E(\varphi_i) = 0$ for all i.

Condition 2: $E(\varphi_j \mid X_1, X_2, \cdots, X_i) = \varphi_i$ for all $j \geq i$, where $E(\varphi_j \mid X_1, X_2, \cdots, X_i)$ stands for the expected value of φ_j under the condition that $X_1, X_2, \cdots, X_i$ are held constant.[1] Define $\varphi_i = \varphi$ and $n = i$ when the event S_i occurs. We shall assume that $E(n)$ is finite.

A problem of central importance in sequential theory may be formulated as follows: What conditions must φ_i satisfy so that $E(\varphi)$ exists and equals zero? We shall prove the following:

THEOREM 2.1. *If there exists a function* $f(x_1, x_2, \cdots, x_k) \geq 0$ *such that* (a) $E[f(X_i)]$ *is finite and* (b) $|\varphi_i| \leq \sum_{d=1}^{i} f(X_d)$ *when* $n \geq i$, *then* $E(\varphi)$ *exists and equals zero.*

Before proceeding to the proof, we consider two consequences of this theorem.

I. Assume that $E(X_{ri}) = a_r$. Let $\varphi_i = \sum_{j=1}^{i} (X_{rj} - a_r)$. It is easily verified that φ_i satisfies conditions 1 and 2. We set $f(x_1, x_2, \cdots, x_k) = | x_r$

[1] Chance variables φ_i satisfying condition 2 have been extensively studied by P. Levy [1] and J. L. Doob [2].

$- a_r |$. Then Theorem 2.1 is applicable and we get $E\varphi = 0$. Now $\varphi = W_r - na_r$ where $W_r = \sum_{i=1}^{n} X_{ri}$. Hence we have

$$E(W_r) = a_r E(n). \tag{2.11}$$

The relationship (2.11) has been proved for $k = 1$ by Wald [3] and subsequently under somewhat more generalized conditions, by one of the authors [4].

II. Let $t_1, t_2, \cdots, t_k$ be any real or complex numbers for which $Ee^{\sum_{r=1}^{k} t_r x_{ri}} = a$ is finite and $|a| \geq 1$. We assume that there exists a positive constant M such that

$$\left| \sum_{i=1}^{m} X_{ri} \right| \leq M, r = 1, 2, \cdots, k, \tag{2.12}$$

when $n > m$. Let

$$\varphi_i = a^{-i} e^{\sum_{j=1}^{i} \sum_{r=1}^{k} t_r x_{rj}} - 1 \tag{2.13}$$

so that

$$\varphi = a^{-n} e^{\sum_{r=1}^{k} t_r W_r} - 1 \tag{2.14}$$

where W_r is defined as above. It is easy to show that φ_i satisfies conditions 1 and 2. Now, in view of (2.12), when $n \geq i$

$$|\varphi_i| \leq |a|^{-i} e^{M \sum_{r=1}^{k} |\tau_r|} e^{\sum_{r=1}^{k} \tau_r x_{ri}} + 1 \leq 1 + Re^{\sum_{r=1}^{k} \tau_r x_{ri}} \tag{2.15}$$

where τ_j is the real part of t_j and $R = e^{M \sum_{r=1}^{k} |\tau_r|}$ is a fixed positive constant. Then, letting

$$f(x_1, x_2, \cdots, x_k) = 1 + Re^{\sum_{r=1}^{k} \tau_r x_{ri}} \tag{2.16}$$

we may apply Theorem 2.1 and obtain

$$E\left(a^{-n} e^{\sum_{r=1}^{n} t_r W_r}\right) = 1 \tag{2.17}$$

which is a generalization of the Fundamental Identity proved by Wald [5] for the case $k = 1$.

PROOF OF THEOREM 2.1. Assume φ_i is real. Define chance variables N_m inductively as follows: $N_0 = 0$. Assuming $N_0, \cdots, N_m$ defined, define $N_{m+1} = N_m + n(X_{N_m+1}, X_{N_m+2}, \cdots)$. Also let $n_m = N_m - N_{m-1}$ and $y_m = f(X_{N_{m-1}+1}) + \cdots + f(X_{N_m})$. It can be shown by induction that N_m is defined for all m with probability one, and that $\{n_m\}$, $\{y_m\}$ are sequences of independent chance variables with identical distributions. Clearly $n_1 = n$.

The Strong Law of Large Numbers asserts that if $z_1, z_2, \cdots$ are independent chance variables with identical distribution, then $\lim_{m\to\infty} \frac{z_1 + z_2 + \cdots + z_m}{m} = c$ with probability one if and only if Ez_1 exists and equals c.

It follows that, with probability one

$$\lim_{m\to\infty} \frac{f(X_1) + \cdots + f(X_m)}{m} = E[f(X_1)] \tag{2.18}$$

and

$$\lim_{m\to\infty} \frac{n_1 + \cdots + n_m}{m} = \lim_{m\to\infty} \frac{N_m}{m} = E(n). \tag{2.19}$$

Since $\frac{y_1 + \cdots + y_m}{n_1 + \cdots + n_m} = \frac{y_1 + \cdots + y_m}{N_m}$ is a subsequence of $\frac{f(X_1) + \cdots + f(X_m)}{m}$, we have with probability one,

$$\lim_{m\to\infty} \frac{y_1 + \cdots + y_m}{N_m} = E[f(X_1)] \tag{2.20}$$

so that

$$\lim_{m\to\infty} \frac{y_1 + \cdots + y_m}{m} = E[f(X_1)]E(n). \tag{2.21}$$

Consequently, $E(y_1)$ exists and equals $Ef(X_1)E(n)$. Since $|\varphi| \leq y_1$, $E(\varphi)$ exists. Also using conditions (2) and (b) which were imposed on φ_i we have

$$\begin{aligned} \left|\int_{S_1+\cdots+S_i} \varphi\, dp\right| &= \left|\sum_{j=1}^{i} \int_{S_j} \varphi_j\, dp\right| = \left|\sum_{j=1}^{i} \int_{S_j} \varphi_i\, dp\right| \\ &= \left|-\int_{n>i} \varphi_i\, dp\right| = \left|\sum_{j>i} \int_{S_j} \varphi_i\, dp\right| \\ &\leq \sum_{j>i} \int_{S_j} |\varphi_i|\, dp \leq \sum_{d>i} \int_{S_j} y_1\, dp \end{aligned} \tag{2.22}$$

which approaches zero as $i \to \infty$. This completes the proof.

If φ_j is a complex valued function, Theorem 2.1 still holds. For writing $\varphi_j = g_j + ih_j$ then Condition 2 becomes $E(g_p + ih_p \mid X_1, \cdots, X_j) = g_j + ih_j$ when $p \geq j$. Hence

$$E(g_p \mid X_1, \cdots, X_j) = g_j \tag{2.23}$$

and

$$E(h_p \mid X_1, \cdots, X_j) = h_j \tag{2.24}$$

when $p \geq j$. Since $|g_j| \leq |\varphi_j|$ and $|h_j| \leq |\varphi_j|$ and φ_j satisfies condition (b) we may apply Theorem 2.1 and get

$$Eg = E(h) = 0. \tag{2.25}$$

Hence $E\varphi = 0$.

3. Applications to the problem of the random walk in k dimensions[2]

3.1. *A theorem concerning decision points.* Let $\{X_j\} = \{(X_{1i}, \cdots, X_{ki})\}$ be a sequence of k-dimensional chance vectors with identical distributions. We assume that X_{ji} $(j = 1, 2, \cdots, k)$, take on a finite number of integral values ranging from $-r_j$ to m_j inclusive, where r_j and m_j are positive integers. We remark that any distribution can be approximated to any degree of accuracy by the distribution of a variate whose values are integral multiples of a constant d, which can be taken as the unit of measurement.

Let $P_{u_1u_2\cdots u_k}$ be the probability that $X_i = (u_1, u_2, \cdots, u_k)$. We define $W_{pi} = \sum_{j=1}^{i} X_{pj}$ and set $U_i = (W_{1i}, W_{2i}, \cdots, W_{ki})$. Then $\{U_i\}$ represents a sequence of points with integral coordinates in a k-dimensional space $S_k = \{(y_1, y_2, \cdots, y_k\}$. Let R be an arbitrary bounded region in S_k. We shall assume, without loss of generality, that the origin is an interior point of R. We now define a random variable n as the smallest subscript i of the sequence $\{U_j\}$ for which W_i is either a boundary point or an exterior point of R. We set $U_n = W = (W_1, W_2, \cdots, W_k)$ and designate W as a decision point of R. Clearly, the number of decision points is finite.

The random variables n and W can be interpreted as follows: Consider a point Q which at the time $t = 0$ is at the origin. At successive intervals of time $t = 1, 2, \cdots$, the point Q moves with integral components in S_k the direction and distance of the motion being determined by chance. The point comes to rest as soon as, but not before it either reaches the boundary of R or falls outside of R. Let U_t be the co-ordinates of the point Q at time t. Then n represents the length of time it takes Q to come to rest, and W represents a possible resting point.[3]

We shall be concerned with the problem of finding the probability distribution of n and W. These will obviously depend on the shape of the region R. In what follows we shall restrict ourselves to the class of regions R which have the property that the intersection of any line parallel to the axes with R is an open interval. In view of the fact that W has integral coordinates, we can without any loss of generality, replace this class of regions by an equivalent class which are bounded by simple polygonal closed surfaces whose vertices have integral coordinates and whose sides are parallel to the planes $y_j = 0$. In the subsequent discussion we assume that the regions R are of this type.

Let

$$\text{(3.10)} \qquad \underset{(y_1, y_2 \cdots y_k) \in R}{\text{l.u.b.}} [y_i]$$

[2] What follows is a generalization of a method previously employed by one of the authors [6] for the case $k = 1$.

[3] That Q will reach a resting point eventually can be asserted with probability one. See A. Wald [5], Lemma 1.

and

$$-b_i = \underset{y_i}{\text{g.l.b.}}\, [(y_1, y_2, \cdots, y_k) \in R] \tag{3.11}$$

then a_i and b_i are positive integers.

We now prove the following:

LEMMA 3.1. *For the given sequence of chance vectors $\{X_i\}$ and the given region R, the number of possible decision points N_R is given by*

$$N_R = \prod_{j=1}^{k} (a_j + b_j + r_j + m_j - 1) - \prod_{j=1}^{k} (a_j + b_j - 1). \tag{3.12}$$

PROOF: We shall first prove this theorem for a rectangular region $R = R_1$ where R_1 is defined by $-b_i < y_i < a_i$, $(i = 1, 2, \cdots, k)$ and then generalize the proof to any region of the class specified.

Let R_2 be a closed rectangular region defined by $-(b_i + r_i - 1) \leq y_i \leq (a_i + m_i - 1)$. Then $R_2 \geq R_1$. Let $S = R_2 - R_1$. It is clear that every integral point of S is a possible decision point. Moreover, no point exterior to R_2 is a possible decision point. For assume, for example, that there exists a point $W = (W_1, W_2, \cdots, W_k)$ which is an exterior point of R_2. Then at least one of its coordinates, say W_j, has the property that $W_j > a_j + m_j - 1$ or $W_j < -(b_j + r_j - 1)$. But since $-(b_j - 1) \leq W_{j,n-1} \leq a_j - 1$, it must follow that X_{jn} took on a value greater than m_j or less than $-r_j$ which is contrary to assumption. Now the total number of integral points contained in R_1 is $\prod_{j=1}^{k} (a_j + b_j + r_j + m_j - 1)$ and the total number of integral points in R_1 which by assumption are not decision points, is $\prod_{j=1}^{k} (a_j + b_j - 1)$. Hence the Lemma is proved if R is a rectangular region.

Now, let R be any polygonal region of the type specified and let R_1 be the corresponding rectangular region. Consider two randomly moving points Q and Q_1, each having coordinates W_t at time t. Let the decision points for Q be defined in terms of R and the decision points of Q_1 in terms of R_1. We shall prove that the number of decision points for Q and Q_1 are the same.

By assumption, every line parallel to the axes intersects R in an open interval. Moreover $R_1 \supseteq R$. Hence the sum of the areas of the segments which compose the boundary of R must equal the area of the boundary of R_1. The same must be true for the total number of integral points on the boundaries of the two regions. Thus, the theorem is true for $r_j = m_j = 1$, $(j = 1, 2, \cdots, k)$. We assume that the theorem is true for $r_j = r_j'$ and $m_j = m_j'$ and prove that it must hold for $m_u = m_u' + 1$ for a fixed but arbitrary u. Now it is obvious that if the range of X_{ui} is increased by unity in the positive direction, the point Q can move an extra unit in the positive direction parallel to the y_u axes. Thus, the total number of additional decision points that Q gains by the unit increase in the range of X_{ui} is identical with the total number that Q_1 gains. This proves the theorem.

It is clear that the smallest rectangular region which includes all the decision points of W is R_2. We now prove the following:

THEOREM 3.1. *For any polygonal region R of the class previously specified, and for any random sequence $\{X_i\}$ in which X_i takes on a finite number of integral values, the number of points in the rectangular region R_2 which are not decision points is always equal to* $\prod_{j=1}^{k} (a_j + b_j - 1)$ *where $a_j + b_j$ are the dimensions of the rectangular region R_1.*

PROOF: This Theorem follows from Lemma 3.1 and the fact that the total number of integral points in R_2 is $\prod_{j=1}^{k} (a_j + b_j + r_j + m_j - 1)$.

3.2. *The distribution of W.* Let $\psi(t_1, \cdots, t_k)$ be the joint generating function of X_{ui}, $(u = 1, 2, \cdots, k)$, and $\varphi(t_1, \cdots, t_k)$ the joint generating function of W_j $(j = 1, 2, \cdots, k)$. Then

$$\psi(t_1, \cdots, t_k) = \sum_{u=-r_1}^{m_1} \cdots \sum_{u_k=-r_k}^{m_k} P_{u_1 \cdots u_k} t_1^{u_1} \cdots t_k^{u_k} \tag{3.21}$$

$$\phi(t_1, \cdots, t_k) = \sum_{v_1=-(b_1+r_1-1)}^{a_1+m_1-1} \cdots \sum_{v_k=-(b_k+r_k-1)}^{a_k+m_k-1} \xi_{v_1 \cdots v_k} t_1^{v_1} \cdots t_k^{v_k} \tag{3.22}$$

where $\xi_{v_1 \cdots v_k}$ is the probability that $W = (v_1, \cdots, v_k)$. In terms of the generating function ψ the Fundamental Identity (3.17) states that

$$E t_1^{W_1} \cdots t_k^{W_k} [\psi(t_1, \cdots, t_k)]^{-n} = 1 \tag{3.23}$$

for all $t_1, \cdots, t_k$ for which $|\psi(t_1, \cdots, t_k)| \geq 1$. Hence, it follows that for $t_1, \cdots, t_k$ for which $\psi(t_1, \cdots, t_k) = 1$, $\varphi(t_1, \cdots, t_k) = 1$. Let

$$f(t_1, \cdots, t_k) = t_1^{r_1} \cdots t_k^{r_k} [\psi(t_1, \cdots, t_k) - 1] \tag{3.24}$$

and

$$g(t_1, \cdots, t_k) = t_1^{b_1+r_1-1} \cdots t_k^{b_k+r_k-1} [\varphi(t_1, \cdots, t_k) - 1]. \tag{3.25}$$

Then $f(t_1, \cdots, t_k)$ is a polynomial of degree $r_j + m_j$ in t_j and $g(t_1, \cdots, t_k)$ is a polynomial of degree $(a_j + b_j + r_j + m_j - 2)$ in t_j.

We shall assume that $f(t_1, \cdots, t_k)$ is an irreducible polynomial. Then, since $g(t_1, \cdots, t_k)$ vanishes for all values of $t_1, \cdots, t_k$ for which $f(t_1, \cdots, t_k)$ vanishes, it follows[4] that f is a factor of g. That is

$$g(t_1, \cdots, t_k) = f(t_1, \cdots, t_k) \sum_{s_1=0}^{a_1+b_1-1} \cdots \sum_{s_k=0}^{a_k+b_k-1} C_{s_1 \cdots s_k} t_1^{s_1} \cdots t_k^{s_k} \tag{3.26}$$

where the $C_{s_1, \cdots, s_k}$ are unknown. Equating coefficients on both sides of (3.26) we get

[4] See, for example, Bôcher [7], Theorem 7, Chapter 16.

$$\xi_{v_1-b_1-r_1+1\cdots v_k-b_k-r_k+1} = \sum_{u_1=0}^{v_1}\cdots\sum_{u_k=0}^{v_k}(P_{u_1-r_1\cdots u_k-r_k} - \delta_{u_jr_j})\,C_{v_1-r_1\cdots v_k-r_k} + \prod_{j=1}^{k}\delta_{v_j,b_j+r_j-1} \tag{3.27}$$

where δ_{ij} is the Kronecker delta. But by Theorem 3.1, $\prod_{j=1}^{k}(a_j + b_j - 1)$ of the $\xi_{v_1\cdots v_k}$ in $\varphi(t_1, \cdots, t_k)$ are zero since they correspond to values of W which are non-decision points. Hence $\prod_{j=1}^{k}(a_j + b_j - 1)$ terms in (3.27) are zero with the exception of the term $\xi_{b_1+r_1-1\cdots b_k+r_k-1}$ (corresponding to the non-decision point (0, 0)) which is -1. Hence, we have the required number of equations to solve for the unknown C's and consequently for the ξ's provided the determinant of the coefficients is different from zero.

As an illustration, let $R = R_1$, then the C's are obtained by solving the set of linear equations

$$\sum_{u_1=0}^{v_1}\cdots\sum_{u_k=0}^{v_k}\left(\prod_{j=1}^{k}\delta_{u_jr_j} - P_{u_1-r_1\cdots u_k-r_k}\right)C_{v_1-r_1\cdots v_k-r_k} = \prod_{j=1}^{k}\delta_{v_j,b_j+r_j-1} \tag{3.28}$$

where v_j takes on all integral values from r_j to $a_j + b_j + r_j - 2$ inclusive.

3.3. *The distribution of n.* For any random variable U, let $E_{v_1\cdots v_k}U$ stand for the expected value of U under the restriction that $W = (v_1, v_2, \cdots, v_k)$. Let $\varphi_1(t_1, \cdots, t_k; \tau)$ be the joint generating function of $W_1, W_2, \cdots, W_k$, and n. Then

$$\varphi_1(t_1, \cdots, t_k; \tau) = \sum_{u_1}\cdots\sum_{u_k}\xi_{u_1\cdots u_k}t_1^{u_1}\cdots t_k^{u_k}E_{u_1\cdots u_k}\tau^n. \tag{3.31}$$

Let

$$\psi_1(t_1, \cdots, t_k; \tau) = \tau\psi(t_1, t_2, \cdots, t_k) - 1 \tag{3.32}$$

where $\psi(t_1, \cdots, t_k)$ is the joint generating function of $X_{1i}, \cdots, X_{ki}$ and is given by (3.21) and let

$$\psi_2(t_1, \cdots, t_k; \tau) = \varphi_1(t_1, \cdots, t_k; \tau) - 1. \tag{3.33}$$

Then, if we fix τ so that $|\tau| \leq 1$, we see by (3.23) that for all values of $t_1, \cdots, t_k$ for which ψ_1 vanishes, ψ_2 also vanishes. Let

$$f_1(t_1, \cdots, t_k; \tau) = t_1^{r_1}\cdots t_k^{r_k}\psi(t_1, \cdots, t_k; \tau) \tag{3.34}$$

and

$$f_2(t_1, \cdots, t_k; \tau) = t_1^{b_1+r_1-1}\cdots t_k^{b_k+r_k-1}\psi_2(t_1, \cdots, t_k; \tau). \tag{3.35}$$

Then for τ fixed, f_1 is a polynomial of degree $r_j + m_j$ in t_j and f_2 is a polynomial of degree $a_j + b_j + r_j + m_j - 2$ in t_j. Since f_2 vanishes for all values of $t_1, \cdots, t_k$ for which f_1 vanishes then if f_1 is irreducible, f_1 will be a factor of f_2. That is f_2 can then be written as

$$(3.36) \qquad f_2(t_1, \cdots, t_k; \tau) = f_1(t_1, \cdots, t_k; \tau) \sum_{v_1=1}^{a_1+b_1-2} \cdots \sum_{v_k=1}^{a_k+b_k-2} d_{v_1} \cdots {}_{v_k} t_1^{v_1} \cdots t_k^{v_k}.$$

The rest of the argument is identical with that employed in section 3.3. The unknowns in the present case, however, are $\xi_{v_1 \cdots v_k} E_{v_1 \cdots v_k} \tau^n$. When $\xi_{v_1 \cdots v_k} E_{v_1 \cdots v_k} \tau_n$ is expanded in a power series in τ, the coefficient of τ^m is the probability that $W = (v_1, \cdots, v_k)$ in exactly m steps. We shall, therefore, examine the validity of the expansion of the above function in the neighborhood of $\tau = 0$.

Let us first consider the rectangular region $R = R_1$. In this case the d's are obtained from the equations

$$(3.37) \qquad \sum_{u_1=1}^{v_1} \cdots \sum_{u_k=1}^{v_k} \left(\prod_{j=1}^{k} \delta_{u_j r_j} - \tau P_{u_1 - r_1 \cdots u_k - r_k} \right) d_{v_1 - r_1 \cdots v_k - r_k} = \prod_{j=1}^{k} \delta_{v_j, b_j + r_j - 1},$$

$$(v_j = r_j, \qquad r_j + 1, \cdots, \qquad a_j + b_j + r_j - 2),$$

so that $\xi_{v_1 \cdots v_k} E_{v_1 \cdots v_k} \tau^n$ will be given as a ratio of two polynomials in τ the denominator of which will be the determinant of the coefficients of (3.37). But this determinant equals unity when $\tau = 0$. Hence the validity of the expansion is established for a rectangular region.

If R is not a rectangle, the value of the determinant of the equations in d will still be unity. This follows from the fact that the number of non-decision points in R_2 is precisely the same as the number of non-decision points contained in R_1, hence by rearranging of the equations they an be made to assume the form (3.37).

REFERENCES

[1] P. Levy, *Theorie de l'Addition des Variables Aleatoires*, Paris, 1937.

[2] J. L. Doob "Regularity properties of certain families of chance variables," *Trans. Amer. Math. Soc.*, Vol. 47 (1940), pp. 455–486.

[3] A. Wald, "Some generalizations of the theory of cumulative sums of random variables," *Annals of Math. Stat.*, Vol. 16 (1945).

[4] D. Blackwell, "On an equation of Wald", *Annals of Math Stat.*, Vol. 17 (1946).

[5] A. Wald, "On cumulative sums of random variables," *Annals of Math. Stat.*, Vol. 15, (1944).

[6] M. A. Girshick, "Contributions to the theory of sequential analysis," *Annals of Math. Stat.*, Vol. 17 (1946).

[7] Maxime Bôcher, *Introduction to Higher Algebra*, Macmillan Co.