The Annals of Probability
1976, Vol. 4, No. 4, 662–666

OSCILLATIONS OF CONTINUOUS SYMMETRIC RANDOM WALK

BY J. P. IMHOF

University of Geneva

Oscillations are defined for n steps of the random walk formed by partial sums of variables with continuous cdf. When the summands are independent, identically and symmetrically distributed, several distribution free results are obtained relative to the number of oscillations and their lengths. Analogy with the behavior of records in a random sequence is used to obtain limit laws.

1. Oscillations. The type of oscillations that we define arose as a natural tool in a constructive proof [3] for a known duality result relating the behavior of a restricted and of an absorbed random walk [4]. Consider the random walk $S_0 = 0, S_1, \cdots, S_n$ where $S_j = X_1 + \cdots + X_j$, the X_j all have a continuous cdf and n is fixed for the time being. With probability one, $S_i \neq S_j$ when $i \neq j$, so one need only consider that situation. For $0 \leqq i < k \leqq n$ denote by $S(i, k)$ the portion of the walk $(S_i, S_{i+1}, \cdots, S_k)$. It is called *up-extremal* (*down-extremal*) if $S_i \leqq S_j < S_k$ ($S_i \geqq S_j > S_k$) for $i \leqq j < k$, and *extremal* if it is so either up or down. Its *amplitude* is then $|S_k - S_i|$. Let $\alpha_0 = 0$ and definine the indices $0 < \alpha_1 < \alpha_2 < \cdots$ by letting

$$\alpha_j = \max\{k : \alpha_{j-1} < k \leqq n \text{ and } S(\alpha_{j-1}, k) \text{ is extremal}\}. \tag{1}$$

Let I and J be the indices where the maximum and minimum of $S(0, n)$ occur. The indices of first and last extremum are then respectively

$$\mu = \min\{I, J\}, \qquad \nu = \max\{I, J\}. \tag{2}$$

Write $S_{\alpha_j} = H_j$. If for instance $S_1 > 0$ and $J > 0$, H_1 is the maximum attained by the walk before it hits $(-\infty, 0)$. If $I > \alpha_1$, H_2 is the minimum attained after α_1 and before the walk hits (H_1, ∞). Thus $0 < H_1 - H_0 < |H_2 - H_1|$, and so forth until one comes to the index V_n defined by either one of the equivalent equations

$$\alpha_{V_n - 1} = \mu, \qquad \alpha_{V_n} = \nu. \tag{3}$$

Suppose for instance that $\nu = J$. Then H_{V_n+1} is the maximum attained by $S(\nu, n)$. If $V_{n+1} < n$, H_{V_n+2} is the minimum attained by $S(V_n + 1, n)$ so that $|H_{V_n} - H_{V_n-1}| > H_{V_n+1} - H_{V_n}$. This goes on until one comes to the last α which has index T_n defined by

$$\alpha_{T_n} = n.$$

Received October 10, 1974; revised December 29, 1975.
AMS 1970 *subject classifications.* Primary 60G50; Secondary 60C05.
Key words and phrases. Random walk, oscillations, fluctuation theory.

The portion $S(\alpha_{j-1}, \alpha_j)$, $0 < j \leqq T_n$, constitutes the jth *oscillation* of $S(0, n)$. It has length L_j and amplitude A_j,

$$L_j = \alpha_j - \alpha_{j-1}, \qquad A_j = |H_j - H_{j-1}| .$$

It is easily seen by adapting the above example that, irrespective of the sign of S_1 and of which is the smaller of I, J, one has in all cases

$$0 < A_1 < \cdots < A_{V_n} > A_{V_n+1} > \cdots > A_{T_n} . \tag{4}$$

Let $W_n = T_n + 1 - V_n$. In more descriptive terms, (4) shows that V_n and W_n are respectively the number of oscillations of increasing and of decreasing amplitude, the maximal oscillation $S(\mu, \nu)$ being counted in both cases, while T_n is the total number of oscillations. One has $0 < V_n, W_n, T_n \leqq n$.

2. The case of i.i.d **symmetric summands.** Suppose from now on that the variables $X_1, \cdots, X_n$ are independent with common continuous cdf F such that $F(u) = 1 - F(-u)$ for all u. A vector $Z = (Z_1, \cdots, Z_k)$ of integer valued variables $Z_j = Z_j(S_1, \cdots, S_n)$ is distribution free if its law does not depend on the particular choice of F satisfying the conditions stated. A technique for determining if Z is distribution free is due to Sparre-Andersen [7] and has been used in particular by Hobby and Pyke [2] and Burdick [1]. To a point $x = (x_1, \cdots, x_n)$ in numerical n-space one associates the $2^n n!$ points $x(\varepsilon, \sigma) = (\varepsilon_1 x_{\sigma_1}, \cdots, \varepsilon_n x_{\sigma_n})$ obtained when $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_n)$ ranges over the 2^n vectors given by all choices $\varepsilon_i = \pm 1$, $i = 1, \cdots, n$ and $\sigma = (\sigma_1, \cdots, \sigma_n)$ ranges over the $n!$ permutations of $(1, \cdots, n)$. To each such point correspond successive sums $s_0(\varepsilon, \sigma) = 0$, $s_i(\varepsilon, \sigma) = \varepsilon_1 x_{\sigma_1} + \cdots + \varepsilon_i x_{\sigma_i}$, $i = 1, \cdots, n$, giving a sample path. With the hypothesis of continuous cdf one need only consider the case of *admissible* x, when for each (ε, σ) all sums $s_i(\varepsilon, \sigma)$, $i = 0, \cdots, n$, are distinct. One may also assume all $x_i > 0$. Let $z = (z_1, \cdots, z_k)$ and let $f_x(z)$ be the number of paths generated by x for which $Z_j(s_1(\varepsilon, \sigma), \cdots, s_n(\varepsilon, \sigma)) = z_j$, $j = 1, \cdots, k$. If $f_x(z) = f(z)$ for all admissible x and for all z, the vector Z is distribution free and

$$P(Z = z) = f(z)/(2^n n!) . \tag{5}$$

We shall construct paths by juxtaposition of successive portions. Suppose one takes $y' = (y_1, \cdots, y_r)$ and $y'' = (y_{r+1}, \cdots, y_n)$, to which correspond portions of paths $(0, y_1, y_1+y_2, \cdots, y_1+\cdots+y_r)$ and $(0, y_{r+1}, y_{r+1}+y_{r+2}, \cdots, y_{r+1}+\cdots+y_n)$. Juxtaposing the two portions means forming the path $(0, y_1, y_1 + y_2, \cdots, y_1 + \cdots + y_n)$ which corresponds to juxtaposition of y' and y'' to form $y = (y_1, \cdots, y_n)$. The portion of path corresponding to $-y' = (-y_1, \cdots, -y_r)$ will be called the opposite of the portion corresponding to y'.

Introduce now for $S(0, n)$ the variables

$$N(i) = \text{"number of oscillations of length } i\text{,"} \qquad i = 1, \cdots, n .$$

We shall show that $(V_n, W_n, N(1), \cdots, N(n))$ is distribution free and obtain its law. It is easier however to formulate the argument only in terms of positive

$N(i)$'s. Let therefore $L_1 = \min\{i: 0 < i \leqq n \text{ and } N(i) > 0\}$, $L_2 = \min\{i: L_1 < i \leqq n \text{ and } N(i) > 0\}$, etc. until the sequence ends with L_Q having index $Q = \max\{i: i \leqq n \text{ and } N(i) > 0\}$. Write $M_1 = N(L_1), \cdots, M_Q = N(L_Q)$. There is clearly one to one correspondence between $(N(1), \cdots, N(n))$ and $\Phi = (Q, L_1, \cdots, L_Q, M_1, \cdots, M_Q)$. The latter describes a random T_n-partition of the path-length n, i.e. a partition into $T_n = M_1 + \cdots + M_Q$ parts of which M_1 have length $L_1, \cdots, M_Q$ have length L_Q so that $L_1M_1 + \cdots + L_QM_Q = n$. A value of Φ will be written $\varphi = [l_1^{m_1} \cdots l_q^{m_q}]$, which describes the t-partition of n corresponding to $Q = q$, $L_j = l_j$ and $M_j = m_j$ for $j = 1, \cdots, q$, with $m_1 + \cdots + m_q = t$, $l_1m_1 + \cdots + l_qm_q = n$ and $0 < l_1 < \cdots < l_q \leqq n$.

THEOREM. *Let r, s be integers $\geqq 0$, such that $r + s + 1 = t \leqq n$ and $\varphi = [l_l^{m_1} \cdots l_q^{m_q}]$ be a t-partition of n. One has*

$$P(V_n = r+1, W_n = s+1, \Phi = \varphi) = 2^{-r-s}\binom{r+s}{r}\Big/\prod_{j=1}^{q}(m_j!\, l_j^{m_j})\,. \tag{6}$$

PROOF. According to (5), it is enough to show that for (V_n, W_n, Φ) the frequency of values $(r+1, s+1, \varphi)$ for the $2^n n!$ paths generated by an admissible x is $2^n n!$ times the above, namely

$$f(r+1, s+1, \varphi) = 2^{1-t}\binom{r+s}{r}H \qquad \text{where} \tag{7}$$

$$H = \frac{1}{\prod (m_j!)}\frac{n!}{\prod (l_j!)^{m_j}}\prod [2^{l_j}(l_j - 1)!]^{m_j}\,,$$

and all products are for $j = 1, \cdots, q$.

Let $h_0 = 0$, $h_j = m_1 + \cdots + m_j$ for $j = 1, \cdots, q$ and for $k = 1, \cdots, t$ put $\lambda_k = l_j$ when $h_{j-1} < k \leqq h_j$. Let $\mathscr{B}$ be the set of all distinct repartitions of $x_1, \cdots, x_n$ into successive batches $B_1, \cdots, B_t$ of sizes $\lambda_1, \cdots, \lambda_t$. Write $B_k = \{y_{k,1}, \cdots, y_{k,\lambda_k}\}$. Call each of the $n!/\prod(\lambda_k!)$ such repartitions a φ-batching of x. In any one of them, the kth batch produces $2^{\lambda_k}\lambda_k!$ points in λ_k-space when effecting all possible sign attachments and all possible coordinate permutations in $(y_{k,1}, \cdots, y_{k,\lambda_k})$, and correspondingly $2^{\lambda_k}\lambda_k!$ path portions of length λ_k. Juxtaposing such path portions arising from $B_1, \cdots, B_t$ gives $\prod(2^{\lambda_k}\lambda_k!) = 2^n \prod(\lambda_k!)$ paths of length n. Doing so for each φ-batching yields the $2^n n!$ paths corresponding to all $x(\varepsilon, \sigma)$. Now consider separately the initial path portions of length λ_1. The frequency form of Baxter's generalized arc sin law shows ([2], formula 2.1 with $m = k = 0$) that for a given φ-batching, exactly $2^{\lambda_1 - 1}(\lambda_1 - 1)!$ of the $2^{\lambda_1}\lambda_1!$ initial portions obtained from B_1 are down-extremal, implying that $2^{\lambda_1}(\lambda_1 - 1)!$ are extremal, and similarly for the successive portions of lengths $\lambda_2, \cdots, \lambda_t$. Hence the set $\mathscr{E}$ of paths which, when cut into successive portions of lengths $\lambda_1, \cdots, \lambda_t$, have each of those extremal, has cardinality $|\mathscr{E}| = \prod[2^{\lambda_k}(\lambda_k - 1)!]n!/\prod(\lambda_k!)$. Suppose now that for each path in $\mathscr{E}$, we permute the above portions so as to reorder them according to increasing amplitudes. Let $\mathscr{O}$ be the set of distinct paths thus obtained. Because paths in $\mathscr{E}$ which differ only by permutations of whole portions of same length lead to a unique path in $\mathscr{O}$, one has $|\mathscr{O}| = |\mathscr{E}|/(m_1! \cdots m_q!)$, equal to the factor H in (7). Divide $\mathscr{O}$ into equivalence classes, each containing 2^t paths which differ only by some of their 1st, $\cdots$, tth portions being opposite. To complete the proof of (7), one just has to notice that each equivalence class permits the construction of $2\binom{r+s}{r}$ paths of type $(r+1, s+1, \varphi)$. In fact, one must choose s among the $r+s$ initial portions, cut them out and reinsert them in reverse order (so as to have decreasing amplitudes) behind the portion of maximal amplitude, and for each choice of the s portions exactly 2 paths in the equivalence class will produce, after said reshuffling, a path with alternately up and down-extremal portions (which therefore become oscillations), the first one being of either type. Finally, each path of type $(r+1, s+1, \varphi)$ is uniquely obtained by this construction.

The joint law of (V_n, W_n) and the law of T_n are now easily deduced from (6). For the (absolute) Stirling numbers of first kind,

$$D_{t,n} = \text{coeff. of } z^t \text{ in } z(z+1)\cdots(z+n-1)\,, \qquad 1 \leqq t \leqq n\,,$$

a classical expression is ([6], page 183)

$$\frac{1}{n!} D_{t,n} = \sum_{\varphi} ({}_1m! \cdots m_q!\, l_1^{m_1} \cdots l_q^{m_q})^{-1}\,,$$

where the sum is over all t-partitions $\varphi = [l_1^{m_1} \cdots l_q^{m_q}]$ of n. Summing (6) over the latter yields therefore:

COROLLARY 1. *For integers r, $s \geqq 0$ with $r+s+1 \leqq n$ one has*

$$P(V_n = r+1,\, W_n = s+1) = 2^{-r-s}\binom{r+s}{r} D_{r+s+1,n}/(n!)\,. \tag{8}$$

Convoluting for $r+s+1 = t$ gives next

COROLLARY 2. *For $t = 1, \cdots, n$,*

$$P(T_n = t) = D_{t,n}/(n!)\,. \tag{9}$$

Let $\mathscr{L}^S(T_n)$ designate the law (9) of T_n under the hypotheses of the present section, and $\mathscr{L}^R(U_n)$ designate the law of the number U_n of upper records in the random sequence $Y_1, \cdots, Y_n$ of i.i.d variables with arbitrary continuous cdf, where Y_1 is counted as a record. One sees, e.g. from [5], that

$$\mathscr{L}^S(T_n) = \mathscr{L}^R(U_n)\,. \tag{10}$$

3. Limit laws. For the number U_n of records in the above sequence $Y_1, \cdots, Y_n$, Rényi [5] noticed that if Z_k is 1 or 0 according as Y_k is a record value or not, the summands in $U_n = Z_1 + \cdots + Z_n$ are independent with $E(Z_k) = k^{-1}$, $\operatorname{Var}(Z_k) = k^{-2}(k-1)$. Thus $E(U_n) \sim \log n$, $\operatorname{Var}(U_n) \sim \log n$ and as $[\operatorname{Var}(U_n)]^{-\frac{3}{2}} \times \sum_1^n E|Z_k - k^{-1}|^3 \to 0$, Liapounov's theorem applies. It follows from (10) that under the hypotheses of Section 2 $(T_n - \log n)(\log n)^{-\frac{1}{2}}$ also has standard normal limit law. One can proceed similarly for V_n. Consider independent pairs (Z_k', Z_k'') of 0–1 variables Z_k' and Z_k'', $k = 2, 3, \cdots$ with law given by $P(Z_k' = Z_k'' = 0) = k^{-1}(k-1)$, $P(Z_k' = 0, Z_k'' = 1) = 0$ and $P(Z_k' = 1, Z_k'' = 0) = P(Z_k' = 1, Z_k'' = 1) = (2k)^{-1}$. Let $Z_k^* = Z_k' Z_k''$.

LEMMA. *One has for* $n > 1$

$$\mathscr{L}^S(V_n) = \mathscr{L}(1 + Z_2^* + \cdots + Z_n^*) .$$

PROOF. One has $\mathscr{L}(Z_k') = \mathscr{L}^R(Z_k)$, where $Z_1, \cdots, Z_n$ are the variables defined above for which $U_n = Z_1 + \cdots + Z_n$, and $Z_1 = 1$ always. It follows from (10) that $\mathscr{L}^S(T_n) = \mathscr{L}(1 + Z_2' + \cdots + Z_n')$. One can thus write

$$\begin{aligned} P(Z_2^* + \cdots Z_n^* = r) & \\ &= \sum_{t=r+1}^n P(Z_2' + \cdots + Z_n' = t - 1) \\ &\qquad \times P(Z_2'' + \cdots + Z_n'' = r | Z_2' + \cdots + Z_n' = t - 1) \\ &= \sum_{t=r+1}^n P(T_n = t) 2^{-t+1} \binom{t-1}{r} , \end{aligned}$$

which according to (8) and (9) equals $P(V_n = r + 1)$.

The variables $Z_2^*, Z_3^*, \cdots$ are independent with $E(Z_k^*) = (2k)^{-1}$, $\mathrm{Var}(Z_k^*) = (2k - 1)(4k^2)^{-1}$. There follows $E(V_n) \sim \frac{1}{2}\log n$, $\mathrm{Var}(V_n) \sim \frac{1}{2}\log n$. Furthermore, Liapounov's condition is satisfied for the Z_k^* as it was for the Z_k. Taking into account that $\mathrm{Var}(V_n) + \mathrm{Var}(W_n) \sim \mathrm{Var}(T_n)$, one concludes therefore

THEOREM. $(V_n - \frac{1}{2}\log n)(\frac{1}{2}\log n)^{-\frac{1}{2}}$ *and* $(W_n - \frac{1}{2}\log n)(\frac{1}{2}\log n)^{-\frac{1}{2}}$ *have independent asymptotic standard normal laws.*

We do not know if, analogously to the result for records, $V_n/(\frac{1}{2}\log n)$ converges to 1 a.s. This is clearly not the case for $W_n/(\frac{1}{2}\log n)$, as W_n drops to 1 whenever S_n reaches a new extremum.

We are much indebted to the referee whose detailed comments and criticisms led to a drastic revision of the initial manuscript.

REFERENCES

[1] BURDICK, D. L. (1972). Distribution free tests for symmetry based on the number of positive sums. *Ann. Math. Statist.* **43** 428–438.

[2] HOBBY, C. and PYKE, R. (1963). Combinatorial results in fluctuation theory. *Ann. Math. Statist.* **34** 1233–1242.

[3] IMHOF, J.-P. (1976). Oscillations of monotone amplitude and application. To appear in *Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes.*

[4] PHATARFORD, R. M., SPEED, T. P. and WALKER, M. (1971). A note on random walks. *J. Appl. Probability* **8** 198–201.

[5] RÉNYI, A. (1962). Théorie des éléments saillants d'une suite d'observations. *Colloquium on Combinatorial Methods in Probability Theory, Aarhus* 104–115.

[6] RIORDAN, J. (1968). *Combinatorial identities.* Wiley, New York.

[7] SPARRE-ANDERSEN, E. (1949). On the number of positive sums of random variables. *Skand. Aktuarietidskr.* **36** 27–36.

SECTION DE MATHÉMATIQUES
UNIVERSITY OF GENEVA
1211 GENEVA 24