The Annals of Mathematical Statistics
1968, Vol. 39, No. 1, 258–260

SOME INVARIANT LAWS RELATED TO THE ARC SINE LAW

BY. J. P. IMHOF

University of Geneva

1. Introduction and summary. Let $\{X_i : i = 1, 2, \cdots\}$ be a sequence of random variables such that $X_1, \cdots, X_n$ are exchangeable and symmetric ($n = 1, 2, \cdots$). Suppose that ties occur with probability zero among the partial sums $S_0 = 0$, $S_k = \sum_1^k X_i$. We study the laws of the variables

J_k, the index of the kth positive sum in the sequence $S_1, S_2, \cdots$ ($k = 1, 2, \cdots$),

N_n', the number of positive sums among $S_0, S_1, \cdots, S_{L_n}$, where L_n is the index of max $\{S_0, S_1, \cdots, S_n\}$.

Brief attention is given to J_k in [2], where the simple form of its law in the symmetric case is however not mentioned. The variable N_n' does not seem to have been considered before. Setting

$$a_k = 2^{-2k}\binom{2k}{k}, \qquad k = 0, 1, 2, \cdots \ (a_0 = 1),$$

we find the probabilities

$$(1.1) \quad q_k(n) = P[J_k = n] = (k/n)a_k a_{n-k}, \qquad n = k, k+1, \cdots,$$

$$(1.2) \quad p_n(i) = P[N_n' = i] = (2ia_i)^{-1}a_n, \qquad i = 1, 2, \cdots, n \ (p_n(0) = a_n).$$

Let $\{X_t, 0 \leqq t \leqq T < \infty\}$ be a measurable, separable stochastic process which is continuous in probability and has exchangeable, symmetric increments. Relative to the bounded time interval $0 \leqq t \leqq T$, introduce the variables

(1.3) U = "time spent in the positive half plane up to the moment when the process reaches its maximum,"

V = "time elapsed until the process reaches its maximum."

Asymptotic evaluations lead to

THEOREM. *U/V is independent of V/T, and for $0 \leqq \alpha, \gamma \leqq 1$,*

$$P[U < \alpha V] = 1 - (1-\alpha)^{\frac{1}{2}}, \qquad P[U < \gamma T] = \gamma^{\frac{1}{2}}.$$

2. The discrete laws. The result (1.1) follows from a formula of Andersen. Using $\binom{\frac{1}{2}}{j} = (-1)^{j-1}a_{j-1}/2j$, $\binom{-\frac{1}{2}}{j} = (-1)^j a_j$, the identity (5.16) of [1] becomes

$$(2.1) \qquad \sum_{s=n-m+1}^{n} (2s)^{-1}a_{s-1}a_{n-s} = (m/n)a_m a_{n-m}, \qquad 0 < m \leqq n.$$

If N_n is the number of positive sums among $S_1, \cdots, S_n$, then from the arc sine law it is known that $P[N_n = i] = a_i a_{n-i}$. Since the event $[J_k > n]$ is equivalent to $[N_n < k]$, it follows that

$$P[J_k > n] = \sum_{i=0}^{k-1} a_i a_{n-i},$$

$$P[J_k = n] = \sum_{i=0}^{k-1} a_i(a_{n-i-1} - a_{n-i}) = \sum_{i=0}^{k-1} (2n - 2i)^{-1} a_i a_{n-i-1}$$

$$= \sum_{s=n-k+1}^{n} (2s)^{-1} a_{s-1}a_{n-s} = (k/n)a_k a_{n-k}.$$

Received 11 April 1967; revised 19 September 1967.

Notice that (1.1) is equivalent to saying that, conditionally on $N_n = k$, the probability is k/n that S_n be one of the k positive sums. The divergence of $\sum_0^\infty a_n$ implies $EJ_k = \infty, k = 1, 2, \cdots$

In order to establish (1.2), we first show that

$$a_n = ia_i \sum_{j=i}^{n} j^{-1} a_{j-i} a_{n-j}, \qquad 0 < i \leqq n. \tag{2.2}$$

When $i = 1$, this coincides with (2.1) in which $m = 1$. We must only check that $R_n(i)$, the right-hand member above, does not depend on i. Using the equality $(2i + 2)a_{i+1} = (2i + 1)a_i$, one has

$$\begin{aligned}(2/a_i)\,[R_n(i) - R_n(i+1)] &= 2i \sum_{j=i}^{n} j^{-1}\, a_{j-i}a_{n-j} \\ &\quad -(2i+1) \sum_{j=i+1}^{n} j^{-1} a_{j-i-1}\, a_{n-j} \\ &= 2a_0a_{n-i} + 2i \sum_{j=i+1}^{n} j^{-1}(a_{j-i} - a_{j-i-1})a_{n-j} \\ &\quad - \sum_{j=i+1}^{n} j^{-1} a_{j-i-1}a_{n-j} \\ &= 2a_{n-i} - i \sum_{j=i+1}^{n} (j(j-i))^{-1} a_{j-i-1}a_{n-j} \\ &\quad - \sum_{j=i+1}^{n} j^{-1} a_{j-i-1}a_{n-j}\,.\end{aligned}$$

Grouping the last two sums and putting $s = j - i$, it follows from (2.1) that

$$a_i^{-1}[R_n(i) - R_n(i+1)] = a_{n-i} - \sum_{s=1}^{n-i} (2s)^{-1} a_{s-1}a_{n-i-s} = 0.$$

Consider now the joint law of (N_n', L_n). For $0 < i \leqq j \leqq n$, the event $[N_n' = i, L_n = j]$ occurs if and only if (a) i of the sums $S_1, \cdots, S_j$ are positive, (b) all sums $S_0, S_1, \cdots, S_{j-1}$ are less than S_j and (c) $S_k - S_j < 0$ for $k = j + 1, \cdots, n$. The probability of simultaneously having (a) and (b) is, according to Baxter's generalized arc sine law [3], $(2j)^{-1}a_{j-i}$. The probability of (c) is a_{n-j}. If $X_1, \cdots, X_n$ are independent, one has therefore

$$p_n(i,j) = P[N_n' = i, L_n = j] = (2j)^{-1}a_{j-i}a_{n-j}, \qquad 0 < i \leqq j < n. \tag{2.3}$$

For exchangeable variables, the same result would be obtained by "counting paths", as in [3]. Finally, $p_n(i) = \sum_{j=i}^{n} p_n(i,j)$ gives, according to (2.2), the result (1.2).

If $i > 1$, one has $p_n(i-1)/p_n(i) = (2i-1)/(2i-2)$, so that

$$a_n = p_n(0) = p_n(1) > p_n(2) > \cdots > p_n(n) = (2n)^{-1}.$$

It is easy to verify that $EN_n' = \frac{1}{2} + (2n-1)(2n)^{-1}EN_{n-1}'$. The general solution of this difference equation is $EN_n' = \frac{1}{3}(n + 1 + ca_n)$, and $EN_1' = \frac{1}{2}$ gives $c = -1$:

$$EN_n' = \tfrac{1}{3}(n + 1 - a_n).$$

When $n \to \infty$, one obtains $EN_n' \sim \frac{1}{3}(n+1)$. If T_n is the number of ladder sums among $S_1, \cdots, S_n$, this can be compared [4] with $ET_n = (2n+2)a_{n+1} - 1 \sim 2[(n+1)/\pi]^{\frac{1}{2}}$.

3. Asymptotic results. Using a standard argument, one deduces from $a_n \sim$

$(\pi n)^{-\frac{1}{2}}$ that for $0 < \gamma < 1$, if $[x]$ is the integral part of x,

$$P[N_n' < \gamma n] = a_n + \sum_{i=1}^{[\gamma n]} (2ia_i)^{-1} a_n \sim \sum_{i=1}^{[\gamma n]} (i/n)^{-\frac{1}{2}} (2n)^{-1}.$$

It follows that

$$\lim_{n\to\infty} P[N_n' < \gamma n] = \tfrac{1}{2} \int_0^\gamma t^{-\frac{1}{2}}\, dt = \gamma^{\frac{1}{2}}. \tag{3.1}$$

In the same fashion, if $0 < \alpha, \beta < 1$, we derive from (2.3) the asymptotic evaluation

$$P[N_n' < \alpha L_n, L_n < \beta n] \sim (2\pi)^{-1} \sum_{j=1}^{[\beta n]} \sum_{i=1}^{[\alpha j]} j^{-1}[(n-j)(j-i)]^{-\frac{1}{2}}$$
$$= (2\pi)^{-1} \sum_{j=1}^{[\beta n]} \sum_{i=1}^{[\alpha j]} (j/n)^{-1}(1-j/n)^{-\frac{1}{2}}(j/n - i/n)^{-\frac{1}{2}}(1/n)^2.$$

It follows that

$$\lim_{n\to\infty} P[N_n' < \alpha L_n, L_n < \beta n]$$
$$= (2\pi)^{-1} \int_0^\beta dx x^{-1} (1-x)^{-\frac{1}{2}} \int_0^{\alpha x} dy (x-y)^{-\frac{1}{2}}$$
$$= [1 - (1-\alpha)^{\frac{1}{2}}] 2\pi^{-1} \text{ arc sin } (\beta^{\frac{1}{2}}) \tag{3.2}$$
$$= \lim_{n\to\infty} P[N_n' < \alpha L_n] P[L_n < \beta n].$$

Now, let t denote time and $\{X_t, 0 \leqq t \leqq T < \infty\}$ be a measurable, separable stochastic process which is continuous in probability and has exchangeable, symmetric increments. For the variables U and V defined in (1.3), (3.1) implies that $P[U < \gamma T] = \lim P[N_n' < \gamma n] = \gamma^{\frac{1}{2}}$, while (3.2) shows that $P[U < \alpha V] = \lim P[N_n' < \alpha L_n] = 1 - (1-\alpha)^{\frac{1}{2}}$, $P[V < \beta T] = 2\pi^{-1}$ arc sin $(\beta^{\frac{1}{2}})$, and U/V is independent of V/T. This proves the theorem stated in Section 1. It provides an illustration to Theorem 1 of [5], relative to products of independent beta variables which have again a beta distribution.

4. Acknowledgment. We had initially derived (1.1) and (1.2) from recurrence relations, which can be obtained much like is done in [3] for Baxter's generalized arc sine probabilities. We thank the referee who drew our attention to Andersen's formula (2.1), showed how it implies (1.1) and suggested obtaining (1.2) from it.

REFERENCES

[1] ANDERSEN, E. SPARRE (1953). On the fluctuations of sums of random variables. *Math. Scand.* **1** 263–285.

[2] BAXTER, GLEN (1961). An analytic approach to finite fluctuation problems in probability. *J. Analyse Math.* **9** 31–70.

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

[4] IMHOF, J. P. (1967). On ladder indices and random walk. To appear in *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.*

[5] JAMBUNATHAN, M. V. (1954). Some properties of beta and gamma distributions. *Ann. Math. Statist.* **25** 401–405.