The Annals of Mathematical Statistics
1971, Vol. 42, No. 3, 1099-1103

SOME JOINT LAWS IN FLUCTUATION THEORY

BY J. P. IMHOF
University of Geneva

1. Summary. Let $S_0 = 0$, $S_n = \sum_1^n X_i$, $n = 1, 2, \cdots$ be the sequence of partial sums of independent, identically distributed real valued random variables $X_1, X_2, \cdots$ The sum S_k is a (strict) ladder sum if $S_j < S_k$ for $0 \leqq j < k$. Contrary to usual practice, we always count S_0 as a ladder sum. A run of ladder sums of length i starts at S_k if $S_k, S_{k+1}, \cdots, S_{k+i-1}$ are ladder sums but S_{k-1} and S_{k+i} are not. Thus S_0 is always the beginning of a run, of length one if $S_1 \leqq 0$. Except at the beginning of Section 3, we assume that X_1 has continuous law, symmetric with respect to 0. Thus ties $S_i = S_j$ can be disregarded. Let

$$L_n = \text{"index of } \max\{S_i, 0 \leqq i \leqq n\}\text{,"}$$
$$G_n = \text{"number of ladder sums among } 0, S_1, \cdots, S_n\text{,"} \tag{1}$$
$$R_n = \text{"number of runs they form."}$$

Two sets of probabilities concerning those variables are obtained, namely the joint law

$$p_n(k+1, m+1) = P(G_n = k+1, R_n = m+1) = 2^{-2n+k}\binom{2n-2k}{n-k-m}\binom{k}{m} \tag{2}$$

where $0 \leqq m \leqq k \leqq n$ and $k+m \leqq n$, and the probabilities

$$p_n^*(k+1, m+1) = P(G_n = k+1, R_n = m+1, L_n = n) = \frac{m}{n-k}\, p_n(k+1, m+1), \tag{3}$$

valid for $1 \leqq m \leqq k \leqq n$ and $k+m \leqq n$. If $m = 0$ and $0 \leqq k < n$, then clearly $p_n^*(k+1, 1) = 0$, while on the other hand $p_n^*(n+1, 1) = P(0 < S_1 < \cdots < S_n) = 2^{-n}$.

Considering similarly the behavior beyond the maximum, let $X_i' = -X_{n-i+1}$, $S_i' = \sum_{k=1}^i X_k' = S_{n-1} - S_n$, $i = 1, \cdots, n$, and

$$G_n' = \text{"number of ladder sums among } 0, S_1', \cdots, S_n'\text{,"} \tag{4}$$
$$R_n' = \text{"number of runs they form."}$$

so that G_n' counts descending ladders from S_{L_n} on. Various limit laws are obtained, in particular it is shown that $(2L_n)^{-\frac{1}{2}}G_n$, $(2n-2L_n)^{-\frac{1}{2}}G_n'$, $n^{-1}L_n$, $G_n^{-\frac{1}{2}}(2R_n - G_n)$ and $G_n'^{-\frac{1}{2}}(2R_n' - G_n')$ are asymptotically independent, the first two having a limiting χ_2 law and the latter two a standard normal one.

2. Derivation via difference equations. A useful technique for obtaining various invariant probabilities is fully described by Hobby and Pyke [2], and applied to

Received May 14, 1970; revised November 16, 1970.

another problem in [3]. Briefly stated, one takes a numerical vector $x = (x_1, \cdots, x_n)$ and considers the set $\{x_{\sigma,\varepsilon} = (\varepsilon_1 x_{\sigma_1}, \cdots, \varepsilon_n x_{\sigma_n})\}$ of $2^n n!$ vectors derived from x by all possible sign attachments $\varepsilon_i = \pm 1$ and all possible coordinate permutations $\sigma = (\sigma_1, \cdots, \sigma_n)$. Each $x_{\sigma,\varepsilon}$ generates a sample path in the Cartesian plane, going from the origin successively to the points $(k, \sum_1^k \varepsilon_i x_{\sigma_i})$, $k = 1, \cdots, n$. In all those paths, one may assume the partial sums $\sum_1^k \varepsilon_i x_{\sigma_i}$ are all different and $\neq 0$. Let $u_n(k+1, m+1)$ and $u_n^*(k+1, m+1)$ be the numbers of paths fulfilling respectively the conditions described in (2) and (3): then $2^n n! p_n(k+1, m+1) = u_n(k+1, m+1)$, and similarly with the *. The method now consists in obtaining a recurrence relation by examining how all paths of length n counted e.g. by $u_n(k+1, m+1)$ can be obtained from paths of length $n-1$ through insertion of a minutely ascending, or descending, segment. One finds in this fashion

$$(5) \qquad u_n(k+1, m+1) = (k-m+1)u_{n-1}(k+1, m) + ku_{n-1}(k, m+1) + (2n-2k+m-1)u_{n-1}(k+1, m+1),$$

valid for $n > 1$ with the initial values $u_1(1, 1) = u_1(2, 1) = 1$, and

$$(6) \qquad u_n^*(k+1, m+1) = (k-m+1)u_{n-1}^*(k+1, m) + ku_{n-1}^*(k, m+1) + (2n-2k+m-2)u_{n-1}^*(k+1, m+1),$$

valid for $n > 2$ and $k \geqq 1$, with the initial conditions $u_2^*(1, 1) = u_2^*(2, 1) = 0$, $u_2^*(2, 2) = 1$ and $u_2^*(3, 1) = 2$. In each case, terms on the right for which the restrictions on k and m (relative to $n-1$) listed for (2) and (3) fail to hold are set equal to zero. One checks easily that the solutions are respectively

$$u_n(k+1, m+1) = 2^{-n+k} n! \binom{2n-2k}{n-k-m}\binom{k}{m}, \qquad u_n^*(k+1, m+1) = \frac{m}{n-k} u_n(k+1, m+1),$$

so that (2) and (3) are established.

The marginal law of G_n, and the marginal probabilities in (3) for fixed k, are well known:

$$(7) \qquad P(G_n = k+1) = q_n(k+1) = 2^{-2n+k}\binom{2n-k}{n},$$

$$(8) \qquad \sum_m p_n^*(k+1, m+1) = P(G_n = k+1, L_n = n) = q_n^*(k+1) = \frac{k}{2n-k} q_n(k+1).$$

It does not seem, on the other hand, that the law of R_n is expressible in compact form. However, using (3) and (8) one finds after some algebra that $E(R_n) = 1 + na_n - (2n)^{-1}E(G_{n-1}^2)$, where $a_n = 2^{-2n}\binom{2n}{n}$. Results in [1] then yield

$$E(R_n) = (n+1)a_n \sim (n/\pi)^{\frac{1}{2}}.$$

It is known that $E(G_n) = (2n+1)a_n$, thus as must be $\lim \{E(G_n)/E(R_n)\} = 2$, the expected length of a run in the infinite sequence $\{S_n\}$.

3. Use of generating functions. We drop for a while the assumption of symmetric law for the X_i. If E_n is the event "S_n is a ladder sum," $\{E_n, n > 0\}$ defines a recurrent event. Let

$$f_1 = P(E_1) = P(X_1 > 0), \qquad f_j = P(E_1{}^c \cdots E_{j-1}^c E_j) \quad \text{or} \quad j > 1, \qquad q_j = \sum_{j+1}^{\infty} f_k,$$
$$F(s) = \sum_1^{\infty} f_j s^j, \qquad Q(s) = \sum_0^{\infty} q_j s^j = (1-s)^{-1}\{1-F(s)\},$$

where we have assumed in the last equality that the recurrent event is persistent, i.e. $q_0 = 1$.

Notice first that all probabilities considered can be obtained from the case $k = m$. In fact

$$p_n(k+1, m+1) = f_1{}^{k-m}\tbinom{k}{m} p_{n-k+m}(m+1, m+1), \tag{9}$$

and the same relation holds for the $p_n{}^*$, because there are $\binom{k}{m}$ ways of splitting $k-m$ X_i's into $m+1$ batches (some maybe empty) which are then inserted after each beginning of a run for the sums $0, S_1, \cdots, S_{n-k+m}$ where all runs are assumed of length one. This will provide, for n summands, $m+1$ runs with a total of $k+1$ ladders if the $k-m$ inserted X_i's are positive, for which the probability is $f_1{}^{k-m}$. On the other hand, $p_n(m+1, m+1)$ can be determined by reviewing the indices $0, i_1, i_1+i_2, \cdots, i_1+\cdots+i_m$ of occuring E_j's; the occurrence of E_j excludes that of E_{j+1}, hence all i_j are greater than one and

$$p_n(m+1, m+1) = \sum_{\left\{ i_1+\cdots+i_m+j=n \atop i_1>1,\cdots,i_m>1 \right\}} f_{i_1} f_{i_2} \cdots f_{i_m} q_j.$$

The factor q_j guarantees that ladder number $m+2$ occurs beyond n. The $p_n{}^*(m+1, m+1)$ correspond to $j = 0$. Hence the generating functions of the $p_n(m+1, m+1)$ and $p_n{}^*(m+1, m+1)$, $n \geqq 2m$, are respectively

$$g_m(s) = (F(s)-f_1 s)^m Q(s), \qquad g_m{}^*(s) = (F(s)-f_1 s)^m, \tag{10}$$

the first for $m \geqq 0$ and the second for $m > 0$. Notice also that a review of possible values of L_n gives, with the notation of (7) and the value $q_0(1) = 1$,

$$p_n(m+1, m+1) = \sum_{i=2m}^{n} p_i{}^*(m+1, m+1) q_{n-i}(1), \quad 0 < m, \quad 2m \leqq n. \tag{11}$$

Returning to the symmetric case where $f_1 = \frac{1}{2}$, $F(s) = 1-(1-s)^{\frac{1}{2}}$ and $Q(s) = (1-s)^{-\frac{1}{2}}$, (2) and (3) imply for $|s| < 1$ the expansions

$$\{1-\tfrac{1}{2}s-(1-s)^{\frac{1}{2}}\}^m (1-s)^{-\frac{1}{2}} = \sum_{n=2m}^{\infty} 2^{-2n+m}\tbinom{2n-2m}{n} s^n,$$

$$\{1-\tfrac{1}{2}s-(1-s)^{\frac{1}{2}}\}^m = \sum_{n=2m}^{\infty} 2^{-2n+m} \frac{m}{n-m}\tbinom{2n-2m}{n} s^n.$$

If one could obtain those directly, the detour via (5) and (6) would not be necessary. The identity (11) now becomes the apparently non-standard binomial identity

$$\binom{2n-2m}{n} = \sum_{i=2m}^{n} \frac{m}{i-m}\binom{2i-2m}{i}\binom{2n-2i}{n-i}, \qquad 0 < m, \quad 2m \leqq n.$$

Quite generally, finding the probability of a specific pattern concerning lengths of runs amounts to solving a corresponding occupancy problem for $k-m$ balls placed into $m+1$ boxes. Thus for instance, if $0 \leqq m < k \leqq n$ and $k+m \leqq n$, one has in the case $f_1 = \frac{1}{2}$

$$P(G_n = k+1,\ R_n = m+1 \quad \text{of which } m-r \text{ have length one})$$
$$2^{-2n+k}\binom{2n-2k}{n-k-m}\binom{m+1}{m-r}\binom{k-m-1}{r}, \qquad r = 0, 1, \cdots, k-m-1.$$

4. Limit laws. The results above for the symmetric case are directly applicable, as is well known, to $2n$ steps of simple, symmetric random walk provided one translates ladder sum into return to equilibrium and run of ladders into run of returns to equilibrium (two steps apart). One can however obtain from (3) more detailed information involving both the variables (1) and (4), which does not have an equivalent for random walk. Clearly,

$$\begin{aligned} &P(G_n = k+1, R_n = m+1, L_n = j, G_n' = k'+1, R_n' = m'+1) \\ (12)\qquad &= P(G_j = k+1, R_j = m+1, L_j = j)P(G_{n-j} = k'+1, R_{n-j} = m'+1, \\ &L_{n-j} = n-j), \end{aligned}$$

with the obvious restrictions on the arguments. Let $U_n = (2n)^{-\frac{1}{2}}G_n$, $U_n' = (2n)^{-\frac{1}{2}}G_n'$, $V_n = G_n^{-\frac{1}{2}}(2R_n - G_n)$, $V_n' = G_n'^{-\frac{1}{2}}(2R_n' - G_n')$, $T_n = n^{-1}L_n$ and for $0 < a < b < 1$, $0 < a' < b' < 1$ consider the event

$$A_n\colon a < U_n < b, \qquad a' < U_n' < b', \qquad c < V_n < d,$$
$$c' < V_n' < d', \qquad e < T_n < f.$$

When k, m and j satisfy in (12) the conditions imposed by A_n, one has for $n \to \infty$ the asymptotic equivalence $2m \sim k$, $(j-k)^{-1}m \sim \frac{1}{2}j^{-1}k$ and also, from standard results in [1],

$$2^{-2j+2k}\binom{2j-2k}{j-k-m} \sim (j\pi)^{-\frac{1}{2}}\exp\left\{-\frac{m^2}{j}\right\}, \qquad 2^{-k}\binom{k}{m} \sim (\tfrac{1}{2}\pi k)^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\frac{(2m-k)^2}{k}\right\}$$

all holding uniformly within A_n. Substituting these expressions into the first right-hand factor of (12) as given by (3) and (2), proceeding similarly with the second factor, then taking sums as specified by A_n and passing to the limit, one obtains

$$\lim P(A_n) = \frac{1}{\pi}\int_a^b du\int_{a'}^{b'} du'\int_e^f dt\ \frac{uu'}{[t(1-t)]^{\frac{3}{2}}}\exp\left\{-\frac{1}{2}\left(\frac{u^2}{t}+\frac{u'^2}{1-t}\right)\right\}$$
$$\cdot\frac{1}{2\pi}\int_c^d \exp(-\tfrac{1}{2}v^2)\,dv\int_{c'}^{d'} \exp(-\tfrac{1}{2}v'^2)\,dv'.$$

Thus V_n, V_n' and the set (U_n, U_n', T_n) are asymptotically independent, V_n and V_n' have limiting $N(0, 1)$ law while the limit law of (U_n, U_n', T_n) is given by the density

$$p(u, u', t) = \frac{1}{\pi}\frac{uu'}{[t(1-t)]^{\frac{3}{2}}}\exp\left\{-\frac{1}{2}\left(\frac{u^2}{t}+\frac{u'^2}{1-t}\right)\right\}, \qquad 0 \leqq u, u', 0 \leqq t \leqq 1. \tag{13}$$

Integration in u', then either in u or in t yields the known limiting arcsine and χ_1 densities for T_n and U_n respectively. Performing the change of variables $w = t^{-\frac{1}{2}}u$, $w' = (1-t)^{-\frac{1}{2}}u'$, one sees that $W_n = (2L_n)^{-\frac{1}{2}}G_n$, $W_n' = (2n-2L_n)^{-\frac{1}{2}}G_n'$ and T_n are also asymptotically independent, W_n and W_n' having limiting χ_2 law with density $p(w) = w \exp(-\frac{1}{2}w^2)$, $w \geqq 0$.

It is not obvious how (13) can be integrated in t, so we derive the result *ab initio*. On the set of $2^n n!$ paths described in Section 2, consider the map which sends each $(\varepsilon_1 x_{\sigma_1}, \cdots, \varepsilon_n x_{\sigma_n}) = (y_1, \cdots, y_n)$ into $(y_1, \cdots, y_{L_n}, -y_n, -y_{n-1}, \cdots, -y_{L_n+1})$. It shows that $P(G_n = k+1, G_n' = k'+1) = P(G_n = k+k'+1, L_n = n)$. Using (8) and (7), then passing to the limit as above, one obtains for (U_n, U_n') the limiting density

$$p(u, u') = \left(\frac{2}{\pi}\right)^{\frac{1}{2}}(u+u')\exp\{-\tfrac{1}{2}(u+u')^2\}, \qquad u, u' \geqq 0.$$

Simple calculations show that the corresponding correlation coefficient is $\rho = (\pi-4)(2\pi-4)^{-1}$, and that the limit laws of $U_n + U_n'$, $U_n - U_n'$ are respectively χ_3 and $N(0, 1)$. Finally, convergence in law of $U_n^{-\frac{1}{2}}V_n = (2n)^{\frac{1}{4}}(1-2R_n/G_n)$ implies that $2R_n/G_n \to 1$ in probability.

REFERENCES

[1] FELLER, WILLIAM (1967). *An Introduction to Probability Theory and its Applications* **1**, 2nd ed. Wiley, New York.

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

[3] IMHOF, JEAN-PIERRE (1967). On ladder indices and random walk. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **9** 10–15.