ON THE FLUCTUATIONS OF SUMS OF RANDOM VARIABLES II

ERIK SPARRE ANDERSEN

1. Let $X_1, X_2, \ldots$ be random variables and let $S_i = X_1 + \ldots + X_i$, $i = 1, 2, \ldots$. We shall study here some random variables which are defined in terms of the sums $S_0 = 0, S_1, S_2, \ldots$. These random variables are:

$1°$ the index L_n at which S_i, $i = 0, \ldots, n$, attains, for the first time, the maximum value $\max(S_0, S_1, \ldots, S_n)$,

$2°$ the index M_n at which S_i, $i = 0, \ldots, n$, attains, for the last time, the minimum value $\min(S_0, S_1, \ldots, S_n)$,

$3°$ the number N_n of sums $S_1, \ldots, S_n$, which are > 0,

$4°$ the number H_n of values i, $i = 1, \ldots, n-1$, for which S_i coincides with the largest convex minorant of the sequence $S_0, S_1, \ldots, S_n$.

The variables L_n, M_n and N_n have been treated in several papers, see e. g. P. Lévy [10, in particular pp. 303–304], P. Erdös and M. Kac [8], K. L. Chung and W. Feller [6], E. Sparre Andersen [1][2], D. A. Darling [7], G. Maruyama [13], M. Lipschutz [11][12], M. Udagawa [14], E. Sparre Andersen [3][4]. In sections 3–5 of the present paper we obtain certain new results for these variables in the case where the variables X_i are independent and have the same distribution function $F(x)$. In particular, we obtain explicit expressions for the probabilities $\Pr\{K_n = m\}$ and the conditional probabilities $\Pr\{K_n = m \mid S_{n+1} > 0\}$, where K_n is one of the random variables L_n, M_n or N_n, in terms of the probabilities $a_i = \Pr\{S_i > 0\}$. In the case in which a_n converges to a limit a, $0 < a < 1$, we obtain the limiting distribution of K_n/n, which for $a = \frac{1}{2}$ becomes the arcsine law found by Erdös and Kac [8] for independent random variables, which are not necessarily identically distributed, but which satisfy the relations $E(X_n) = 0$, $E(X_n^2) = 1$ and to which the central limit theorem is applicable. Our result for $a = \frac{1}{2}$ contains the result found by Udagawa [14], but does not contain Erdös and Kac's theorem, since we have here assumed that $X_1, X_2, \ldots$ are identically distributed.

In section 6 a result is proved which supplements the results in [3] on symmetrically dependent random variables.

In sections 7–9 we study the variables H_n under conditions, which in particular are satisfied when the variables X_i are independent and have the same continuous distribution function $F(x)$. We find the distribution of the variables H_n. It turns out that the distribution is independent of $F(x)$.

2. As sample space E we take the product set $(R, R, \ldots)$, where R is the set of real numbers. By $[\ldots]$, where $\ldots$ indicates a number of relations involving $X_1, X_2, \ldots$, we denote the event in E at which the relations are satisfied.

If the joint distribution function $\Pr\{\bigcap_{i=1}^{n}[X_i \leq x_i]\}$ of $X_1, \ldots, X_n$ is a symmetric function of $x_1, \ldots, x_n$, we shall say that the random variables $X_1, \ldots, X_n$ are symmetrically dependent. If the sequence $X_1, X_2, \ldots$ is infinite and $X_1, \ldots, X_n$ are symmetrically dependent for $n = 2, 3, \ldots$, we shall say that the random variables $X_1, X_2, \ldots$ are symmetrically dependent.

If an event C is invariant under permutations of the variables $x_1, \ldots, x_n$ or $x_1, x_2, \ldots$ we shall say that the event is symmetric with respect to $X_1, \ldots, X_n$ or $X_1, X_2, \ldots$, respectively.

By an s-permutation of n numbers, $a_1, \ldots, a_n$, we shall understand a permutation, of the n numbers, $\varepsilon_1 a_1, \ldots, \varepsilon_n a_n$, where $\varepsilon_i = \pm 1$, $i = 1, \ldots, n$, are arbitrarily chosen.

Let the random variables $X_1, \ldots, X_n$ satisfy the following conditions: a) the joint distribution of $X_1, \ldots, X_n$ is invariant under s-permutations of the variables, i. e. for any set $x_1, \ldots, x_n$ of real numbers, the probability

$$\Pr\left\{\bigcap_{i=1}^{n} [\min(0, x_i) \leq X_i \leq \max(0, x_i)]\right\}$$

is invariant under s-permutations of the numbers $x_1, \ldots, x_n$; and b)

$$\Pr\{S_i = 0\} = 0, \qquad i = 1, \ldots, n.$$

Then we shall say that the random variables $X_1, \ldots, X_n$ satisfy conditions (S). If the sequence $X_1, X_2, \ldots$ is infinite, and $X_1, \ldots, X_n$ satisfy conditions (S) for $n = 1, 2, \ldots$, we shall say that the random variables $X_1, X_2, \ldots$ satisfy conditions (S).

From the papers [3] and [4], we shall quote the following results which we shall use in the proofs below (cf. [3, p. 125], [3, p. 128], [3, p. 133] and [4, p. 269] for Theorems A, B, C and D, respectively).

THEOREM A. *Let $X_1, X_2, \ldots$ be symmetrically dependent, and let C_n be an event which is symmetric with respect to $X_1, \ldots, X_n$. Then*

$$\Pr\{[N_n = m]C_n\} = \Pr\{[L_n = m]C_n\} = \Pr\{[M_n = n-m]C_n\},$$
$$m = 0, 1, \ldots, n.$$

THEOREM B. *Let $X_1, \ldots, X_n$ be identically distributed and independent, and let K_n be one of the variables L_n, M_n or N_n. Then*

$$\Pr\{K_n = m\} = \Pr\{K_m = m\}\,\Pr\{K_{n-m} = 0\}, \qquad m = 0, 1, \ldots, n.$$

THEOREM C. *Let $X_1, X_2, \ldots$ satisfy conditions (S), and let K_n be one of the variables L_n, M_n or N_n. Then, if C_n is an event which is invariant under s-permutations of the variables $X_1, \ldots, X_n$, we have*

$$\Pr\{[K_n = m]C_n\} = (-1)^n \binom{-\frac{1}{2}}{m}\binom{-\frac{1}{2}}{n-m}\Pr\{C_n\}, \qquad m = 0, 1, \ldots, n,$$

and when $\Pr\{C_n\} > 0$,

$$\Pr\{K_n = m \mid C_n\} = (-1)^n \binom{-\frac{1}{2}}{m}\binom{-\frac{1}{2}}{n-m}, \qquad m = 0, 1, \ldots, n.$$

THEOREM D. *Let $X_1, \ldots, X_{n+1}$ be symmetrically dependent, and let C_{n+1} be an event which is symmetric with respect to $X_1, \ldots, X_{n+1}$. Let $N_n{}^*$ be the number of points (j, S_j), $j = 1, \ldots, n$, which lie above the straight line from $(0, 0)$ to $(n+1, S_{n+1})$. Then, for $\Pr\{C_{n+1}\} > 0$,*

$$\Pr\{N_n{}^* = m \mid C_{n+1}\} = (n+1)^{-1}, \qquad m = 0, 1, \ldots, n,$$

if and only if

$$\Pr\{i^{-1}S_i = (n+1)^{-1}S_{n+1} \mid C_{n+1}\} = 0, \qquad i = 1, 2, \ldots, n.$$

3. In this section, we shall prove a theorem which gives explicit formulae for $\Pr\{K_n = m\}$, $\Pr\{K_n = m \mid S_{n+1} > 0\}$ and $\Pr\{K_n = m \mid S_{n+1} \leqq 0\}$ in terms of $\Pr\{S_n > 0\}$, $n = 1, 2, \ldots$. In order to simplify the formulae, we use the symbol Σ^*. This symbol shall indicate that the summation is restricted to those values of the summation variables $\alpha_1, \ldots, \alpha_n$ which are non-negative and satisfy the relation $\alpha_1 + 2\alpha_2 + \ldots + n\alpha_n = n$.

THEOREM 1. *Let $X_1, X_2, \ldots$ be independent and identically distributed, and let $a_n = \Pr\{S_n > 0\}$, $n = 1, 2, \ldots$. Let K_n be N_n, L_n, or $n - M_n$, and let $p_n = \Pr\{K_n = n\}$ and $q_n = \Pr\{K_n = 0\}$, for $n = 1, 2, \ldots$. Put $p_0 = q_0 = 1$ and $q_{-1} = 0$. Then, for $n = 1, 2, \ldots$ and $m = 0, \ldots, n$,*

$$(3.1) \qquad \Pr\{K_n = n\} = p_n = \sum_{\alpha_1,\ldots,\alpha_n}^* \prod_{i=1}^n (\alpha_i!)^{-1} (a_i i^{-1})^{\alpha_i},$$

$$(3.2) \qquad \Pr\{K_n = 0\} = q_n = \sum_{\alpha_1,\ldots,\alpha_n}^* \prod_{i=1}^n (\alpha_i!)^{-1} ((1-a_i)i^{-1})^{\alpha_i},$$

$$(3.3) \qquad \Pr\{K_n = m\} = p_m q_{n-m},$$

$$(3.4) \qquad \Pr\{K_n = m \mid S_{n+1} > 0\} = a_{n+1}^{-1} \sum_{k=m}^n p_{k+1}(q_{n-k} - q_{n-k-1}),$$

$$(3.5) \qquad \Pr\{K_n = m \mid S_{n+1} \leq 0\} = (1-a_{n+1})^{-1} \sum_{k=m}^n q_{n-k}(p_k - p_{k+1}).$$

Furthermore the generating functions

$$A(s) = \sum_{n=1}^\infty a_n s^n, \quad P(s) = \sum_{n=0}^\infty p_n s^n \quad and \quad Q(s) = \sum_{n=0}^\infty q_n s^n$$

satisfy, for $|s| < 1$, *the relations*

$$(3.6) \qquad P(s) = \exp\left(\int_0^s \sigma^{-1} A(\sigma) d\sigma\right),$$

$$(3.7) \qquad Q(s) = (1-s)^{-1} \exp\left(-\int_0^s \sigma^{-1} A(\sigma) d\sigma\right) = (1-s)^{-1}(P(s))^{-1}.$$

PROOF: It follows from Theorem A that it is sufficient to prove the theorem in the case in which K_n is N_n. Since we have

$$[N_{n+1} \geq m+1] = [N_n = m][S_{n+1} > 0] \cup [N_n \geq m+1],$$

where the two components are non-overlapping, we obtain the relation

$$\Pr\{[N_n = m][S_{n+1} > 0]\} = \Pr\{N_{n+1} \geq m+1\} - \Pr\{N_n \geq m+1\}$$

$$(3.8) \qquad = \sum_{k=m}^n \Pr\{N_{n+1} = k+1\} - \sum_{k=m}^n \Pr\{N_n = k+1\},$$

and, using Theorem B, the formula,

$$(3.9) \qquad \Pr\{[N_n = m][S_{n+1} > 0]\} = \sum_{k=m}^n p_{k+1}(q_{n-k} - q_{n-k-1}).$$

We now sum equation (3.9) for $m = 0, 1, \ldots, n$ and obtain

$$(3.10) \quad \sum_{m=0}^n \Pr\{[N_n = m][S_{n+1} > 0]\} = \sum_{k=0}^n (k+1) p_{k+1}(q_{n-k} - q_{n-k-1}).$$

The left hand side is $\Pr\{S_{n+1} > 0\} = a_{n+1}$. If we introduce this in (3.10)

and multiply by s^n and thereafter sum for $n = 0, 1, 2, \ldots$, we obtain for $|s| < 1$ the relation

$$\sum_{n=0}^{\infty}{}' s^n a_{n+1} = \sum_{n=0}^{\infty}{}' s^n \sum_{k=0}^{n}{}' (k+1) p_{k+1}(q_{n-k}-q_{n-k-1})$$

or

$$(3.11) \quad s^{-1}A(s) = \left(\sum_{n=1}^{\infty} n\,p_n s^{n-1}\right)\left(\sum_{n=0}^{\infty}(q_n-q_{n-1})s^n)\right) = P'(s)\,Q(s)\,(1-s)\,.$$

We now consider the product $P(s)Q(s)$ and obtain for $|s| < 1$ the relation

$$P(s)\,Q(s) = \sum_{n=0}^{\infty}{}'\left(\sum_{m=0}^{n}{}' p_m q_{n-m}\right)s^n\,.$$

Since $p_m q_{n-m} = \Pr\{N_n=m\}$, we obtain

$$\sum_{m=0}^{n}{}' p_m q_{n-m} = \sum_{m=0}^{n}{}' \Pr\{N_n=m\} = \Pr\{E\} = 1\,,$$

so that

$$(3.12) \qquad\qquad P(s)\,Q(s) = (1-s)^{-1}\,.$$

From (3.11) and (3.12), it follows that $s^{-1}A(s) = P'(s)\big(P(s)\big)^{-1}$ and

$$(3.13) \qquad \int_{0}^{s} \sigma^{-1}A(\sigma)d\sigma = \log P(s), \qquad |s| < 1\,.$$

Formula (3.6) follows from (3.13), and, using (3.12), formula (3.7) also follows. The formulae (3.1) and (3.2) are derived from (3.6) and (3.7), respectively, by evaluation of the power series. The formula (3.3) is the conclusion of Theorem B, while (3.4) follows from (3.9), if we change to conditional probabilities. From (3.3) and (3.9),

$$\Pr\{[N_n=m][S_{n+1}\leqq 0]\} = p_m q_{n-m} - \sum_{k=m}^{n} p_{k+1}(q_{n-k}-q_{n-k-1})$$
$$= \sum_{k=m}^{n} q_{n-k}(p_k-p_{k+1})$$

follows. When we change to conditional probabilities, we obtain the relation (3.5).

4. We shall now consider the special case where we have $a_1=a_2=\ldots=a$. We shall assume that $0<a<1$, since we have in the cases $a=0$ and $a=1$ that $\Pr\{K_n=0\} = 1$ and $\Pr\{K_n=n\} = 1$, respectively, for all n. We prove the following:

THEOREM 2. *Let* $X_1, \ldots, X_{n+1}$ *be independent and identically distributed, and suppose that* $\Pr\{S_k > 0\} = a$ *for* $k = 1, \ldots, n+1$, *where* $0 < a < 1$. *Let* K_n *be* N_n, L_n, *or* $n - M_n$. *Then, for* $m = 0, 1, \ldots, n$,

$$(4.1) \qquad \Pr\{K_n = m\} = (-1)^n \binom{-a}{m} \binom{a-1}{n-m},$$

$$(4.2) \qquad \Pr\{K_n = m \mid S_{n+1} > 0\}$$
$$= a^{-1}(-1)^{n+1}(n+1)^{-1}(m+1) \binom{-a}{m+1} \binom{a-1}{n-m},$$

$$(4.3) \qquad \Pr\{K_n = m \mid S_{n+1} \leqq 0\}$$
$$= (1-a)^{-1}(-1)^{n+1}(n+1)^{-1}(n+1-m) \binom{-a}{m} \binom{a-1}{n+1-m},$$

$$(4.4) \qquad \left| \Pr\{K_n \leqq m\} - \pi^{-1}\sin(\pi a) \int_0^\mu x^{a-1}(1-x)^{-a}\, dx \right|$$
$$< \big(\Gamma(a)\big)^{-1}(n+1)^{a-1} + \big(\Gamma(1-a)\big)^{-1}(n+1)^{-a},$$

$$(4.5) \qquad \left| \Pr\{K_n \leqq m \mid S_{n+1} > 0\} - (\pi a)^{-1}\sin(\pi a) \int_0^\mu x^a(1-x)^{-a}\, dx \right|$$
$$< \big(\Gamma(1+a)\big)^{-1}(n+1)^{a-1},$$

$$(4.6) \qquad \left| \Pr\{K_n \leqq m \mid S_{n+1} \leqq 0\} - (\pi - \pi a)^{-1}\sin(\pi a) \int_0^\mu x^{a-1}(1-x)^{1-a}\, dx \right|$$
$$< \big(\Gamma(2-a)\big)^{-1}(n+1)^{-a},$$

where μ *stands for* $(m+1)/(n+1)$.

REMARK. It is evident that we have $a_n = \frac{1}{2}$ for all n if the common distribution function of $X_1, X_2, \ldots$ is symmetric and continuous. (When $a = \frac{1}{2}$ the integrals in formulae (4.4), (4.5) and (4.6) can be evaluated, and we obtain the corollary below.) Examples where $a_1 \doteq a_2 = \ldots = a$ and $a \neq \frac{1}{2}, 0, 1$ are rarer. There does, however, exist at least one simple example. Let $X_1, X_2, \ldots$ be independent, and let them all have the same Cauchy distribution with distribution function,

$$F(x) = \pi^{-1}\arctan\big(x + \cot(\pi a)\big) + \tfrac{1}{2};$$

then $a_n = \Pr\{S_n > 0\} = a$ for all n.

COROLLARY. *Let* $X_1, \ldots, X_{n+1}$ *be independent and identically distributed, and suppose that* $\Pr\{S_k > 0\} = \frac{1}{2}$ *for* $k = 1, \ldots, n+1$. *Let* K_n *be* N_n, L_n, *or* $n - M_n$. *Then, for* $m = 0, 1, \ldots, n$,

(4.7) $\left| \Pr\{K_n \leqq m\} - 2\pi^{-1}\arcsin\mu^{\frac{1}{2}} \right| < 2\pi^{-\frac{1}{2}}(n+1)^{-\frac{1}{2}}$,

(4.8) $\left| \Pr\{K_n \leqq m \mid S_{n+1} > 0\} - 2\pi^{-1}(\arcsin\mu^{\frac{1}{2}} - (\mu - \mu^2)^{\frac{1}{2}}) \right|$
$$< 2\pi^{-\frac{1}{2}}(n+1)^{-\frac{1}{2}} ,$$

(4.9) $\left| \Pr\{K_n \leqq m \mid S_{n+1} \leqq 0\} - 2\pi^{-1}(\arcsin\mu^{\frac{1}{2}} + (\mu - \mu^2)^{\frac{1}{2}}) \right|$
$$< 2\pi^{-\frac{1}{2}}(n+1)^{-\frac{1}{2}} ,$$

where μ again stands for $(m+1)/(n+1)$.

PROOF OF THEOREM 2: From Theorem 1, it follows that $\Pr\{K_n = m\}$, $\Pr\{K_n = m \mid S_{n+1} > 0\}$, and $\Pr\{K_n = m \mid S_{n+1} \leqq 0\}$ are functions of $a_1, \ldots, a_n$ and do not depend on $a_{n+1}, a_{n+2}, \ldots$. We can therefore in this proof replace a_k by a for $k = n+1, n+2, \ldots$. We then obtain, from Theorem 1, since now

$$A(s) = \sum_{n=1}^{\infty} a_n s^n = a s(1-s)^{-1} ,$$

the relations,

(4.10) $P(s) = \exp\left(\int_0^s \sigma^{-1} A(\sigma)d\sigma \right) = \exp\left(a \int_0^s (1-\sigma)^{-1}d\sigma \right)$
$$= \exp(-a\log(1-s)) = (1-s)^{-a} ,$$

(4.11) $Q(s) = (1-s)^{-1}(P(s))^{-1} = (1-s)^{a-1}$.

When we introduce the binomial series for $(1-s)^{-a}$ and $(1-s)^{a-1}$, we obtain, from (4.10) and (4.11), the formulae

$$p_n = (-1)^n \binom{-a}{n}, \quad q_n = (-1)^n \binom{a-1}{n} \quad \text{and} \quad q_n - q_{n-1} = (-1)^n \binom{a}{n}.$$

Formula (4.1) follows from (3.3), and the formula,

$$\Pr\{K_n = m \mid S_{n+1} > 0\} = a^{-1}(-1)^{n+1}\sum_{k=m}^{n} \binom{-a}{k+1}\binom{a}{n-k}$$
$$= a^{-1}(-1)^{n+1}\sum_{i=0}^{n-m} \binom{a}{i}\binom{-a}{n+1-i} ,$$

follows from (3.4). This formula can be simplified by using the formula

(4.12) $\displaystyle\sum_{i=0}^{k} \binom{a}{i}\binom{-a}{n-i} = n^{-1}(n-k)\binom{a-1}{k}\binom{-a}{n-k}, \qquad \begin{matrix} n = 1, 2, \ldots, \\ k = 0, 1, \ldots, n \end{matrix}$,

which can be proved by induction, see [5]. We then obtain (4.2). The formula (4.3) follows from (4.1), (4.2), and

$$\Pr\{K_n = m \mid S_{n+1} \leqq 0\} = (1-a)^{-1}\big(\Pr\{K_n = m\} - a\Pr\{K_n = m \mid S_{n+1} > 0\}\big).$$

Since there does not seem to exist summation formulae for

$$\sum_{i=0}^{k}\binom{-a}{i}\binom{a-1}{n-i}, \quad \sum_{i=0}^{k}(i+1)\binom{-a}{i+1}\binom{a-1}{n-i},$$

and

$$\sum_{i=0}^{k}(n+1-i)\binom{-a}{i}\binom{a-1}{n+1-i}$$

corresponding to the formula (4.12), we are not able to obtain equally simple expressions for

$$\Pr\{K_n \leqq m\}, \quad \Pr\{K_n \leqq m \mid S_{n+1} > 0\}, \quad \text{and} \quad \Pr\{K_n \leqq m \mid S_{n+1} \leqq 0\}\,.$$

We shall, however, prove the inequalities (4.4), (4.5) and (4.6).

For $0 < a < 1$ we have the formula of Gauss,

$$(4.13) \qquad (-1)^n\binom{-a}{n}n^{1-a} \to \big(\Gamma(a)\big)^{-1} \quad \text{for} \quad n \to \infty\,.$$

As is easily seen, the converging sequence is monotonically increasing. Thus we obtain from (4.13) the inequality

$$(4.14) \quad (-1)^n\binom{-a}{n} < \big(\Gamma(a)\big)^{-1}n^{a-1}, \quad n = 1, 2, \ldots \quad (0 < a < 1)\,.$$

For $m = 0, 1, \ldots, n-1$, the inequalities,

$$(4.15) \quad \Pr\{K_n = m \mid S_{n+1} > 0\}$$
$$< a^{-1}(n+1)^{-1}(m+1)\big(\Gamma(a)\big)^{-1}(m+1)^{a-1}\big(\Gamma(1-a)\big)^{-1}(n-m)^{-a}$$
$$= (\pi a)^{-1}\sin(\pi a)\,(n+1)^{-1}(m+1)^a(n-m)^{-a}$$

follow from (4.2) and (4.14); and, for $m = n$, the inequality

$$(4.16) \ \Pr\{K_n = n \mid S_{n+1} > 0\} = a^{-1}(-1)^{n+1}\binom{-a}{n+1} < \big(\Gamma(1+a)\big)^{-1}(n+1)^{a-1}$$

follows from (4.14). As is easily seen from (4.2), we have

$$\Pr\{K_n = m \mid S_{n+1} > 0\} < \Pr\{K_n = m+1 \mid S_{n+1} > 0\}$$

for $m = 0, 1, \ldots, n-1$. We therefore obtain, from (4.16), the inequalities

$$(4.17) \ \Pr\{K_n = m \mid S_{n+1} > 0\} < \big(\Gamma(1+a)\big)^{-1}(n+1)^{a-1}, \quad m = 0, 1, \ldots, n\,.$$

By summation of (4.15), we obtain, for $m = 0, 1, \ldots, n$,

$$(4.18) \ \Pr\{K_n < m \mid S_{n+1} > 0\} < (\pi a)^{-1}\sin(\pi a)\,(n+1)^{-1}\sum_{k=0}^{m-1}(k+1)^a(n-k)^{-a}\,.$$

For $m = 0, 1, \ldots, n$,

$$(4.19) \quad \Pr\{K_n \leqq m \mid S_{n+1} > 0\} < (\pi a)^{-1} \sin(\pi a) (n+1)^{-1} \sum_{k=0}^{m-1} (k+1)^a (n-k)^{-a}$$
$$+ \left(\Gamma(1+a)\right)^{-1} (n+1)^{a-1}$$

follows from (4.17) and (4.18).

By summation of (4.15), we obtain, using (4.16), for $m = 0, 1, \ldots, n$,

$$(4.20) \quad \Pr\{K_n > m \mid S_{n+1} > 0\} < (\pi a)^{-1} \sin(\pi a) (n+1)^{-1} \sum_{k=m+1}^{n-1} (k+1)^a (n-k)^{-a}$$
$$+ \left(\Gamma(1+a)\right)^{-1} (n+1)^{a-1} .$$

We shall now compare $\sum (k+1)^a (n-k)^{-a}$ with $\int x^a (1-x)^{-a} dx$. Since the function $x^a(1-x)^{-a}$, for $0 < a < 1$, is strictly increasing in the interval $0 < x < 1$, we obtain, for $k = 0, 1, \ldots, n-1$, the relation,

$$(n+1)^{-1} (k+1)^a (n-k)^{-a} < \int_{(k+1)/(n+1)}^{(k+2)/(n+1)} x^a (1-x)^{-a} dx ,$$

and, therefore, for $m = 0, 1, \ldots, n-1$,

$$(4.21) \quad \Pr\{K_n \leqq m \mid S_{n+1} > 0\}$$
$$< (\pi a)^{-1} \sin(\pi a) \int_{1/(n+1)}^{(m+1)/(n+1)} x^a (1-x)^{-a} dx + \left(\Gamma(1+a)\right)^{-1} (n+1)^{a-1}$$

and

$$(4.22) \quad \Pr\{K_n > m \mid S_{n+1} > 0\}$$
$$< (\pi a)^{-1} \sin(\pi a) \int_{(m+2)/(n+1)}^{1} x^a (1-x)^{-a} dx + \left(\Gamma(1+a)\right)^{-1} (n+1)^{a-1} .$$

Since $\int_0^1 x^a (1-x)^{-a} dx = (\pi a)(\sin(\pi a))^{-1}$, we obtain from (4.22) the relation

$$(4.23) \quad \Pr\{K_n \leqq m \mid S_{n+1} > 0\} = 1 - \Pr\{K_n > m \mid S_{n+1} > 0\}$$
$$> (\pi a)^{-1} \sin(\pi a) \int_{0}^{(m+2)/(n+1)} x^a (1-x)^{-a} dx - \left(\Gamma(1+a)\right)^{-1} (n+1)^{a-1} .$$

The inequalities (4.5) follow from (4.21) and (4.23). The inequalities (4.6) are proved in a similar way and, thereafter, (4.4) is obtained from (4.5) and (4.6) by using the formula

$$\Pr\{K_n \leqq m\} = a \Pr\{K_n \leqq m \mid S_{n+1} > 0\} + (1-a) \Pr\{K_n \leqq m \mid S_{n+1} \leqq 0\} .$$

5. We shall now consider the more general case where we do not have $\Pr\{S_n>0\} = a$ for all n, but only $\Pr\{S_n>0\} = a_n \to a$ for $n \to \infty$. We shall first prove the following:

LEMMA. *Let the value of* $\Pr\{K_n \leqq m\}$ *corresponding to the sequence* $a_1, \ldots, a_n$ *be denoted by* $b_{n,m}(a_1, \ldots, a_n)$. *Then*

$$(5.1) \qquad\qquad b_{n,m}(a_1, \ldots, a_n) \geqq b_{n,m}(b_1, \ldots, b_n),$$

if $0 \leqq a_i \leqq b_i \leqq 1$ *for* $i = 1, \ldots, n$.

REMARK. Since $b_{n,m}(a_1, \ldots, a_n)$ is a polynomial in $a_1, \ldots, a_n$, we may extend the domain to the whole n-dimensional cube $0 \leqq a_i \leqq 1$, $i = 1, \ldots, n$, although $\Pr\{K_n \leqq m\}$ is only defined if there exist random variables X_k such that $a_i = \Pr\{S_i>0\}$, $i = 1, \ldots, n$.

PROOF: We introduce, for $n = 1, 2, \ldots$, the notations

$$p_n(a_1, \ldots, a_n) = \sum_{\alpha_1, \ldots, \alpha_n}^{*} \prod_{i=1}^{n} (\alpha_i!)^{-1}(a_i i^{-1})^{\alpha_i},$$

$$q_n(a_1, \ldots, a_n) = \sum_{\alpha_1, \ldots, \alpha_n}^{*} \prod_{i=1}^{n} (\alpha_i!)^{-1}((1-a_i)i^{-1})^{\alpha_i},$$

and put $p_0 = q_0 = 1$. For $k = 1, \ldots, n$ we have

$$\frac{\partial}{\partial a_k} p_n(a_1, \ldots, a_n) = \sum_{\alpha_1, \ldots, \alpha_n}^{*} \frac{\partial}{\partial a_k} \prod_{i=1}^{n} (\alpha_i!)^{-1}(a_i i^{-1})^{\alpha_i}.$$

When we differentiate, the terms with $\alpha_k = 0$ disappear. The terms with $\alpha_k > 0$ give new terms of the form

$$k^{-1} \prod_{i=1}^{n} (\beta_i!)^{-1}(a_i i^{-1})^{\beta_i},$$

where $\beta_i = \alpha_i$ for $i = 1, \ldots, k-1, k+1, \ldots, n$, while $\beta_k = \alpha_k-1$. The numbers β_i are non-negative and satisfy the relation

$$\beta_1 + 2\beta_2 + \ldots + n\beta_n = n-k,$$

so that $\beta_i = 0$ for $i = n-k+1, \ldots, n$. We evidently obtain all sets $\beta_1, \ldots, \beta_{n-k}$ for which $\beta_i \geqq 0$ and which satisfy

$$\beta_1 + 2\beta_2 + \ldots + (n-k)\beta_{n-k} = n-k.$$

Each set is obtained once and only once. Therefore the terms in

$$\partial p_n(a_1, \ldots, a_n)/\partial a_k$$

are exactly the terms in the sum

$$\sum_{\beta_1,\ldots,\beta_{n-k}}^* \quad k^{-1} \prod_{i=1}^{n-k} (\beta_i!)^{-1} (a_i i^{-1})^{\beta_i},$$

and we have

(5.2) $\dfrac{\partial}{\partial a_k} p_n(a_1,\ldots,a_n) = k^{-1} p_{n-k}(a_1,\ldots,a_{n-k}),\qquad k = 1,\ldots,n.$

Since $q_n(a_1,\ldots,a_n) = p_n(1-a_1,\ldots,1-a_n)$, we obtain

(5.3) $\dfrac{\partial}{\partial a_k} q_n(a_1,\ldots,a_n) = -\,k^{-1}\, p_{n-k}(1-a_1,\ldots,1-a_{n-k})$

$$= -\,k^{-1}\, q_{n-k}(a_1,\ldots,a_{n-k}),\qquad k = 1,\ldots,n.$$

The formula $b_{n,m}(a_1,\ldots,a_n) = \sum_{i=0}^m p_i(a_1,\ldots,a_i)\, q_{n-i}(a_1,\ldots,a_{n-i})$ follows from (3.3) by summation, and is valid, by extension, in the whole n-dimensional cube $0 \leq a_i \leq 1$, $i = 1,\ldots,n$. Using (5.2) and (5.3), we obtain

(5.4) $\dfrac{\partial}{\partial a_k} b_{n,m}(a_1,\ldots,a_n)$

$$= \sum_{i=0}^m \left(p_i(a_1,\ldots,a_i)\frac{\partial}{\partial a_k} q_{n-i}(a_1,\ldots,a_{n-i}) + \right.$$

$$\left. + q_{n-i}(a_1,\ldots,a_{n-i})\frac{\partial}{\partial a_k} p_i(a_1,\ldots,a_i) \right)$$

$$= -k^{-1} \sum_{i=0}^{\min(m,n-k)} p_i(a_1,\ldots,a_i)\, q_{n-i-k}(a_1,\ldots,a_{n-i-k}) +$$

$$+ k^{-1} \sum_{i=k}^{m} q_{n-i}(a_1,\ldots,a_{n-i})\, p_{i-k}(a_1,\ldots,a_{i-k})$$

$$= -k^{-1}\left(\sum_{i=0}^{\min(m,n-k)} p_i(a_1,\ldots,a_i)\, q_{n-i-k}(a_1,\ldots,a_{n-i-k}) - \right.$$

$$\left. - \sum_{i=0}^{m-k} q_{n-i-k}(a_1,\ldots,a_{n-i-k})\, p_i(a_1,\ldots,a_i) \right)$$

$$= -k^{-1} \sum_{i=\max(0,\,m-k+1)}^{\min(m,n-k)} p_i(a_1,\ldots,a_i)\, q_{n-i-k}(a_1,\ldots,a_{n-i-k})$$

$$\leq 0,$$

if $0 \leq a_i \leq 1$ for $i = 1,\ldots,n$. Formula (5.1) follows from (5.4), for $k = 1,\ldots,n$, since we obtain

$$b_{n,m}(a_1,\ldots,a_n) \geq b_{n,m}(b_1,a_2,\ldots,a_n) \geq b_{n,m}(b_1,b_2,a_3,\ldots,a_n) \geq \ldots$$

$$\ldots \geq b_{n,m}(b_1,\ldots,b_n).$$

Using the lemma and formula (4.4), we shall now obtain the limiting distribution of K_n/n for $n \to \infty$ in the case in which $a_n \to a$ for $n \to \infty$.

THEOREM 3. *Let $X_1, X_2, \ldots$ be independent and identically distributed, and suppose that $\Pr\{S_n > 0\} = a_n \to a$ for $n \to \infty$, where $0 < a < 1$. Let K_n be N_n, L_n, or $n - M_n$. Then, for $0 < \alpha < 1$,*

$$(5.5) \qquad \lim_{n \to \infty} \Pr\{K_n \leq \alpha n\} = \pi^{-1} \sin(\pi a) \int_0^\alpha x^{a-1}(1-x)^{-a} dx .$$

REMARK. For $a = \frac{1}{2}$, the integral in (5.5) can be evaluated, and we obtain

$$(5.6) \qquad \lim_{n \to \infty} \Pr\{K_n \leq \alpha n\} = 2\pi^{-1} \arcsin \alpha^{\frac{1}{2}} .$$

PROOF: Let $m = [\alpha n]$. We then have $\Pr\{K_n \leq \alpha n\} = b_{n,m}(a_1, \ldots, a_n)$. From the lemma, it follows that

$$(5.7) \qquad b_{n,m}(b_1, \ldots, b_n) \geq \Pr\{K_n \leq m\} \geq b_{n,m}(c_1, \ldots, c_n)$$

if $0 \leq b_k \leq a_k \leq c_k \leq 1$ for $k = 1, \ldots, n$. Let now $b_k = \min(a_k, a - \delta)$ and $c_k = \max(a_k, a + \delta)$ for $k = 1, 2, \ldots$. It then follows that, for $\delta < \min(a, 1-a)$, we have

$$(5.8) \qquad \begin{aligned} &\limsup_{n \to \infty} b_{n,m}(b_1, \ldots, b_n) \geq \limsup_{n \to \infty} \Pr\{K_n \leq m\} \\ &\geq \liminf_{n \to \infty} \Pr\{K_n \leq m\} \geq \liminf_{n \to \infty} b_{n,m}(c_1, \ldots, c_n) . \end{aligned}$$

Since $\pi^{-1} \sin(\pi a) \int_0^\alpha x^{a-1}(1-x)^{-a} dx$ is a continuous function of a in the interval $0 < a < 1$ for any α in the interval $0 < \alpha < 1$, it follows from (5.8), which holds for arbitrarily small δ, that it is sufficient, in order to prove Theorem 3, to show that,

$$(5.9) \qquad \lim_{n \to \infty} b_{n,m}(a_1, \ldots, a_n) = \pi^{-1} \sin(\pi a) \int_0^\alpha x^{a-1}(1-x)^{-a} dx ,$$

if $0 < a_n < 1$ for all n and $a_n = a$ for $n > N$, where N is an arbitrary but fixed number. Since we know already from (4.4) that (5.9) holds if $N = 0$, it is sufficient to prove that

$$(5.10) \qquad b_{n,m}(a_1, \ldots, a_N, a, \ldots, a) - b_{n,m}(a, \ldots, a) \to 0 \quad \text{for} \quad n \to \infty .$$

Let now

$$A(s) = \sum_{n=1}^N a_n s^n + \sum_{n=N+1}^\infty a s^n .$$

Then $b_{n,m}(a_1, \ldots, a_N, a, \ldots, a) = \sum_{k=0}^m p_k q_{n-k}$, where p_ν and q_ν are determined by

$$P(s) = \sum_{n=0}^\infty p_n s^n = \exp\left(\int_0^s \sigma^{-1} A(\sigma)\,d\sigma\right)$$

and

$$Q(s) = \sum_{n=0}^\infty q_n s^n = (1-s)^{-1}\left(P(s)\right)^{-1}.$$

We evidently have

$$\int_0^s \sigma^{-1} A(\sigma)\,d\sigma = \int_0^s \left(\sum_{n=1}^\infty a\,\sigma^{n-1} + \sum_{n=1}^N (a_n-a)\sigma^{n-1}\right) d\sigma$$

$$= -a\log(1-s) + \sum_{n=1}^N n^{-1}(a_n-a)s^n.$$

We therefore obtain

(5.11)
$$P(s) = (1-s)^{-a} \exp\left(\sum_{n=1}^N n^{-1}(a_n-a)s^n\right)$$

and

(5.12)
$$Q(s) = (1-s)^{a-1} \exp\left(-\sum_{n=1}^N n^{-1}(a_n-a)s^n\right).$$

Let now

$$\exp\left(\sum_{n=1}^N n^{-1}(a_n-a)s^n\right) = \sum_{n=0}^\infty b_n s^n$$

and

$$\exp\left(-\sum_{n=1}^N n^{-1}(a_n-a)s^n\right) = \sum_{n=0}^\infty c_n s^n.$$

If we introduce these series in (5.11) and (5.12), we obtain

(5.13)
$$\sum_{n=0}^\infty p_n s^n = \left(\sum_{n=0}^\infty (-1)^n \binom{-a}{n} s^n\right)\left(\sum_{n=0}^\infty b_n s^n\right)$$

and

(5.14)
$$\sum_{n=0}^\infty q_n s^n = \left(\sum_{n=0}^\infty (-1)^n \binom{a-1}{n} s^n\right)\left(\sum_{n=0}^\infty c_n s^n\right).$$

From these formulae,

$$p_n = \sum_{i=0}^n (-1)^i \binom{-a}{i} b_{n-i}, \qquad q_n = \sum_{j=0}^n (-1)^j \binom{a-1}{j} c_{n-j}$$

and

$$(5.15) \quad b_{n,m}(a_1, \ldots, a_N, a, \ldots, a) = \sum_{k=0}^{m} p_k q_{n-k}$$

$$= \sum_{k=0}^{m} \sum_{i=0}^{k} (-1)^i \binom{-a}{i} b_{k-i} \sum_{j=0}^{n-k} (-1)^j \binom{a-1}{j} c_{n-k-j}$$

$$= \sum_{i=0}^{m} \sum_{j=0}^{n-i} (-1)^i \binom{-a}{i} (-1)^j \binom{a-1}{j} \sum_{k=i}^{\min(n-j,m)} b_{k-i} c_{n-k-j}$$

$$= \sum_{i=0}^{m} \sum_{j=0}^{n-i} (-1)^i \binom{-a}{i} (-1)^j \binom{a-1}{j} \sum_{k=0}^{\min(n-j-i,m-i)} b_k c_{n-j-i-k},$$

follow. From the definitions of b_n and c_n, it follows that

$$\left(\sum_{n=0}^{\infty} b_n s^n \right) \left(\sum_{n=0}^{\infty} c_n s^n \right) = 1 ;$$

therefore $\sum_{k=0}^{n} b_k c_{n-k}$ equals 1 if $n = 0$ and 0 if $n = 1, 2, \ldots$. We thus obtain

$$\sum_{k=0}^{\min(n-j-i,m-i)} b_k c_{n-j-i-k} = \begin{cases} 1 & if \quad i+j = n \\ 0 & if \quad j \geqq n-m, \ i < m \end{cases}$$

(for other values of i and j no reduction seems possible). If we use this, we obtain from (5.15) the relation

$$(5.16) \quad b_{n,m}(a_1, \ldots, a_N, a, \ldots, a)$$

$$= \sum_{i=0}^{m} (-1)^n \binom{-a}{i} \binom{a-1}{n-i} + \sum_{i=0}^{m} \sum_{j=0}^{n-m-1} (-1)^{i+j} \binom{-a}{i} \binom{a-1}{j} \sum_{k=0}^{m-i} b_k c_{n-j-i-k}$$

$$= b_{n,m}(a, \ldots, a) + \sum_{i=0}^{m} \sum_{j=0}^{n-m-1} (-1)^{i+j} \binom{-a}{i} \binom{a-1}{j} \sum_{k=0}^{m-i} b_k c_{n-j-i-k} .$$

We now use the fact that $\sum_{n=0}^{\infty} b_n s^n$ and $\sum_{n=0}^{\infty} c_n s^n$ are regular functions in the s-plane, so that we have $|b_n| < C 8^{-n}$ and $|c_n| < C 8^{-n}$, where the constant C depends on $a_1, \ldots, a_N, a$ but not on n. We then obtain from (5.16) the inequality

$$(5.17) \quad \left| b_{n,m}(a_1, \ldots, a_N, a, \ldots, a) - b_{n,m}(a, \ldots, a) \right|$$

$$< \sum_{i=0}^{m} \sum_{j=0}^{n-m-1} (-1)^{i+j} \binom{-a}{i} \binom{a-1}{j} \sum_{k=0}^{m-i} C 8^{-k} C 8^{-n+j+i+k}$$

$$= 8^{-n} C^2 \left(\sum_{i=0}^{m} (-1)^i \binom{-a}{i} (m-i+1) 8^i \right) \left(\sum_{j=0}^{n-m-1} (-1)^j \binom{a-1}{j} 8^j \right).$$

In the two sums in the last line of (5.17), each term is less than half the following term. Each sum is therefore smaller than twice its last term.

When we use this and the inequality (4.14) together with the inequality derived from (4.14) by replacing a by $1-a$, we obtain

$$(5.18) \quad |b_{n,m}(a_1, \ldots, a_N, a, \ldots, a) - b_{n,m}(a, \ldots, a)|$$
$$< 8^{-n} C^2 2 \big(\Gamma(a)\big)^{-1} m^{a-1} 8^m 2 \big(\Gamma(1-a)\big)^{-1} (n-m-1)^{-a} 8^{n-m-1}$$
$$= \tfrac{1}{2} C^2 \pi^{-1} \sin(\pi a) \, m^{a-1} (n-m-1)^{-a} .$$

This completes the proof of Theorem 3, since (5.10) follows from (5.18).

6. If in Theorem C we put $C_n = E$, we obtain the same value for $\Pr\{K_n = m\}$ as the value which we obtain from Theorem 2 for $a = \frac{1}{2}$. We shall now prove a theorem which shows that the formulae in Theorem 2 (for $a = \frac{1}{2}$) and in the corollary are also valid, if we replace the assumption "$X_1, \ldots, X_{n+1}$ are independent and identically distributed and satisfy $\Pr\{S_n > 0\} = \frac{1}{2}$" by the assumption "$X_1, \ldots, X_{n+1}$ satisfy conditions (S)".

THEOREM 4. *Let* $X_1, \ldots, X_{n+1}$ *satisfy conditions* (S), *and let* K_n *be* N_n, L_n, *or* $n - M_n$. *Then* (4.1), (4.2), (4.3) *of Theorem 2 hold with* $a = \frac{1}{2}$. *Furthermore* (4.7), (4.8), (4.9) *of the corollary hold.*

PROOF: The formula (4.1) follows from Theorem C and is true also if we replace n by $n+1$. From (3.8), which holds for any set $X_1, \ldots, X_{n+1}$ of random variables, it follows that when two different sets of assumptions on $X_1, \ldots, X_{n+1}$ lead to the same formulae for $\Pr\{K_n = m\}$ and $\Pr\{K_{n+1} = m\}$, then the two sets of assumptions lead also to the same formulae for $\Pr\{K_n = m \mid S_{n+1} > 0\}$ and $\Pr\{K_n = m \mid S_{n+1} \leq 0\}$. The inequalities (4.7), (4.8), (4.9) are consequences of the formulae (4.1), (4.2), (4.3). Therefore the inequalities hold when conditions (S) are satisfied, and we have finished the proof of Theorem 4.

7. Let $b_0, \ldots, b_n$ be a sequence of real numbers. We shall say that the sequence is convex if the sequence $b_1 - b_0, b_2 - b_1, \ldots, b_n - b_{n-1}$ of the differences is nondecreasing. If $a_0, \ldots, a_n$ is an arbitrary sequence of real numbers, then there exists a unique, largest, convex minorant sequence $b_0, \ldots, b_n$. We evidently have $b_0 = a_0$ and $b_n = a_n$. Furthermore, for $i = 1, \ldots, n-1$, we have either $b_i = a_i$ or

$$b_i = (k-j)^{-1} \big((k-i)a_j + (i-j)a_k\big) ,$$

where j is the largest subscript ν less than i for which $a_\nu = b_\nu$ and k is the smallest subscript ν larger than i for which $a_\nu = b_\nu$.

The number of equalities $b_i = a_i$, $i = 1, \ldots, n-1$, is, to some extent,

a measure of the fluctuations of the sequence $a_0, \ldots, a_n$. We shall therefore study the probability distribution of the number of these equalities when $a_0, \ldots, a_n$ is replaced by the sequence of sums $S_0, S_1, \ldots, S_n$ of the random variables $X_1, \ldots, X_n$. We first prove the following:

THEOREM 5. *Let the random variables* $X_1, \ldots, X_n$ *be symmetrically dependent, and let the joint distribution satisfy*

$$(7.1) \qquad \Pr\{i^{-1}S_i = j^{-1}S_j\} = 0, \qquad 1 \leq i < j \leq n .$$

Let $T_0 = 0, T_1, \ldots, T_{n-1}, T_n = S_n$ *be the largest convex minorant sequence to the sequence* $S_0 = 0, S_1, \ldots, S_{n-1}, S_n$, *and let* H_n *be the number of equalities* $S_i = T_i$, $i = 1, \ldots, n-1$. *Then for* $m = 0, \ldots, n-1$

$$(7.2) \qquad \Pr\{H_n = m\} = \big((m+1)!\big)^{-1} \sum_{j_0, \ldots, j_m}^{[n]} \prod_{v=0}^{m} j_v^{-1} ,$$

where $\sum^{[n]}$ *indicates that the summation is restricted to those values of the summation variables* $j_0, \ldots, j_m$ *which are positive and satisfy* $j_0 + \ldots + j_m = n$.

Furthermore we have, for $|s| < 1$ *and* $0 < |t| < 1$,

$$(7.3) \qquad H(s, t) = \sum_{n=1}^{\infty} \sum_{m=0}^{n-1} \Pr\{H_n = m\} s^n t^m = t^{-1}\big((1-s)^{-t} - 1\big) .$$

The condition (7.1) *evidently is satisfied if the random variables* $X_1, \ldots, X_n$ *are independent and have the same continuous distribution.*

PROOF OF (7.2): Let $A^{(n)}_{i_1, \ldots, i_m}$ for $0 < i_1 < \ldots < i_m < n$ denote the event in which the equalities $S_i = T_i$ are satisfied for $i = i_1, \ldots, i_m$ and for no other values of i in the interval $1 \leq i \leq n-1$. (If $m = 0$, then there are no lower indices.) Then $[H_n = m]$ is the union of all possible $A^{(n)}_{i_1, \ldots, i_m}$. Since the events $A^{(n)}_{i_1, \ldots, i_m}$ are non-overlapping, we obtain

$$(7.4) \qquad \Pr\{H_n = m\} = \sum{}' \Pr\{A^{(n)}_{i_1, \ldots, i_m}\} ,$$

where the summation is over all possible sets of m integers $i_1, \ldots, i_m$ for which $0 < i_1 < \ldots < i_m < n$. In the following, we shall, when convenient, write i_0 for 0 and i_{m+1} for n.

We now consider the event $A^{(n)}_{i_1, \ldots, i_m}$ for a fixed set of indices $i_1, \ldots, i_m$. If follows from the definition of $T_0, \ldots, T_n$ that $A^{(n)}_{i_1, \ldots, i_m}$ is the event at which the inequalities

$$(7.5) \qquad (i_1 - i_0)^{-1}(S_{i_1} - S_{i_0}) < (i_2 - i_1)^{-1}(S_{i_2} - S_{i_1}) < \ldots$$
$$\ldots < (i_{m+1} - i_m)^{-1}(S_{i_{m+1}} - S_{i_m})$$

and

(7.6.0) $\qquad S_i > T_i \quad$ for $\quad i = 1, \ldots, i_1-1$,

(7.6.1) $\qquad S_i > T_i \quad$ for $\quad i = i_1+1, \ldots, i_2-1$,

$\ldots \qquad\qquad\qquad \ldots$

(7.6.m) $\qquad S_i > T_i \quad$ for $\quad i = i_m+1, \ldots, n-1$,

are satisfied. If we denote the event at which (7.5) is satisfied by $B^{(n)}_{i_1,\ldots,i_m}$ (for $m = 0$ we have $B^{(n)} = E$), and for $\nu = 0, \ldots, m$ the event at which (7.6.ν) is satisfied by D_ν (for $i_{\nu+1} = i +1$ we have $D_\nu = E$), then we obtain

(7.7) $$A^{(n)}_{i_1,\ldots,i_m} = B^{(n)}_{i_1,\ldots,i_m} D_0 \ldots D_m .$$

We now use the fact that the event $B^{(n)}_{i_1,\ldots,i_m} D_k \ldots D_m$, for $k = 1, \ldots, m+1$, is symmetric with respect to $X_{i_{k-1}+1}, \ldots, X_{i_k}$, so that we may apply Theorem D, first to the events

$$D_0 = [N_{i_1-1}{}^* = i_1-1] \quad \text{and} \quad B^{(n)}_{i_1,\ldots,i_m} D_1 \ldots D_m ,$$

thereafter to the events D_k and $B^{(n)}_{i_1,\ldots,i_m} D_{k+1} \ldots D_m$ for $k = 1, \ldots, m$. We then obtain

(7.8) $\Pr\{A^{(n)}_{i_1,\ldots,i_m}\} = (i_1-i_0)^{-1} \Pr\{B^{(n)}_{i_1,\ldots,i_m} D_1 \ldots D_m\}$

$\qquad\qquad\qquad = (i_1-i_0)^{-1}(i_2-i_1)^{-1} \Pr\{B^{(n)}_{i_1,\ldots,i_m} D_2 \ldots D_m\}$

$\qquad\qquad\qquad \ldots$

$\qquad\qquad\qquad = \left(\prod_{\nu=0}^{m} (i_{\nu+1}-i_\nu)^{-1} \right) \Pr\{B^{(n)}_{i_1,\ldots,i_m}\} .$

Before we apply Theorem D to D_k and $B^{(n)}_{i_1,\ldots,i_m} D_{k+1} \ldots D_m$, we must of course perform the permutation

$$X_i \to X_{i+n-i_k} \quad \text{for} \quad i = 1, \ldots, i_k ,$$
$$X_i \to X_{i-i_k} \quad \text{for} \quad i = i_k+1, \ldots, n ,$$

on the variables $X_1, \ldots, X_n$. This permutation, which carries D_k into $[N_{i_{k+1}-i_k-1}{}^* = i_{k+1}-i_k-1]$, does not change the probabilities we are considering.

From (7.4) and (7.8) it follows that $\Pr\{H_n=m\}$ can be expressed as a linear combination of $\Pr\{B^{(n)}_{i_1,\ldots,i_m}\}$, where $i_1, \ldots, i_m$ run through all possible values for which $0 < i_1 < \ldots < i_m < n$.

The event $B^{(n)}_{i_1,\ldots,i_m}$ defined by (7.5) depends only on the random variables $(i_{k+1}-i_k)^{-1}(S_{i_{k+1}}-S_{i_k})$, $k = 0, \ldots, m$. We shall denote these variables by $W_k^{(i_{k+1}-i_k)}$, $k = 0, \ldots, m$. The upper index, which we shall

omit when no ambiguities can arise, indicates the number of random variables X_i used in the definition of W_k.

As an illustration of the idea in the following more complicated argument let us first consider the case in which $i_1-i_0 = i_2-i_1 = \ldots = i_{m+1}-i_m$. Then the random variables $W_0, W_1, \ldots, W_m$ are symmetrically dependent, and, from (7.1), it follows that $\Pr\{W_i = W_j\} = 0$ for $i \neq j$. It therefore follows that

$$(7.9) \qquad \Pr\{B^{(n)}{}_{i_1,\ldots,i_m}\} = ((m+1)!)^{-1},$$

since we have $B^{(n)}{}_{i_1,\ldots,i_m} = [W_0 < W_1 < \ldots < W_m]$, and evidently the probability that the $m+1$ symmetrically dependent random variables $W_0, \ldots, W_m$ assume values which form an increasing sequence is $((m+1)!)^{-1}$. The relation (7.9) is not used in the following.

If we do not assume that $i_1-i_0 = i_2-i_1 = \ldots = i_{m+1}-i_m$, then (7.9) does not hold in general. This is easily seen from the following example. Let $n = 3$, $m = 1$, $i_1 = 1$. Then $B^{(3)}{}_1 = [X_1 < \frac{1}{2}(X_2+X_3)]$, and if the random variables X_1, X_2, X_3 assume the values 0, 1, 3 in arbitrary order such that each of the 6 permutations have probability 1/6, then $\Pr\{B^{(3)}{}_1\} = \Pr\{X_1 \neq 3\} = 2/3$.

We may, however, show that if we consider together with $B^{(n)}{}_{i_1,\ldots,i_m}$ the events $B^{(n)}{}_{k_1,\ldots,k_m}$ for which $k_1, k_2-k_1, \ldots, n-k_m$ is a permutation of $i_1, i_2-i_1, \ldots, n-i_m$, then

$$(7.10) \qquad \sideset{}{'}\sum_{k_1,\ldots,k_m} \Pr\{B^{(n)}{}_{k_1,\ldots,k_m}\} = p(i_1,\ldots,i_m, n)\,((m+1)!)^{-1},$$

the summation being over all sets $k_1, \ldots, k_m$ for which $k_1, k_2-k_1, \ldots, n-k_m$ is a permutation of $i_1, i_2-i_1, \ldots, n-i_m$, and where $p(i_1,\ldots,i_m,n)$ is the number of such permutations. In the illustration given above, the equalities $i_1-i_0 = i_2-i_1 = \ldots = i_{m+1}-i_m$ imply $p(i_1,\ldots,i_m,n) = 1$ so that (7.9) is contained in (7.10).

Before we give the proof of (7.10), we shall use this relation to prove (7.2). We first sum (7.8) over all $k_1, \ldots, k_m$ for which $k_1, k_2-k_1, \ldots, n-k_m$ is a permutation of $i_1, i_2-i_1, \ldots, n-i_m$, and obtain, since $\prod_{\nu=0}^{m}(i_{\nu+1}-i_\nu)^{-1}$ is a common factor,

$$(7.11) \quad \sideset{}{'}\sum_{k_1,\ldots,k_m} \Pr\{A^{(n)}{}_{k_1,\ldots,k_m}\} = \left(\prod_{\nu=0}^{m}(i_{\nu+1}-i_\nu)^{-1}\right) \sideset{}{'}\sum_{k_1,\ldots,k_m} \Pr\{B^{(n)}{}_{k_1,\ldots,k_m}\}$$

$$= \left(\prod_{\nu=0}^{m}(i_{\nu+1}-i_\nu)^{-1}\right) p(i_1,\ldots,i_m,n)\,((m+1)!)^{-1}.$$

On the left hand side of (7.11), we have all events $A^{(n)}{}_{k_1,\ldots,k_m}$, where $k_1, k_2-k_1, \ldots, n-k_m$, if we disregard the order, is the same partition

of n into $m+1$ positive integers. If we now sum all the equations (7.11) corresponding to different partitions of n into $m+1$ positive integers, we obtain on the left hand side $\sum \Pr\{A^{(n)}{}_{k_1,\ldots,k_m}\}$, where the summation is over all sets of integers $k_1,\ldots,k_m$ for which

$$0 < k_1 < k_2 < \ldots < k_m < n \ .$$

It follows from (7.4) that the sum of the left hand sides of the equations (7.11) is $\Pr\{H_n=m\}$. The sum of the right hand sides is the sum over all partitions of n into $m+1$ positive integers. Each term, however, may be considered as the sum of equal terms of the form

$$\left(\prod_{\nu=0}^{m} (i_{\nu+1}-i_\nu)^{-1}\right)\left((m+1)!\right)^{-1} ,$$

each term corresponding to a permutation of the $m+1$ numbers in the partition. The sum of the right hand sides of the equations (7.11), therefore, is the sum, over all ordered partitions $i_1, i_2-i_1, \ldots, n-i_m$ of n into $m+1$ positive integers, of terms of the form

$$\left(\prod_{\nu=0}^{m} (i_{\nu+1}-i_\nu)^{-1}\right)\left((m+1)!\right)^{-1} .$$

If we make the substitution $j_\nu = i_{\nu+1}-i_\nu$, we obtain the right hand side of (7.2).

We shall now prove the relation (7.10) in order to complete the proof of (7.2). We first consider a fixed set $i_1,\ldots,i_m$ and show that if exactly p of the numbers $i_1, i_2-i_1, \ldots, i_m-i_{m-1}$ equal $n-i_m$, then

$$(7.12) \quad (p+1)\sum_{k_1,\ldots,k_m} \Pr\{B^{(n)}{}_{k_1,\ldots,k_m}\} = \sum_{k_1{}',\ldots,k_{m-1}{}'} \Pr\{B^{(i_m)}{}_{k_1{}',\ldots,k_{m-1}{}'}\} ,$$

where on the left hand side the sum is over those $k_1,\ldots,k_m$, for which $k_1, k_2-k_1, \ldots, n-k_m$ is a permutation of $i_1, i_2-i_1, \ldots, n-i_m$, while the sum on the right hand side is over those $k_1{}',\ldots,k_{m-1}{}'$ for which

$$k_1{}', k_2{}'-k_1{}', \ldots, i_m-k_{m-1}{}'$$

is a permutation of $i_1, i_2-i_1, \ldots, i_m-i_{m-1}$. Choosing μ, $0 \leqq \mu \leqq m$, such that $k_{\mu+1}-k_\mu = n-i_m$, the transformation

$$(7.13) \quad \begin{aligned} X_i &\to X_i & \text{for} \quad i &= 1,\ldots,k_\mu , \\ X_i &\to X_{i+i_m-k_\mu} & \text{for} \quad i &= k_\mu+1,\ldots,k_{\mu+1} , \\ X_i &\to X_{i+i_m-n} & \text{for} \quad i &= k_{\mu+1}+1,\ldots,n , \end{aligned}$$

transforms the random variables W_ν in the following way:

$$W_\nu \to W_\nu \qquad \text{for} \quad \nu = 0, \ldots, \mu-1\,,$$
$$W_\mu \to W_m\,,$$
$$W_\nu \to W_{\nu-1} \qquad \text{for} \quad \nu = \mu+1, \ldots, m\,,$$

and carries the event $B^{(n)}{}_{k_1,\ldots,k_m}$ into the new event,

(7.14)
$$B^{(im)}{}_{k_1,\ldots,k_\mu,k_{\mu+2}+i_m-n,\ldots,k_m+i_m-n}\,[W_{\mu-1}{}^{(k_\mu-k_{\mu-1})} < W_m{}^{(n-i_m)} < W_\mu{}^{(k_{\mu+2}-k_{\mu+1})}]\,,$$

with the same probability. The event $[W_{\mu-1} < W_m < W_\mu]$ for $\mu = 0$ or $\mu = m$ is to be understood as $[W_m < W_0]$ or $[W_{m-1} < W_m]$, respectively; for $\mu = m-1$ and $\mu = m$, the event $B^{(im)}{}_{k_1,\ldots,k_\mu,k_{\mu+2}+i_m-n,\ldots,k_m+i_m-n}$ is to be understood as $B^{(im)}{}_{k_1,\ldots,k_{m-1}}$. For each $B^{(n)}{}_{k_1,\ldots,k_m}$, we have $p+1$ transformations of the form (7.13). Furthermore any event of the form (7.14), where $B^{(im)}{}_{k_1,\ldots,k_\mu,k_{\mu+2}+i_m-n,\ldots,k_m+i_m-n}$ is one of the events on the right hand side of (7.12), is obtained once and only once when we apply the transformations (7.13) to the events $B^{(n)}{}_{k_1,\ldots,k_m}$. We therefore obtain

(7.15) $(p+1) \displaystyle\sum_{k_1,\ldots,k_m} \Pr\{B^{(n)}{}_{k_1,\ldots,k_m}\}$
$$= \sum_{k_1',\ldots k_{m-1}'} \sum_{\mu=0}^{m} \Pr\{B^{(im)}{}_{k_1',\ldots,k_{m-1}'}\,[W_{\mu-1} < W_m < W_\mu]\}\,.$$

Since the events $[W_{\mu-1} < W_m < W_\mu]$, $\mu = 0, \ldots, m$, are non-overlapping and $\Pr\{W_i = W_j\} = 0$ for $i \neq j$, we obtain

$$\sum_{\mu=0}^{m} \Pr\{W_{\mu-1} < W_m < W_\mu\} = 1$$

and

(7.16) $\displaystyle\sum_{\mu=0}^{m} \Pr\{B^{(im)}{}_{k_1',\ldots,k_{m-1}'}\,[W_{\mu-1} < W_m < W_\mu]\} = \Pr\{B^{(im)}{}_{k_1',\ldots,k_{m-1}'}\}\,.$

Formula (7.12) follows from (7.15) and (7.16).

We now assume that (7.10) holds if we replace n by i_m and the set $i_1, \ldots, i_m$ by $i_1, \ldots, i_{m-1}$. We then get, from (7.12), the equation

(7.17) $\displaystyle\sum_{k_1,\ldots,k_m} \Pr\{B^{(n)}{}_{k_1,\ldots,k_m}\} = (p+1)^{-1} \sum_{k_1',\ldots,k'_{m-1}} \Pr\{B^{(im)}{}_{k_1',\ldots,k_{m-1}'}\}$
$$= (p+1)^{-1}\, p(i_1, \ldots, i_m)\, (m!)^{-1}\,.$$

Since $p(i_1, \ldots, i_m)$ and $p(i_1, \ldots, i_m, n)$ are the number of permutations of the numbers $i_1, i_2-i_1, \ldots, i_m-i_{m-1}$; $i_1, i_2-i_1, \ldots, i_m-i_{m-1}, n-i_m$,

respectively, and there are p of the numbers $i_1, i_2-i_1, \ldots, i_m-i_{m-1}$, which equal $n-i_m$, we obtain

(7.18) $p(i_1, \ldots, i_m, n) = (m+1)\,(p+1)^{-1}\,p(i_1, \ldots, i_m)$.

If we eliminate $p(i_1, \ldots, i_m)$ from (7.17) and (7.18), we obtain (7.10). We can therefore use induction with respect to the number of lower indices i_m to prove (7.10), since we have $\Pr\{B^{(k)}\} = 1$ for all k.

REMARK. An alternative form of (7.2) is the following:

(7.19) $$\Pr\{H_n=m\} = \sum_{\alpha_1,\ldots,\alpha_n}^{(m)} \prod_{i=1}^{n} (\alpha_i!)^{-1}\, i^{-\alpha_i} ,$$

where $\sum^{(m)}$ is the summation over those values of $\alpha_1, \ldots, \alpha_n$ for which

$$\alpha_1+2\alpha_2+\ldots+n\alpha_n = n, \qquad \alpha_1+\ldots+\alpha_n = m+1 ,$$

and $\alpha_i \geqq 0$. This relation may be derived from (7.2) as follows. In each term of (7.2), we replace $j_0, \ldots, j_m$ by their values. The exponents α_i indicate the multiplicity of the value i. The condition

$$\alpha_1+\ldots+\alpha_n = m+1$$

insures that we have, in (7.19), $m+1$ factors in each term corresponding to $j_0, \ldots, j_m$; whereas the condition $\alpha_1+2\alpha_2+\ldots+n\alpha_n = n$ corresponds to $j_0+\ldots+j_m = m$. The number of terms for which $\prod_{\nu=0}^{m} j_\nu^{-1}$ represents $\prod_{i=1}^{n} i^{-\alpha_i}$ is $(m+1)!\, \prod_{i=1}^{n}(\alpha_i!)^{-1}$. We therefore obtain from (7.2) the relation (7.19).

PROOF OF (7.3): We have, for $|s| < 1$ and $0 < |t| < 1$,

(7.20) $$H(s, t) = \sum_{n=1}^{\infty} \sum_{m=0}^{n-1} \Pr\{H_n=m\}\, s^n t^m$$

$$= \sum_{n=1}^{\infty} \sum_{m=0}^{n-1} ((m+1)!)^{-1} \sum_{j_0,\ldots,j_m}^{(n)} \left(\prod_{\nu=0}^{m} j_\nu^{-1} \right) s^n t^m$$

$$= \sum_{m=0}^{\infty} ((m+1)!)^{-1} t^m \sum_{n=m+1}^{\infty} \sum_{j_0,\ldots,j_m}^{(n)} \prod_{\nu=0}^{m} j_\nu^{-1} s^{j_\nu}$$

$$= \sum_{m=0}^{\infty} ((m+1)!)^{-1} t^m \sum_{j_0=1}^{\infty} \cdots \sum_{j_m=1}^{\infty} \prod_{\nu=0}^{m} j_\nu^{-1} s^{j_\nu}$$

$$= \sum_{m=0}^{\infty} ((m+1)!)^{-1} t^m \prod_{\nu=0}^{m} \sum_{j_\nu=1}^{\infty} j_\nu^{-1} s^{j_\nu} =$$

$$= \sum_{m=0}^{\infty} ((m+1)!)^{-1} t^m \prod_{\nu=0}^{m} \left(-\log(1-s) \right)$$

$$= \sum_{m=0}^{\infty} ((m+1)!)^{-1} t^m \left(-\log(1-s) \right)^{m+1}$$

$$= t^{-1} \left(\exp\left(-t \log(1-s) \right) - 1 \right)$$

$$= t^{-1} \left((1-s)^{-t} - 1 \right).$$

The operations performed are all valid, since we have, for

$$|s| < 1, \qquad 0 < |t| < 1,$$

absolute convergence of the infinite series which occur in the operations.

8. The formulae (7.2) and (7.19) are both rather useless for numerical calculations. We shall, however, derive, $1°$ a method for calculation of the probability distribution of H_n for a fixed n, and $2°$ the limiting form of the distribution of H_n. We expand the function $t^{-1}((1-s)^{-t} - 1)$ in a power series in s; then the coefficient of s^n is the generating function

$$H_n(t) = \sum_{m=0}^{n-1} \Pr\{H_n = m\} t^m$$

of the sequence $\Pr\{H_n=0\}$, $\Pr\{H_n=1\}, \ldots, \Pr\{H_n=n-1\}$. We therefore obtain for $n = 1, 2, \ldots$ the relation

$$(8.1) \quad H_n(t) = \sum_{m=0}^{n-1} \Pr\{H_n = m\} t^m = t^{-1}(-1)^n \binom{-t}{n} = n^{-1} \prod_{m=1}^{n-1} (1 + m^{-1}t).$$

If we compare (8.1) for two consecutive values of n, we obtain

$$(8.2) \qquad H_{n+1}(t) = H_n(t)\, n(n+1)^{-1}(1 + n^{-1}t).$$

From the relation (8.2),

$$(8.3) \quad \Pr\{H_{n+1} = m\} = (n+1)^{-1} \left(\Pr\{H_n = m-1\} + n \Pr\{H_n = m\} \right),$$

$$m = 0, 1, \ldots, n,$$

follows by comparison of coefficients of t^m on both sides.

We shall now use (8.3) to show, by induction with respect to n, that if $s_m(n) = \sum_{i=1}^{n-1} i^{-m}$, then, for $n = 1, 2, \ldots$ and $m = 0, 1, \ldots, n-1$,

$$(8.4) \qquad m \Pr\{H_n = m\} = \sum_{i=1}^{m} (-1)^{i+1} s_i(n) \Pr\{H_n = m-i\}.$$

The relation (8.4) evidently holds for $n = 1$, since both sides equal 0 in this case. As for $m = 0$, the sum on the right hand side is empty. We assume that (8.4) holds and shall show that the corresponding relation with n replaced by $n+1$ also holds. We have

$$\sum_{i=1}^{m} (-1)^{i+1} s_i(n+1) \Pr\{H_{n+1}=m-i\}$$

$$= \sum_{i=1}^{m} (-1)^{i+1} s_i(n+1) (n+1)^{-1}\big(\Pr\{H_n=m-i-1\} + n \Pr\{H_n=m-i\}\big)$$

$$= \sum_{i=1}^{m} (-1)^{i+1} s_i(n) (n+1)^{-1} \Pr\{H_n=m-i-1\} +$$

$$+ \sum_{i=1}^{m} (-1)^{i+1} n^{-i}(n+1)^{-1} \Pr\{H_n=m-i-1\} +$$

$$+ \sum_{i=1}^{m} (-1)^{i+1} s_i(n) (n+1)^{-1} n \Pr\{H_n=m-i\} +$$

$$+ \sum_{i=1}^{m} (-1)^{i+1} n^{-i}(n+1)^{-1} n \Pr\{H_n=m-i\}$$

$$= (n+1)^{-1}(m-1) \Pr\{H_n=m-1\} + (n+1)^{-1}nm \Pr\{H_n=m\} +$$
$$+ (n+1)^{-1} \Pr\{H_n=m-1\}$$

$$= m(n+1)^{-1}\big(\Pr\{H_n=m-1\} + n \Pr\{H_n=m\}\big) = m \Pr\{H_{n+1}=m\}\,.$$

The formula (8.4) is useful, since, as we shall see below, $E(H_n) \sim \log n$ and $V(H_n) \sim \log n$, for $n \to \infty$, so that $\Pr\{H_n=m\}$ is small if m is considerably larger than $\log n$.

The generating function,

$$H_n(t) = n^{-1} \prod_{m=1}^{n-1} (1 + m^{-1}t) = \prod_{m=1}^{n-1} \big((m+t)(m+1)^{-1}\big)\,,$$

also shows that H_n has the same distribution function as a random variable which is the sum of $n-1$ independent random variables $Y_1, \ldots, Y_{n-1}$, where Y_i, for $i = 1, \ldots, n-1$, assumes the values 0 and 1 with probabilities $i(i+1)^{-1}$ and $(i+1)^{-1}$. Since $E(Y_i) = (i+1)^{-1}$ and $E(Y_i^2) = (i+1)^{-1}$, we obtain $V(Y_i) = i(i+1)^{-2}$. We therefore have

$$E(H_n) = \sum_{i=1}^{n-1} (i+1)^{-1}$$

and

$$V(H_n) = \sum_{i=1}^{n-1} i\,(i+1)^{-2} = \sum_{i=1}^{n-1} (i+1)^{-1} - \sum_{i=1}^{n-1} (i+1)^{-2}\,.$$

If we apply the central limit theorem to the random variables $Y_1, \ldots, Y_{n-1}$, it follows that the distribution of H_n is, for large values of n, approximately normal. Since the values of H_n are integers, and we have $E(H_n) \sim \log n$, $V(H_n) \sim \log n$ for $n \to \infty$, we also get an approximation to the distribution of H_n if we use a Poisson distribution with mean $\log n$. The convergence to the normal or the Poisson distribution, however, is very slow, since the mean of H_n increases only as $\log n$.

REMARK. It is known (see e. g. Feller [9, p. 205]) that the distribution of the number of cycles in a random permutation of n elements, if we assume that all $n!$ permutations are equally probable, has the same distribution as the sum of n independent random variables $Z_1, \ldots, Z_n$, where Z_i assumes the values 0 and 1 with probabilities $(i-1)i^{-1}$ and i^{-1}. If we let $k_{n,m}$ denote the number of permutations of n elements for which the number of cycles is m, we therefore obtain

$$k_{n,m} = n!\,\mathrm{Pr}\,\{H_n = m-1\}\,.$$

9. We shall consider the distribution of the random variable H_n under the assumption that the random variables $W_\nu^{(i_{\nu+1}-i_\nu)}$ defined in section 7, p. 211, assume values which belong to some given sets. These sets may e. g. be $[W_\nu^{(k)} > 0]$. It is now necessary to assume that the random variables $X_1, \ldots, X_n$ are independent and identically distributed, since we should otherwise have to introduce probabilities, e. g. of the type $\mathrm{Pr}\,\{\bigcap_{\nu=0}^{i}[W_\nu^{(k)} > 0]\}$, in the theorem below.

THEOREM 6. *Let the random variables* $X_1, X_2, \ldots$ *be independent and identically distributed, and let the common distribution function be continuous. Let* H_n *be defined as in Theorem 5, and let* $F_1, F_2, \ldots$ *be given point sets on the real axis; let* $a_k = \mathrm{Pr}\,\{k^{-1}S_k \in F_k\}$. *Let* $C_{n,m}$ *be the event in which* $H_n = m$ *and, furthermore, for* $\nu = 0, \ldots, m$, *the random variables* $W_\nu^{(i_{\nu+1}-i_\nu)}$ *assume values belonging to* $F_{i_{\nu+1}-i_\nu}$. *Then*

$$(9.1) \qquad \mathrm{Pr}\,\{C_{n,m}\} = \big((m+1)!\big)^{-1} \sum_{j_1,\ldots,j_m}^{[n]} \prod_{\nu=0}^{m} (a_{j_\nu}\, j_\nu^{-1})\,,$$

where $\sum^{[n]}$ *indicates summation over those values of* $j_0, \ldots, j_m$ *for which we have* $j_\nu > 0$ *and* $j_0 + \ldots + j_m = n$.

Furthermore, we have, for $|s| < 1$ *and* $0 < |t| < 1$,

$$(9.2) \quad H^*(s, t) = \sum_{n=1}^{\infty} \sum_{m=0}^{n-1} \Pr\{C_{n,m}\} \, s^n t^m = t^{-1} \left(\exp\left(t \int_0^s \sigma^{-1} A(\sigma) d\sigma \right) - 1 \right),$$

where $A(s)$ is the generating function $\sum_{n=1}^{\infty} a_n s^n$ of the sequence $a_1, a_2, \ldots$.

PROOF OF (9.1): This proof is a modification of the proof of (7.2). We shall therefore, when possible, refer to the corresponding steps in the proof of (7.2). Symbols introduced in section 7 will have the same meaning here.

Corresponding to (7.4), we have

$$(9.3) \quad \Pr\{C_{n,m}\} = \sum_{i_1, \ldots, i_m} \Pr\left\{ A^{(n)}{}_{i_1, \ldots, i_m} \bigcap_{\nu=0}^{m} [W_\nu^{(i_{\nu+1}-i_\nu)} \in F_{i_{\nu+1}-i_\nu}] \right\},$$

where the summation again is over all possible sets of m integers $i_1, \ldots, i_m$ for which $0 < i_1 < \ldots < i_m < n$. Since the events $A^{(n)}{}_{i_1, \ldots, i_m}$ and $\bigcap_{\nu=0}^{m} [W_\nu^{(i_{\nu+1}-i_\nu)} \in F_{i_{\nu+1}-i_\nu}]$ may be non-overlapping, some of the events in the summation may be empty.

The event $\bigcap_{\nu=0}^{m} [W_\nu^{(i_{\nu+1}-i_\nu)} \in F_{i_{\nu+1}-i_\nu}]$ is symmetric with respect to the variables $X_{i_\nu+1}, \ldots, X_{i_{\nu+1}}$ for $\nu = 0, \ldots, m$, so that we may use Theorem D as in the proof of (7.8) and thereby obtain

$$(9.4) \quad \Pr\left\{ A^{(n)}{}_{i_1, \ldots, i_m} \bigcap_{\nu=0}^{m} [W_\nu^{(i_{\nu+1}-i_\nu)} \in F_{i_{\nu+1}-i_\nu}] \right\}$$

$$= \left(\prod_{\nu=0}^{m} (i_{\nu+1} - i_\nu)^{-1} \right) \Pr\left\{ B^{(n)}{}_{i_1, \ldots, i_m} \bigcap_{\nu=0}^{m} [W_\nu^{(i_{\nu+1}-i_\nu)} \in F_{i_{\nu+1}-i_\nu}] \right\}.$$

Corresponding to (7.10), we have the relation

$$(9.5) \quad \sum_{k_1, \ldots, k_m} \Pr\left\{ B^{(n)}{}_{k_1, \ldots, k_m} \bigcap_{\nu=0}^{m} [W_\nu^{(k_{\nu+1}-k_\nu)} \in F_{k_{\nu+1}-k_\nu}] \right\}$$

$$= p(i_1, \ldots, i_m, n) \left((m+1)! \right)^{-1} \prod_{\nu=0}^{m} (a_{i_{\nu+1}-i_\nu}),$$

where $\sum_{k_1, \ldots, k_m}$ has the same meaning as in (7.10). Before we turn to the proof of this relation, we note that (9.1) follows from the relations (9.3), (9.4), and (9.5) in the same way in which (7.2) followed from (7.4), (7.8), and (7.10).

The proof of (9.5) also follows the proof of (7.10) with small modifications arising from the events $[W_\nu^{(k_{\nu+1}-k_\nu)} \in F_{k_{\nu+1}-k_\nu}]$. Corresponding to (7.12), we first prove, for a fixed set of indices $i_1, \ldots, i_m$, the relation

$$(9.7) \qquad (p+1) \sum_{k_1,\ldots,k_m} \Pr \left\{ B^{(n)}{}_{k_1,\ldots,k_m} \bigcap_{\nu=0}^{m} [W_\nu^{(k_{\nu+1}-k_\nu)} \in F_{k_{\nu+1}-k_\nu}] \right\}$$

$$= a_{n-i_m} \sum_{k_1',\ldots,k_{m-1}'} \Pr \left\{ B^{(i_m)}{}_{k_1',\ldots,k_{m-1}'} \bigcap_{\nu=0}^{m-1} [W_\nu^{(k_{\nu+1}'-k_\nu')} \in F_{k_{\nu+1}'-k_\nu'}] \right\},$$

where $\sum_{k_1,\ldots,k_m}$ and $\sum_{k_1',\ldots,k_{m-1}'}$ have the same meaning as in (7.12), and where we assume that p of the numbers $i_1, i_2-i_1, \ldots, i_m-i_{m-1}$ are equal to $n-i_m$. We consider one of the events

$$(9.8) \qquad B^{(n)}{}_{k_1,\ldots,k_m} \bigcap_{\nu=0}^{m} [W_\nu^{(k_{\nu+1}-k_\nu)} \in F_{k_{\nu+1}-k_\nu}] .$$

The transformation (7.13) carries this event into the new event

$$(9.9) \qquad B^{(i_m)}{}_{k_1,\ldots,k_\mu,k_{\mu+2+i_m-n}} \bigcap_{\nu=0}^{\mu=1} [W_\nu^{(k_{\nu+1}-k_\nu)} \in F_{k_{\nu+1}-k_\nu}]$$

$$\bigcap_{\nu=\mu}^{m-1} [W_\nu^{(k_{\nu+2}-k_{\nu+1})} \in F_{k_{\nu+2}-k_{\nu+1}}][W_m^{(n-i_m)} \in F_{n-i_m}][W_{\mu-1} < W_m < W_\mu] .$$

For each of the events (9.8), we have $p+1$ transformations of the form (7.13). Furthermore any event of the form (9.9), where

$$B^{(i_m)}{}_{k_1,\ldots,k_\mu,k_{\mu+2+i_m-n}} \bigcap_{\nu=0}^{\mu-1} [W_\nu^{(k_{\nu+1}-k_\nu)} \in F_{k_{\nu+1}-k_\nu}] \bigcap_{\nu=\mu}^{m-1} [W_\nu^{(k_{\nu+2}-k_{\nu+1})} \in F_{k_{\nu+2}-k_{\nu+1}}]$$

$$(9.10)$$

is one of the events on the right hand side of (9.7), is obtained once and only once, if we disregard the cases where the events (9.8) and (9.10) are both empty. We obtain, analogously to (7.15) and (7.16), the relations

$$(9.11) \qquad (p+1) \sum_{k_1,\ldots,k_m} \Pr \{ B^{(n)}{}_{k_1,\ldots,k_m} \}$$

$$= \sum_{k_1',\ldots,k_{m-1}'} \sum_{\mu=0}^{m} \Pr \left\{ B^{(i_m)}{}_{k_1',\ldots,k_{m-1}'} \bigcap_{\nu=0}^{m-1} [W_\nu^{(k_{\nu+1}'-k_\nu')} \in F_{k_{\nu+1}'-k_\nu'}] \right.$$

$$\left. [W_m^{(n-i_m)} \in F_{n-i_m}][W_{\mu-1} < W_m < W_\mu] \right\}$$

and

$$(9.12) \qquad \sum_{\mu=0}^{m-1} \Pr\left\{ B^{(im)}{}_{k_1{}',\ldots,k_{m-1}{}'} \bigcap_{\nu=0}^{m-1} [W_\nu^{(k_{\nu+1}{}'-k_\nu{}')} \in F_{k_{\nu+1}{}'-k_\nu{}'}] \right.$$

$$\left. [W_m^{(n-im)} \in F_{n-im}][W_{\mu+1} < W_m < W_\mu] \right\}$$

$$= \Pr\left\{ B^{(im)}{}_{k_1{}',\ldots,k_{m-1}{}'} \bigcap_{\nu=0}^{m-1} W_\nu^{(k_{\nu+1}{}'-k_\nu{}')} \in F_{k_{\nu+1}{}'-k_\nu{}'}][W_m^{(n-im)} \in F_{n-im}] \right\}.$$

We now use the fact that we have assumed, in this section, that the random variables $X_1, X_2, \ldots$ are independent and that the event $B^{(im)}{}_{k_1{}',\ldots,k_{m-1}{}'} \bigcap_{\nu=0}^{m-1} [W_\nu^{(k_{\nu+1}{}'-k_\nu{}')} \in F_{k_{\nu+1}{}'-k_\nu{}'}]$ depends only on $X_1, \ldots, X_{i_m}$, while $[W_m^{(n-im)} \in F_{n-im}]$ depends only on $X_{i_m+1}, \ldots, X_n$, so that we may use the multiplication rule for probabilities of independent events and thereby obtain (9.7) from (9.11) and (9.12).

The relation (9.5) now follows from (9.7) in almost the same way in which (7.10) followed from (7.12). We need only observe that the factor $\prod_{\nu=0}^{m} a_{i_{\nu+1}-i_\nu}$ goes into the induction together with $p(i_1, \ldots, i_m, n)$ and use the fact that $\Pr\{B^{(k)}[W_0^{(k)} \in F_k]\} = \Pr\{W_0^{(k)} \in F_k\} = a_k$.

An alternative form of $\Pr\{C_{n,m}\}$ corresponding to (7.19) is

$$(9.13) \qquad \Pr\{C_{n,m}\} = \sum_{\alpha_1,\ldots,\alpha_n}^{(m)} \prod_{i=1}^{n} (\alpha_i!)^{-1}(a_i i^{-1})^{\alpha_i};$$

the proof is an easy modification of the proof of (7.19).

PROOF OF (9.2): We have, for $|s| < 1$, $0 < |t| < 1$,

$$H^*(s, t) = \sum_{n=1}^{\infty} \sum_{m=0}^{n-1} \Pr\{C_{n,m}\} s^n t^m = t^{-1}\left(\exp\left(t \sum_{j=1}^{\infty} a_j j^{-1} s^j \right) - 1 \right),$$

which may be seen by calculations similar to those of (7.20). We do, however, have $A(s) = \sum_{j=1}^{\infty} a_j s^j$ and $\int_0^s \sigma^{-1} A(\sigma) d\sigma = \sum_{j=1}^{\infty} a_j j^{-1} s^j$, so that we obtain (9.2).

We shall now consider the probability of the event

$$C_n = C_{n,0} \cup \ldots \cup C_{n,n-1}.$$

In general this event is connected with the random variables $X_1, \ldots, X_n$ and the sums $S_0, S_1, \ldots, S_n$ in a rather complicated way. When, however, we choose the point sets F_k to be, for all k, the set of all positive numbers, then it follows that C_n is the event at which the largest convex minorant sequence $T_0, T_1, \ldots, T_n$ to the sequence $S_0, S_1, \ldots, S_n$ assumes only

positive values (except, of course, $T_0 = 0$). Since we have $S_i \geqq T_i$, it follows that in C_n the sums $S_0, \ldots, S_n$ assume only positive values. Conversely, if $S_1, \ldots, S_n$ are all positive, then $T_1, \ldots, T_n$ are also positive. Therefore, in this case, we have

$$C_n = \bigcap_{i=1}^{n} [S_i > 0] = [N_n = n] = [M_n = 0] .$$

From the definition of C_n and the relation (9.13), the formula

$$(9.14) \qquad \Pr\{C_n\} = \sum_{m=0}^{n-1} \sum_{\alpha_1, \ldots, \alpha_n}^{(m)} \prod_{i=1}^{n} (\alpha_i!)^{-1} (a_i i^{-1})^{\alpha_i}$$

$$= \sum_{\alpha_1, \ldots, \alpha_n}^{*} \prod_{i=1}^{n} (\alpha_i!)^{-1} (a_i i^{-1})^{\alpha_i} ,$$

follows.

We have, in this way, obtained a new proof of the formula (3.1) of Theorem 1. (In the case $K_n = L_n$, we use Theorem A.) This proof is valid only if the common distribution function of $X_1, \ldots, X_n$ is continuous. It may, however, be extended to the general case, if we introduce auxiliary random variables $Y_1, Y_2, \ldots$ and use the method indicated in [3, section 7].

REFERENCES

1. E. Sparre Andersen, *On the number of positive sums of random variables*, Skand. Aktuarie-tidskr. 32 (1949), 27–36.
2. E. Sparre Andersen, *On the frequency of positive partial sums of a series of random variables*, Mat. Tidsskr. B 1950, 33–35.
3. E. Sparre Andersen, *On sums of symmetrically dependent random variables*, Skand. Aktuarietidskr. 36 (1953), 123–138.
4. E. Sparre Andersen, *On the fluctuations of sums of random variables*, Math. Scand. 1 (1953), 263–285.
5. E. Sparre Andersen, *Two summation formulae for product sums of binomial coefficients*, Math. Scand. 1 (1953), 261–262.
6. K. L. Chung and W. Feller, *On fluctuations in coin-tossing*, Proc. Nat. Acad. Sci. U. S. A. 35 (1949), 605–608.
7. D. A. Darling, *Sums of symmetrical random variables*, Proc. Amer. Math. Soc. 2 (1951), 511–517.
8. P. Erdös and M. Kac, *On the number of positive sums of independent random variables*, Bull. Amer. Math. Soc. 53 (1947), 1011–1020.
9. W. Feller, *An introduction to probability theory and its applications* I, New York, 1950.
10. P. Lévy, *Sur certaines processus stochastiques homogènes*, Compositio Math. 7 (1939), 283–339.
11. M. Lipschutz, *Generalization of a theorem of Chung and Feller*, Proc. Amer. Math. Soc. 3 (1952), 659–670.

12. M. Lipschutz, *On strong laws for certain types of events connected with sums of independent random variables,* Ann. of Math. 57 (1953), 318–330.
13. G. Maruyama, *Note on the arcsine law in the theory of probability,* Nat. Sci. Rep. Ochanomizu Univ. 2 (1951), 25–27.
14. M. Udagawa, *On numbers of positive sums of independent random variables,* Kōdai Math. Sem. Rep. 1952, 45–50.

THE LIFE INSURANCE COMPANY
PENSIONS- OG LIVRENTE-INSTITUTET AF 1919 A/S,
COPENHAGEN, DENMARK