SOME GENERALIZATIONS OF THE THEORY OF CUMULATIVE SUMS OF RANDOM VARIABLES

By Abraham Wald

Columbia University

1. Introduction. In a previous paper [1] the author dealt with the following problem: Let $\{z_i\}$ $(i = 1, 2, \cdots$, ad inf.) be a sequence of independently distributed random variables each having the same distribution. Let a be a given positive constant, b a given negative constant and denote by n the smallest positive integer for which either

$$z_1 + \cdots + z_n \geq a \tag{1}$$

or

$$z_1 + \cdots + z_n \leq b \tag{2}$$

holds. The main problems treated in [1] were: (1) Derivation of the probability that the cumulative sum reaches the boundary a before the boundary b is reached; (2) Derivation of the characteristic function and the distribution function of n.

In this paper we shall consider the following more general problem: Let $K = \{k_i(z_1, \cdots, z_i)\}$ $(i = 1, 2, \cdots$, ad inf.) be a given sequence of functions and let n be the smallest positive integer for which either

$$k_n(z_1, \cdots, z_n) \geq 1 \tag{3}$$

or

$$k_n(z_1, \cdots, z_n) \leq -1 \tag{4}$$

holds. No restrictions are imposed on the sequence K except that it must be such that the probability that $n < \infty$ is equal to one. The purpose of this paper is to derive some theorems concerning the probability that $k_n(z_1, \cdots, z_n) \geq 1$ and concerning the expected value of n. Obviously, the problem formulated here is a generalization of that considered in [1], since the latter can be obtained by putting $k_i(x_1, \cdots, x_i) = \dfrac{2}{a-b}(z_1 + \cdots + z_i) - \dfrac{a+b}{a-b}$.

2. The conjugate distribution of z. Let z be a random variable whose distribution is equal to the common distribution of z_i. In this section we shall introduce the notion of the conjugate distribution of z which will be used later. According to Lemma 2 in [1], under some weak restrictions on the distribution of z there exists exactly one real value $h_0 \neq 0$ such that

$$E(e^{zh_0}) = 1 \tag{5}$$

where $E(u)$ denotes the expected value of u for any random variable u.

For simplicity we shall assume that z has a continuous distribution admitting a probability density everywhere, or that z has a discrete distribution. By the probability distribution $f(z)$ of z we shall mean the probability density of z, if the distribution of z is continuous. In the discrete case $f(z)$ will denote the probability that the random variable takes the value z. From (5) it follows that

$$f^*(z) = e^{zh_0}f(z) \tag{6}$$

is a probability distribution. We shall call $f^*(z)$ the conjugate distribution of z. For any random variable u we shall denote by $E^*(u)$ the expected value of u under the assumption that the distribution of z is given by $f^*(z)$. The expected values $E(u)$ and $E^*(u)$ may depend on the sequence $K = \{k_i(z_1, \cdots, z_i)\}$ $(i = 1, 2, \cdots, \text{ad inf.})$. Occasionally we shall put this dependence in evidence by writing $E(u \mid K)$ and $E^*(u \mid K)$, respectively.

3. Two theorems. In this section we shall derive two theorems. The first theorem is concerned with the probability that $k_n(z_1, \cdots, z_n) \geq 1$ and the second theorem with the expected value of n. In what follows the operator E_1 will mean conditional expected value under the restriction that $k_n(z_1, \cdots, z_n) \geq 1$ and E_2 will mean conditional expected value under the restriction that $k_n(z_1, \cdots, z_n) \leq -1$. If the distribution of z is given by $f^*(z)$, these conditional expected values will be denoted by the operators E_1^* and E_2^*, respectively.

THEOREM 1. *Let $K = \{k_i(z_1, \cdots, z_i)\}$ be a sequence such that the probability that $n < \infty$ is equal to one under both distributions $f(z)$ and $f^*(z)$. Let γ denote the probability that $k_n(z_1, \cdots, z_n) \geq 1$ when $f(z)$ is the distribution of z, and let γ^* denote the probability of the same event when $f^*(z)$ is the distribution of z. Then*

$$E_1(e^{Z_n h_0} \mid K) = \frac{\gamma^*}{\gamma}; \qquad E_2(e^{Z_n h_0} \mid K) = \frac{1 - \gamma^*}{1 - \gamma} \tag{7}$$

and

$$E_1^*(e^{-Z_n h_0} \mid K) = \frac{\gamma}{\gamma^*}; \qquad E_2^*(e^{-Z_n h_0} \mid K) = \frac{1 - \gamma}{1 - \gamma^*} \tag{8}$$

where $Z_n = z_1 + \cdots + z_n$.

PROOF: From (6) it follows that

$$e^{Z_n h_0} = \frac{f^*(z_1) \cdots f^*(z_n)}{f(z_1) \cdots f(z_n)} \tag{9}$$

and

$$e^{-Z_n h_0} = \frac{f(z_1) \cdots f(z_n)}{f^*(z_1) \cdots f^*(z_n)}. \tag{10}$$

A set $(z_1, \cdots, z_n)$ will be said to be of type 1 if and only if $-1 < k_m(z_1, \cdots, z_m) < 1$ for $m = 1, \cdots, n - 1$ and $k_n(z_1, \cdots, z_n) \geq 1$. Similarly a set $(z_1, \cdots z_n)$ will be said to be of type 2 if and only if $-1 < k_m(z_1, \cdots, z_m) < 1$ for $m = 1, \cdots, n - 1$ and $k_n(z_1, \cdots, z_n) \leq -1$.

We shall prove Theorem 1 under the assumption that the distribution of z is discrete. Because of (9) we have

$$(11)\quad E_1(e^{Z_n h_0} \mid K) = E_1\left(\frac{f^*(z_1) \cdots f^*(z_n)}{f(z_1) \cdots f(z_n)} \,\middle|\, K\right) = \frac{\sum_{(z_1, \cdots, z_n)} f^*(z_1) \cdots f^*(z_n)}{\sum_{(z_1, \cdots, z_n)} f(z_1) \cdots f(z_n)}$$

where the summation is to be taken over all sets $(z_1, \cdots, z_n)$ of type 1. But the last expression is obviously equal to $\frac{\gamma^*}{\gamma}$ and, therefore, the first equation in (7) is proved. The second equation in (7) follows in the same manner if we take into account the fact that the probability that $n < \infty$ is equal to one. Similarly, equation (8) can be obtained from (10). The proof can easily be extended to the case when the distribution of z is continuous. Hence, Theorem 1 is proved.

THEOREM 2. *If $Ez \neq 0$, the relation*

$$(12)\qquad E(n \mid K) = \frac{E(Z_n \mid K)}{Ez}$$

holds for any sequence $K = \{k_i(z_1, \cdots, z_i)\}$ for which one of the following two conditions is fulfilled:

(a) *There exists an integer N such that the probability that $n < N$ is equal to one.*

(b) *$E(n \mid K) < \infty$ and the first four moments of z are finite.*

PROOF: First we shall show that condition (a) implies the validity of (12). For any integer i we shall denote $z_1 + \cdots + z_i$ by Z_i. Since the probability that $n < N$ is equal to 1, we have

$$(13)\qquad E(Z_n \mid K) + E(z_{n+1} + \cdots + z_N) = EZ_N = NEz.$$

Since the conditional expected value of $(z_{n+1} + \cdots + z_N)$ for a given value of n is equal to $(N - n)Ez$, we have

$$(14)\qquad E(z_{n+1} + \cdots + z_N) = E(N - n \mid K)Ez = NEz - E(n \mid K)Ez.$$

Equation (12) follows from (13) and (14).

Now we shall show that condition (b) implies (12). Denote by P_N the probability that $n \leq N$. Let the operator E_N denote conditional expected value under the restriction that $n \leq N$, and let the operator E'_N denote conditional expected value under the restriction that $n > N$. Then we have

$$(15)\qquad P_N E_N(Z_N) + (1 - P_N)E'_N(Z_N) = E(Z_N) = NEz.$$

Since

$$\begin{aligned} & = E_N(Z_n \mid K) + E_N(z_{n+1} + \cdots + z_N \mid K) \\ (16)\quad E_N(Z_N) & = E_N(Z_n \mid K) + E_N(N - n \mid K)Ez \\ & = E_N(Z_n \mid K) + NEz - E_N(n \mid K)Ez, \end{aligned}$$

we obtain from (15)

(17) $P_N\{E_N(Z_n \mid K) + NEz - E_N(n \mid K)Ez\} + (1 - P_N)E'_N(Z_N) = NEz.$

From $E(n \mid K) < \infty$ it follows that

$$\lim_{N=\infty} (1 - P_N)N = 0. \tag{18}$$

Now we shall show that (18) implies the validity of

$$\lim_{N=\infty} (1 - P_N)E'_N(Z_N) = 0. \tag{19}$$

Let $T_N = Z_N - NEz$. Because of (18), (19) is proved if we can show that

$$\lim_{N=\infty} (1 - P_N)E'_N(T_N) = 0. \tag{20}$$

Denote by R_N the set of all points $(z_1, \cdots, z_N)$ for which $n > N$. Then the probability measure of R_N is equal to $1 - P_N$ and

$$(1 - P_N)E'_N(T_N) = \int_{R_N} T_N f(z_1) \cdots f(z_N)\, dz_1 \cdots dz_N. \tag{21}$$

Let R_N^1 be the part of R_N in which $T_N < -N$, R_N^2 the part of R_N in which $T_N > N$ and R_N^3 the part of R_N in which $-N \leq T_N \leq N$. Because of (18) we have

$$\lim_{N=\infty} \left| \int_{R_N^3} T_N f(z_1) \cdots f(z_N)\, dz_1 \cdots dz_N \right| \leq \lim_{N=\infty} (1 - P_N)N = 0. \tag{22}$$

Denote the cumulative distribution function of T_N by $F_N(T_N)$. Clearly,

$$\int_{R_N^2} T_N f(z_1) \cdots f(z_N)\, dz_1 \cdots dz_N \leq \int_N^\infty T_N\, dF_N(T_N) \leq \frac{1}{N^3}\int_N^\infty T_N^4\, dF_N(T_N). \tag{23}$$

Since the first four moments of z are finite, the 4-th moment of $\frac{T_N}{\sqrt{N}}$ converges to $3\sigma^4$ where σ is the standard deviation of z. Hence

$$\lim_{N=\infty} \int_{-\infty}^{+\infty} \frac{1}{N^2} T_N^4 dF_N(T_N) = 3\sigma^4. \tag{24}$$

From (23) and (24) it follows that

$$\lim_{N=\infty} \int_{R_N^2} T_N f(z_1) \cdots f(z_N)\, dz_1 \cdots dz_N = 0. \tag{25}$$

Similarly we can prove that

$$\lim_{N=\infty} \int_{R'_N} T_N f(z_1) \cdots f(z_N)\, dz_1 \cdots dz_N = 0. \tag{26}$$

Equation (20) follows from (21), (22), (25) and (26). Hence (19) is proved. From (17), (18) and (19) we obtain

$$\lim_{N=\infty} P_N\{E_N(Z_n \mid K) - E_N(n \mid K)Ez\} = 0. \tag{27}$$

Since $Ez \neq 0$, $\lim P_N = 1$, $\lim E_N(n \mid K) = E(n \mid K)$ and $\lim E_N(\mathbf{Z}_n \mid K) = E(\mathbf{Z}_n \mid K)$, equation (12) follows from (27). Hence condition (b) implies (12) and Theorem 2 is proved.

4. Lower limit of $\mathbf{E}(n \mid K)$. In this section we shall derive a lower limit for $E(n \mid K)$. First we shall prove the following lemma.

LEMMA 1. *For any random variable u we have*

$$e^{E(u)} \leq Ee^{u}. \tag{28}$$

PROOF: Inequality (28) can be written as

$$1 \leq Ee^{u'} \tag{29}$$

where $u' = u - Eu$. Lemma 1 is proved if we show that (29) holds for any random variable u' whose mean is zero. Expanding $e^{u'}$ in a Taylor series around $u' = 0$, we obtain

$$e^{u'} = 1 + u' + \frac{u'^2}{2} e^{\xi(u')} \qquad \text{where } 0 \leq \xi(u') \leq u'.$$

Hence

$$Ee^{u'} = 1 + \tfrac{1}{2}Eu'^2 e^{\xi(u')} \geq 1$$

and Lemma 1 is proved.

Now we are able to prove the following theorem.

THEOREM 3. *Let $K = \{K_i(z_1, \cdots, z_i)\}$ be a sequence of functions such that the probability that $n < \infty$ is one under both distributions $f(z)$ and $f^*(z)$ of z. Let γ be the probability that $K_n(z_1, \cdots, z_n) \geq 1$ when $f(z)$ is the distribution of z, and let γ^* be the probability of the same event when $f^*(z)$ is the distribution of z. Then*

$$E(n \mid K) \geq \frac{1}{h_0 Ez}\left[\gamma \log \frac{\gamma^*}{\gamma} + (1 - \gamma) \log \frac{1 - \gamma^*}{1 - \gamma}\right] \tag{30}$$

and

$$E^*(n \mid K) \geq \frac{1}{h_0 Ez^*}\left[\gamma^* \log \frac{\gamma^*}{\gamma} + (1 - \gamma^*) \log \frac{1 - \gamma^*}{1 - \gamma}\right], \tag{31}$$

provided that Ez and Ez^* are not equal to zero.

PROOF: First we shall prove Theorem 3 in the case when there exists an integer N such that the probability that $n < N$ is one. According to Theorem 2 we have

$$E(n \mid K) = \frac{E(\mathbf{Z}_n \mid K)}{Ez} = \frac{1}{Ez}[\gamma E_1(Z_n \mid K) + (1 - \gamma)E_2(\mathbf{Z}_n \mid K)]. \tag{32}$$

From Lemma 1 and Theorem 1 it follows that

$$h_0 E_1(Z_n \mid K) \leq \log \frac{\gamma^*}{\gamma} \qquad \text{and } h_0 E_2(\mathbf{Z}_n \mid K) \leq \log \frac{1 - \gamma^*}{1 - \gamma}. \tag{33}$$

From (32) and (33) we obtain

$$(34)\qquad \begin{aligned} h_0 Ez\, E(n \mid K) &= h_0[\gamma E_1(Z_n \mid K) \\ &\quad + (1-\gamma)E_2(Z_n \mid K)] \le \gamma \log \frac{\gamma^*}{\gamma} + (1-\gamma)\log\frac{1-\gamma^*}{1-\gamma}. \end{aligned}$$

Inequality (30) follows from (34) if we can show that $h_0E(z) < 0$. From $Ee^{h_0 z} = 1$ and Lemma 1 it follows that $h_0E(z) \le 0$. Since $h_0 \ne 0$ and $E(z) \ne 0$, we must have $h_0E(z) < 0$. Hence (30) is proved. To prove (31) we proceed as follows: From Theorem 2 we obtain

$$(35)\qquad -h_0Ez^*E^*(n \mid K) = -h_0[\gamma^*E_1^*(Z_n \mid K) + (1-\gamma^*)E_2^*(Z_n \mid K)].$$

From Lemma 1 and Theorem 1 it follows that

$$(36)\qquad \begin{aligned} &-h_0[\gamma^*E_1^*(Z_n \mid K) + (1-\gamma^*)E_2^*(Z_n \mid K)] \\ &\qquad \le \gamma^* \log\frac{\gamma}{\gamma^*} + (1-\gamma^*)\log\frac{1-\gamma}{1-\gamma^*}. \end{aligned}$$

From (35) and (36) we obtain

$$(37)\qquad h_0 E^*(z)E^*(n \mid K) \ge \gamma^* \log\frac{\gamma^*}{\gamma} + (1-\gamma^*)\log\frac{1-\gamma^*}{1-\gamma}.$$

Since $E^*e^{-h_0 z} = 1$ it follows from Lemma 1 that $-h_0E^*z \le 0$. Inequality (31) follows from this and (37). Hence Theorem 3 is proved in the special case when there exists an integer N such that the probability that $n < N$ is equal to one.

To prove Theorem 3 in the general case, for any integer N let the sequence $K_N = \{k_{iN}(z_1, \cdots, z_i)\}$ be defined as follows: $k_{iN}(z_i, \cdots, z_i) = k_i(z_1, \cdots, z_i)$ for $i < N$ and $k_{iN}(z_1, \cdots, z_i) = 1$ for $i \ge N$. Denote by γ_N and γ_N^* the values of γ and γ^*, respectively, if the sequence K is replaced by K_N. Then we have

$$(38)\qquad E(n \mid K) \ge E(n \mid K_N) \ge \frac{1}{h_0 Ez}\left[\gamma_N \log\frac{\gamma_N^*}{\gamma_N} + (1-\gamma_N)\log\frac{1-\gamma_N^*}{1-\gamma_N}\right]$$

and

$$(39)\qquad E^*(n \mid K) \ge E^*(n \mid K_N) \ge \frac{1}{h_0 Ez^*}\left[\gamma_N^* \log\frac{\gamma_N^*}{\gamma_N} + (1-\gamma_N^*)\log\frac{1-\gamma_N^*}{1-\gamma_N}\right].$$

Since $\lim_{N=\infty} \gamma_N = \gamma$ and $\lim_{N=\infty} \gamma_N^* = \gamma^*$, inequalities (30) and (31) follow from (38) and (39). Hence the proof of Theorem 3 is completed.

5. Remarks added in proof. The results obtained in the present paper have obvious applications to sequential analysis. These applications are, however, not mentioned here, because at the time the present paper was submitted for publication, sequential analysis constituted classified material. In the meantime, the material on sequential analysis has been released and was published in this Journal, June, 1945. The results obtained in the present paper are more general than those obtained in connection with sequential analysis. Theorem 3, in the present paper, implied the efficiency of the sequential probability ratio test discussed in Section 4.7 of the paper on sequential tests.

REFERENCE

[1] A. WALD, "On cumulative sums of random variables," *Annals of Math. Stat.*, Vol. 15, (1944), pp. 283–296.