

## ON EXCESS OVER THE BOUNDARY<sup>1</sup>

BY GARY LORDEN<sup>2</sup>

Northwestern University

**0. Summary.** A random walk,  $\{S_n\}_{n=0}^\infty$ , having positive drift and starting at the origin, is stopped the first time  $S_n > t \geq 0$ . The present paper studies the "excess,"  $S_n - t$ , when the walk is stopped. The main result is an upper bound on the mean of the excess, uniform in  $t$ . Through Wald's equation, this gives an upper bound on the mean stopping time, as well as upper bounds on the average sample numbers of sequential probability ratio tests. The same elementary approach yields simple upper bounds on the moments and tail probabilities of residual and spent waiting times of renewal processes.

**1. Introduction and main results.** In the study of cumulative sums,  $S_n = X_1 + \dots + X_n$ , of independent random variables with common mean  $m > 0$ , stopping times  $N(t) = \inf\{n: S_n > t\}$  are of central interest. For many applications in probability theory and statistics, e.g. to estimate sample sizes of sequential decision procedures, one is concerned with the first moment,  $EN(t)$ . Wald's equation (1946),  $mEN(t) = ES_{N(t)}$ , holds whenever  $\sup_n E|X_n|$  and  $EN(t)$  are finite, and can be rewritten  $mEN(t) = t + ER_t$ , where  $R_t = S_{N(t)} - t$ . The positive quantity  $R_t$  is called *excess over the boundary* and is often presumed to be negligible. In his fundamental treatise on sequential analysis (1947), Wald gave an upper bound for  $\sup_{t \geq 0} ER_t$  in the case of independent, identically distributed  $X$ 's: the evidently sufficient quantity  $\sup_{r \geq 0} E[X - r | X > r]$ . This bound is exact for the exponential distribution and is quite good in many cases (e.g. normal distributions) where the common distribution function,  $F$ , has monotone increasing *hazard rate*,  $(d/dx) \{-\log(1 - F(x))\}$ , for  $x \geq 0$ . In fact, for this class of distributions the supremum of  $E[X - r | X > r]$  is attained at  $r = 0$  and, since  $\sup_{t \geq 0} ER_t \geq ER_0 \geq (1 - F(0))E[X | X > 0]$ , evidently Wald's bound is too large by at most a factor of  $(1 - F(0))^{-1}$ , typically less than two. On the other hand, Wald's estimate has apparent deficiencies: it may be difficult to calculate; it is frequently much too large (for instance, when the distribution of  $X$  has large "gaps"); in fact, it may be infinite even if  $E(X^+)^2$  is finite, which is well known to be necessary and sufficient for the finiteness of  $\sup_{t \geq 0} ER_t$ .

The problem of finding other useful bounds on  $\sup_{t \geq 0} ER_t$  has proved troublesome because excess over the boundary is the sort of phenomenon which seems to behave in a pleasantly regular fashion only in the limit as  $t$  becomes large. An approach due to Farrell (1964) is of greatest interest in the case where  $E(X^+)^2 = \infty$ , since Farrell's bounds on  $ER_t$  necessarily tend to infinity as  $t$  becomes large. The powerful methods of renewal theory yield impressive asymptotic results in the case

Received February 16, 1968.

<sup>1</sup> Research sponsored in part by the National Science Foundation under grant NSF GP-4972.

<sup>2</sup> Now at California Institute of Technology.

of nonnegative  $X$ 's, where  $EN(t)$ , usually denoted by  $U(t)$ , is the expected number of renewal epochs in  $[0, t]$ , counting the zero epoch. An excellent exposition of these results is contained in Feller's book ((1966) pages 355–357). For nonnegative  $X$  with non-arithmetic distribution, a refinement of the renewal theorem yields  $ER_t \rightarrow \frac{1}{2}EX^2/m$  and  $P[R_t > u] \rightarrow E(X-u)^+/m$  as  $t \rightarrow \infty$ . The asymptotic expansion of  $EN(t)$  has been further refined under suitable assumptions, e.g. in (Stone, 1966).

For arithmetic distributions with span  $h$ ,  $\limsup_{t \rightarrow \infty} ER_t = \frac{1}{2}EX^2/m + \frac{1}{2}h$ , and  $h$  may be as large as  $EX^2/m$ . Considering  $t = 0$  and still assuming  $X$  is nonnegative, we find  $ER_0 = EX = EX^2/m - \text{Var } X/m$  and if  $\text{Var } X$  is small compared to  $m$  then  $\sup_{t \geq 0} ER_t$  is at best nearly  $EX^2/m$ .

Thus, in the nonnegative case, for  $t$  at both ends of  $[0, \infty)$  we may find  $ER_t$  as large as  $EX^2/m$ . The proof that this quantity is an upper bound on  $ER_t$  for all  $t$  is based upon the fact that, just as the asymptotic behavior of  $ER_t$  is fairly regular, so is its *average behavior* over intervals.

**THEOREM 1.** *Suppose  $X_1, X_2, \dots$  are independent, identically distributed random variables with  $EX = m > 0$  and  $E(X^+)^2 < \infty$ .*

*Let  $S_n = X_1 + \dots + X_n$ ,  $N(t) = \inf \{n: S_n > t\}$ , and  $R_t = S_{N(t)} - t$ . Then*

$$\sup_{t \geq 0} ER_t \leq E(X^+)^2/m.$$

**PROOF.** Whatever the values of  $X_1, X_2, \dots$ , the chance function  $\{R_t; t \geq 0\}$  is piecewise linear, all pieces having slope  $-1$ . We consider first the case where the  $X$ 's are nonnegative. It is easy to see that for  $c \geq 0$

$$(1) \quad \int_0^c R_t dt = \int_0^{S_{N(c)}} R_t dt - \int_c^{S_{N(c)}} R_t dt = \frac{1}{2} \sum_{n=1}^{N(c)} X_n^2 - \frac{1}{2} R_c^2.$$

Since  $R_t \geq 0$  for all  $t$ , we have by Fubini's theorem and Wald's equation

$$(2) \quad \int_0^c ER_t dt = \frac{1}{2} EX^2 EN(c) - \frac{1}{2} ER_c^2.$$

(Since  $EN(c)$  is finite under our hypotheses, the summation term in (1) has finite expectation by Wald's equation and since the other terms are nonnegative, they also have finite expectations.)

By Jensen's inequality and Wald's equation

$$(3) \quad \int_0^c ER_t dt \leq \frac{1}{2} m^{-1} EX^2 (c + ER_c) - \frac{1}{2} (ER_c)^2.$$

It is easy to see that for all  $t, u \geq 0$   $EN(t+u) \leq EN(t) + EN(u)$ , since the conditional expectation of  $N(t+u) - N(t)$  given  $N(t) = n, X_1 = x_1, \dots, X_n = x_n$  equals  $EN(u-r)$ , where  $r = x_1 + \dots + x_n - t > 0$  and  $N(u-r)$  is zero if  $r > u$ , so that  $EN(u-r) \leq EN(u)$  in any case. It follows from Wald's equation that  $ER_t$  is likewise a subadditive function of  $t$  and therefore

$$(4) \quad \frac{1}{2} c ER_c \leq \frac{1}{2} c \inf_{0 \leq t \leq \frac{1}{2}c} (ER_t + ER_{c-t}) \leq \int_0^{\frac{1}{2}c} (ER_t + ER_{c-t}) dt = \int_0^c ER_t dt.$$

Combining (3) and (4) and rewriting, we obtain

$$(5) \quad (ER_c)^2 + (c - EX^2/m) ER_c - c(EX^2/m) \leq 0.$$

The left-hand side of (5) is a quadratic in  $ER_c$  which is non-positive only between its roots,  $-c$  and  $EX^2/m$ . Therefore,  $ER_c \leq EX^2/m$ , and since  $c$  is arbitrary the proof is complete for the nonnegative case. The case where  $X_1, X_2, \dots$  may take negative values reduces to the nonnegative case through consideration of the associated sequence of positive ladder variables, which are independent and distributed like  $S_{N(0)}$  (Blackwell, (1953)). Since  $R_t$  is pointwise the same for  $X_1, X_2, \dots$  and the sequence of ladder variables, and clearly  $0 < S_{N(0)} \leq X_{N(0)}^+$ , the result for the nonnegative case implies

$$\sup_{t \geq 0} ER_t \leq \frac{ES_{N(0)}^2}{ES_{N(0)}} \leq \frac{E(X_{N(0)}^+)^2}{ES_{N(0)}} \leq \frac{E[(X_1^+)^2 + \dots + (X_{N(0)}^+)^2]}{E[X_1 + \dots + X_{N(0)}]} = \frac{E(X^+)^2}{m}$$

by Wald's equation, and the proof is complete.

If the  $X$ 's are nonnegative and are considered as waiting times of a renewal process, then Theorem 1 yields the following estimate of  $U\{I\}$ , the expected number of renewal epochs in an interval  $I$  of length  $h$ :

$$|U\{I\} - h/m| \leq EX^2/m^2.$$

The next result is useful for estimating the average sample numbers of a sequential probability ratio test.

**COROLLARY 1.** *Under the assumptions of Theorem 1, if  $a \leq 0 \leq b$  and  $N^* = \inf \{n: S_n \notin [a, b]\}$ , then*

$$EN^* \leq \frac{(1-\alpha)b + \alpha a}{m} + \frac{E(X^+)^2}{m^2},$$

where  $\alpha = P(S_{N^*} < a)$ .

**PROOF.** It is easy to verify that  $S_{N^*} \leq \min(S_{N^*}, b) + S_{N(b)} - b$ . Therefore,

$$ES_{N^*} \leq (1-\alpha)b + \alpha a + E(S_{N(b)} - b) \leq (1-\alpha)b + \alpha a + E(X^+)^2/m,$$

and Wald's equation yields the stated result.

In applying the corollary it is obviously desirable to have a lower bound on  $\alpha$  available. In the case of sequential probability ratio tests the standard argument (Wald (1947)) establishes only an upper bound. This same argument, however, demonstrates that

$$\frac{\beta}{1-\alpha} = E[\exp(-S_{N^*}) | S_{N^*} > b],$$

where  $\beta$  is the error probability when the alternative distribution is true (the upper bound being derived from the observation that the conditional expectation is at most  $\exp(-b)$ ). By applying the conditional version of Jensen's inequality and an argument like that of the corollary, we obtain

$$\frac{\beta}{1-\alpha} \geq \exp(-E[S_{N^*} | S_{N^*} > b]) \geq \exp\left(-\left\{b + \frac{E(X^+)^2}{(1-\alpha)m}\right\}\right).$$

Using the standard upper bound  $\alpha \leq \exp(a)$ ,

$$\beta \geq (1 - \exp(a)) \exp\left(-\left\{b + \frac{E(X^+)^2}{(1 - \exp(a))m}\right\}\right),$$

and a similar calculation using the alternative distribution yields a lower bound on  $\alpha$ .

**2. Generalizations and applications to renewal theory.** In trying to extend the ideas of Theorem 1 to cover variables which are not necessarily independent or identically distributed it is difficult to formulate general conditions under which  $EN(t)$  is subadditive. The following theorem gives weaker results but shows that at least the (uniform) average of  $ER_t$  over intervals  $[0, c]$  is well behaved under quite general circumstances.

**THEOREM 2.** *Suppose  $X_1, X_2, \dots$  are random variables on some probability space with  $N(c)$  and  $R_c$  defined as above. If  $E[(X_n^+)^2 | N(c) \geq n] \leq R E[X_n | N(c) \geq n]$  for  $n = 1, 2, \dots$  and  $ER_c$  or  $EN(c)$  is finite, then*

$$(6) \quad \begin{aligned} c^{-1} \int_0^c ER_t dt &\leq \frac{1}{2}R(1 + R/4c) && \text{if } c \geq \frac{1}{2}R \\ &\leq R - \frac{1}{2}c && \text{if } c < \frac{1}{2}R \end{aligned}$$

and  $ER_c \leq R + (Rc)^{\frac{1}{2}}$ .

**PROOF.** Relation (1) holds if  $X_n^2$  is replaced by  $((S_n - \max_{k < n} S_k)^+)^2$ , which is not larger than  $(X_n^+)^2$ . Therefore,

$$(7) \quad \int_0^c R_t dt \leq \frac{1}{2} \sum_{n=1}^{N(c)} (X_n^+)^2 - \frac{1}{2}R_c^2.$$

Define 
$$I_n = \begin{cases} 1 & \text{if } N(c) \geq n, \\ 0 & \text{if } N(c) < n, \end{cases} \quad \text{for } n = 1, 2, \dots$$

For  $m = 1, 2, \dots$ , our hypothesis implies

$$\sum_{n=1}^m E(X_n^+)^2 I_n \leq R \sum_{n=1}^m EX_n I_n = RE \sum_{n=1}^m X_n I_n \leq R(c + ER_c),$$

since 
$$\sum_{n=1}^m X_n I_n \leq \begin{cases} c & \text{if } N(c) > m; \\ c + R_c & \text{if } N(c) \leq m. \end{cases}$$

Letting  $m \rightarrow \infty$ ,

$$R(c + ER_c) \geq \sum_{n=1}^{\infty} E(X_n^+)^2 I_n = E \sum_{n=1}^{\infty} (X_n^+)^2 I_n = E \sum_{n=1}^{N(c)} (X_n^+)^2,$$

by monotone convergence.

Assume  $ER_c$  is finite. (It is shown below that this is the case whenever  $EN(c)$  is finite.) Then by (7) and the relation just stated,  $ER_c^2$  is finite and

$$(8) \quad \int_0^c ER_t dt = E \int_0^c R_t dt \leq \frac{1}{2}R(c + ER_c) - \frac{1}{2}ER_c^2 \leq \frac{1}{2}R(c + ER_c) - \frac{1}{2}(ER_c)^2.$$

The right-hand side of (8) is a quadratic in  $ER_c$  with maximum value  $\frac{1}{2}cR(1 + R/4c)$ , attained at  $ER_c = \frac{1}{2}R$ . This proves (6) for  $c \geq \frac{1}{2}R$ . Also, since the left-hand side of

(8) is nonnegative,  $R(c + ER_c) - (ER_c)^2 \geq 0$ , whence  $ER_c \leq \frac{1}{2}(R + (R^2 + 4Rc)^{\frac{1}{2}}) \leq R + (Rc)^{\frac{1}{2}}$ . Clearly  $t > u$  implies  $R_t \geq R_u - (t - u)$ . It follows by a simple computation that  $1/c \int_0^c ER_t dt \leq ER_c + \frac{1}{2}c$ . Now, if for some  $c < \frac{1}{2}R$  (6) does not hold, then evidently  $R - \frac{1}{2}c < c^{-1} \int_0^c ER_t dt \leq ER_c + \frac{1}{2}c$ , so that

$$(9) \quad ER_c > R - c.$$

The quadratic on the right-hand side of (8) is decreasing for  $ER_c > \frac{1}{2}R$  and  $R - c > \frac{1}{2}R$  since  $c < \frac{1}{2}R$ . Thus (8) and (9) imply

$$\int_0^c ER_t dt \leq \frac{1}{2}R^2 - \frac{1}{2}(R - c)^2 = Rc - \frac{1}{2}c^2,$$

contradicting the assumption that (6) does not hold.

It remains to verify that  $ER_c$  is finite whenever  $EN(c)$  is finite. First note that

$$\begin{aligned} E[X_n^+ | N(c) \geq n]^2 &\leq E[(X_n^+)^2 | N(c) \geq n] \leq RE[X_n | N(c) \geq n] \\ &\leq RE[X_n^+ | N(c) \geq n], \end{aligned}$$

so that  $E[X_n^+ | N(c) \geq n] \leq R$ . Then, if  $EN(c)$  is finite,

$$c + ER_c = E \sum_{n=1}^{\infty} X_n I_n \leq E \sum_{n=1}^{\infty} X_n^+ I_n \leq \sum_{n=1}^{\infty} EX_n^+ I_n \leq \sum_{n=1}^{\infty} REI_n,$$

and this last equals  $REN(c)$ , which is finite. Therefore,  $ER_c$  is finite and the proof is complete.

The next result generalizes Theorem 1 by giving upper bounds for arbitrary moments of  $R_t$ . These also provide alternative bounds on  $\sup_{t \geq 0} ER_t$  by virtue of the relation  $(ER_t)^p \leq E(R_t)^p$  valid for  $p > 1$ .

**THEOREM 3.** *Under the assumptions of Theorem 1,*

$$\sup_{t \geq 0} E(R_t)^p \leq \frac{p + 2}{p + 1} \frac{E(X^+)^{p+1}}{m} \quad \text{for all } p > 0.$$

**REMARK.** In the case of nonnegative variables with non-arithmetic distribution, the  $p$ th moment of the limiting distribution of  $R_t$  is  $E(X^+)^{p+1}/(p+1)m$ .

**PROOF.** By the same kind of argument used for Theorems 1 and 2 we obtain

$$(10) \quad \int_0^c ER_t^p dt \leq \frac{E(X^+)^{p+1}}{(p+1)m} (c + ER_c) - \frac{E(R_c)^{p+1}}{p+1}.$$

Letting  $L_t$  denote  $X_{N(t)}$ , we have by a similar computation

$$(11) \quad \int_0^c EL_t^p dt \leq \frac{E(X^+)^{p+1}}{m} (c + ER_c) - E(R_c)^{p+1}.$$

We now show that for all  $u, v \geq 0$

$$(12) \quad R_{u+v} \leq \max(R_u, L_v'),$$

where  $L_v'$  is distributed like  $L_v$ . Let  $N'$  be the smallest  $n > N(u)$  such that

$X_{N(u)+1} + \dots + X_n > v$  and set  $L_v' = X_{N'}$ . If  $R_{u+v} > R_u$ , then  $N(u+v) = N'$  and hence  $L_v' = X_{N(u+v)} \geq R_{u+v}$ . Therefore, (12) holds and it clearly follows that

$$cER_c^p \leq \int_0^c [ER_t^p + EL_{c-t}^p] dt.$$

Combining this last inequality with (10) and (11), and using obvious estimates, we obtain

$$cER_c^p - \frac{(p+2)E(X^+)^{p+1}}{(p+1)m} [c + (ER_c^p)^{1/p}] + \frac{(p+2)(ER_c^p)^{(p+1)/p}}{p+1} \leq 0.$$

The expression on the left-hand side is nonnegative if we set  $ER_c^p = (p+2)E(X^+)^{p+1}/(p+1)m$  and for this value of  $ER_c^p$  and all larger values the expression is seen by differentiation to be an increasing function of  $ER_c^p$ . Therefore, for all  $c \geq 0$   $ER_c^p$  is not larger than  $(p+2)E(X^+)^{p+1}/(p+1)m$ .

There is an alternative approach to the problem of bounding  $ER_t^p$  which uses the relation  $R_{u+v} \leq R_u + R_v'$ , where  $R_v'$  is distributed like  $R_v$ . This leads to the result that  $ER_c^p \leq 2^p E(X^+)^{p+1}/(p+1)m$ , which is sharper than the bound in Theorem 3 only for  $p < 2$ . The approach of Theorem 3 can be used to obtain upper bounds on other functionals of the distribution of  $R_t$ , such as its moment generating function.

In the case of nonnegative i.i.d. variables, one can apply Theorem 3 with  $p = 2$  and extend the argument of Theorem 2 to obtain (with  $R$  defined to be  $EX^2/m$ )

$$\left| \frac{1}{c} \int_A^{A+c} ER_t dt - \frac{1}{2}R \right| \leq \frac{1}{c} \left\{ \frac{2EX^3}{3m} + \frac{R^2}{8} \right\} \quad \text{for all } A, c \geq 0.$$

We now use relation (12) to derive upper bounds on the tail probabilities of  $R_t$ .

**THEOREM 4.** *If  $X_1, X_2, \dots$  are independent, identically distributed random variables with  $EX = m > 0$  and  $E(X^+)^2 \leq Rm < \infty$ , then for all  $t, z \geq 0$*

$$(13) \quad P[R_t \geq z] \leq \frac{1}{m} E[(2X - z)I\{X \geq z\}] \left( \frac{t+R}{t+z} \right),$$

where  $I\{X \geq z\}$  is the indicator function of the set where  $X \geq z$ .

**PROOF.** Applying (12) and using the same kind of argument as in the proofs of Theorems 1 and 3, we have

$$\begin{aligned} tP[R_t \geq z] &\leq \int_0^t P[R_u \geq z] du + \int_0^t P[L_{t-u} \geq z] du \\ &= EN(t)E(X-z)^+ - E(R_t-z)^+ + EN(t)E[XI\{X \geq z\}] \\ &\quad - E[R_t I\{L_t \geq z\}] \\ &\leq EN(t)E[(2X-z)I\{X \geq z\}] - E[R_t I\{R_t \geq z\}] \\ &\leq m^{-1}(t+R)E[(2X-z)I\{X \geq z\}] - zP[R_t \geq z], \end{aligned}$$

using Theorem 1 to estimate  $EN(t)$ . The stated result follows immediately.

A rough indication of the sharpness of (13) in the case of non-arithmetic nonnegative variables comes from the observation that  $E[(2X-z)I\{X \geq z\}]/m$  is smaller than  $E(2X-z)^+/m$ , which is two times  $\lim_{t \rightarrow \infty} P[R_t \geq \frac{1}{2}z]$ . Equality holds if  $X \equiv m$ ,  $z = m$  and  $t$  is an integer multiple of  $m$ .

In renewal theory,  $R_t$  is called the *residual waiting time* at epoch  $t$ , and in the nonnegative, non-arithmetic case it has the same limit distribution as  $Q_t = t - S_{N(t)-1}$ , which is called the *spent waiting time*. We now indicate how our results about  $R_t$  can be proved to hold in the nonnegative case for  $Q_t$  as well, with slight modifications. Using the same notation as in the derivation of (12), we note that

$$(14) \quad Q_v' \leq \max(R_u, L_{u+v}),$$

where  $Q_v' = v - [X_{N(u)+1} + \dots + X_{N'-1}]$  and is distributed like  $Q_v$ , by considering the cases  $N(u+v) < N'$  and  $N(u+v) \geq N'$ .

We have, therefore, for all  $b, z \geq 0$

$$\begin{aligned} bP[Q_v \geq z] &\leq \int_0^b P[R_u \geq z] du + \int_v^{b+v} P[L_u \geq z] du \\ &\leq EN(b+v)E[(2X-z)I\{X \geq z\}] \quad \text{for all } v \geq 0. \end{aligned}$$

Dividing by  $b$  and letting  $b \rightarrow \infty$ , we conclude that

$$(15) \quad P[Q_v \geq z] \leq m^{-1}E[(2X-z)I\{X \geq z\}] \quad \text{for all } v, z \geq 0,$$

which is similar to Theorem 4.

The restriction to nonnegative variables is necessary for (14). Unlike  $R_t$ ,  $Q_t$  may be larger for a (possibly negative) sequence  $X_1, X_2, \dots$ , than for the associated sequence of ladder variables. There is, therefore, no straightforward extension from the nonnegative case.

For nonnegative variables, Theorems 1 and 3 hold with  $Q_t$  in place of  $R_t$ . The proof of Theorem 1 uses the easily verified relation  $Q_v' \leq R_u + Q_{u+v}$ , which implies

$$bEQ_v \leq \int_0^b ER_u du + \int_v^{b+v} EQ_u du \leq \int_0^{b+v} EL_u du \leq EN(b+v)EX^2,$$

and the conclusion follows upon dividing by  $b$  and letting  $b \rightarrow \infty$ . Theorem 3 can be proved either by a similar argument using (14), or by a direct integration, using (15) to estimate tail probabilities. Since  $L_t = R_t + Q_t$ , we have for the nonnegative case  $\sup_{t \geq 0} EL_t \leq 2EX^2/m$  and it is straightforward to derive estimates for the moments and tail probabilities of  $L_t$ .

**Acknowledgment.** The author is grateful to Professor J. Sacks for helpful discussions in the course of this research.

## REFERENCES

- [1] BLACKWELL, D. (1953). Extension of a renewal theorem. *Pacific J. Math.* **3** 315–320.
- [2] FARRELL, R. H. (1964). Limit theorems for stopped random walks. *Ann. Math. Statist.* **35** 1332–1343.
- [3] FELLER, W. (1966). *An Introduction to Probability Theory and its Applications*, **2** Wiley, New York.
- [4] STONE, C. J. (1966). Absolutely continuous components and renewal theory. *Ann. Math. Statist.* **37** 271–275.
- [5] WALD, A. (1946). Differentiation under the expectation sign in the fundamental identity of sequential analysis. *Ann. Math. Statist.* **17** 493–497.
- [6] WALD, A. (1947). *Sequential Analysis*. Wiley, New York.