On Second Moments of Stopping Rules

by

Y.S. Chow* and H. Teicher*

Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series No. 59
October 1965

* Research under National Science Foundation Contract No. GP-4590.
On Second Moments of Stopping Rules

by

Y.S. Chow* and H. Teicher*

Purdue University

Summary. The current investigation is a natural outgrowth of [2], being concerned with the variance of stopping rules and the effect of non-zero means on the variance of a randomly stopped sum. Some martingale generalizations of applications of [2] also appear.

1. Introduction. A stopping rule or stopping variable of a sequence 
\{X_n, n \geq 1\} of random variables defined on a probability space \((\Omega, \mathcal{F}, P)\) is a positive integer-valued random variable \(t\) such that for every \(n \geq 1\) the event \(\{t = n\} \in \mathcal{F}_n\), the Borel field generated by \(X_1, \ldots, X_n\). In contradistinction, a stopping time \(t\) (likewise of a sequence \(\{X_n\}\)) will be defined as a positive integer or \(+\infty\) valued function on \(\Omega\) subject to the same proviso that \(\{t = n\} \in \mathcal{F}_n, n \geq 1\). Thus, a stopping time \(t\) is a stopping variable or stopping rule if and only if \(P\{t < \infty\} = 1\). In numerous problems of probability theory and statistics it is necessary to demonstrate that what is obviously a stopping time is further a stopping variable and even to obtain detailed information about the latter.

2. Comparison of Stopping Rules. Let the basic process \(\{X_n, n \geq 1\}\) consist of independent random variables with \(E X_n = 0, E X_n^2 = 1, P\{X_n | a < \infty\} = 1\) for \(n \geq 1\). Set \(S_n = \sum_{i=1}^{n} X_i\) and define \(t_m(c)\) to be the smallest positive index \(n \geq m (m = 1, 2, \ldots)\) for which \(S_n^2 > c^2 n\) where \(c\) is a positive
constant. For the case of coin tossing \((a = 1)\), it was shown in \([1]\) that for all \(m\), \(E t_m(c)\) is finite or infinite according as \(c < 1\) or \(c \geq 1\) and this was generalized in \([2]\) to the uniformly bounded case.\(^2\) Apropos of these results it may be noted that for \(m = 0\), the lemma of theorem 1 of the next section gives an upper bound for \(E t_1(c)\) when the latter is finite. It will be proved in section 3 that if \(c^2 < 3 - \sqrt{6}\), \(E t_m^2(c) < \infty\), all \(m \geq 1\) while if \(c^2 > 3 - \sqrt{6}\) then \(E t_m^2(c) = \infty\) for all sufficiently large (but not necessarily all) \(m\).

It is clear from a comparison technique that there is a non-increasing sequence of non-negative constants \(\{c_k, k \geq 1\}\) such that \(E t_m^k(c) < \infty\) for \(c < c_k\) (if \(c_k > 0\)) while \(E t_m^k(c) = \infty\) for all sufficiently large \(m\) if \(c > c_k\). Such comparisons may be formalized by the following.

Definition: A stopping time \(t\) will be called "more restrictive" than a stopping time \(s\) if \(\{t = n\} \subset \{s \leq n\}\) for \(n = 1, 2, \ldots\) that is, if \(t \geq s\).

Clearly, if \(t\) is more restrictive than \(s\), and \(t\) is a bonafide stopping variable, so is \(s\); moreover, the finiteness of \(E t^\alpha\) implies that of \(E s^\alpha\) for any \(\alpha > 0\).

Thus, if \(c < c'\), \(E t_m^k(c) \leq E t_m^k(c')\), \((k, m = 1, 2, \ldots\) corroborating the prior statement about the sequence \(c_k\). It is a natural conjecture that \(c_k\) decreases to zero but currently the authors know of no method for attacking this seemingly simple question.

3. **Second Moments.** When \(c^2 < 1\), the situation changes in the coin tossing example \((a = 1)\) alluded to earlier since now \(P[t_m(c) = 1] = 1\) for \(m = 1\). Thus, to allow the second moment to attain an infinite value, it is necessary to dawdle for a while so as to insure that \(S_n\) does not prematurely escape
its parabolic bonds. This accounts for the appearance of the phrase "for all sufficiently large \( m \)" in

Theorem 1: Let \( \{X_n\} \) be independent random variables with \( P[|X_n| \leq a < \infty] = 1 \), \( E X_n = 0 \), \( E X_n^2 = 1 \) for \( n \geq 1 \) and define \( t_m \) smallest integer \( n \geq m \) for which \( s_n^2 > c_n^2 \) (\( n = 1, 2, \ldots \)). If \( c^2 < 3 - \sqrt{5} \), then \( E t_m^2 < \infty \), all \( m \geq 1 \) while if \( c^2 > 3 - \sqrt{5} \), \( E t_m^2 = \infty \) for all sufficiently large \( m \).

Proof: In the case \( c^2 < 3 - \sqrt{5} \) we write \( t \) for \( t_m \). Set \( \gamma_n = E X_n^3 \), \( \beta_n = E X_n^4 \) and \( t' = \min(t, k) \) where \( k > m \). Since \( E t' \sum_{j=1}^{t'} \beta_j < a^4 E t' t'^2 < \infty \), by Theorem 3 or [2],

\[
E s_{t'}^2 = 6 E t' s_{t'}^2 - 3 E t' (t'+1) + 4 E s_{t'}^2 + E \sum_{j=1}^{t'} \gamma_j
\]

whence

\[
E(s_{t'}^2 - c_{2t'}^2) = (6 - 2c^2) E t' s_{t'}^2 - (3 - c^4) E t' t'^2 - 3 E t' + 4 E s_{t'}^2 \sum_{j=1}^{t'} \gamma_j
\]

\[
+ E \sum_{j=1}^{t'} \beta_j
\]

implying

\[
(3 - c^4) E t' t'^2 + (2c^2 - 6) E t' s_{t'}^2 \leq (a^4 - 3) E t' + 4a^3 E t' | s_{t'} | .
\]

Let \( A_k = \{ m < t \leq k \} \). From (2), recalling that \( E t' \leq E t < \infty \) for \( c^2 < 1 \) [2],
\[(3 - c^4)[\int_{t>k} k^2 + \int_{t>k} t^2] + (2c^2 - 6)\left[\int_{t>k} c^2k^2 + \int_{t>k} t(ct^{1/2} + a)^2\right]\]
\[\leq 4a^3\left[\int_{t>k} ck^{3/2} + \int_{t>k} t(ct^{1/2} + a)\right] + o(1).
\]

Consequently,
\[(c^4 - 6c^2 + 3)\left[\int_{t>k} k^2P[t > k] + \int_{t>k} t^2\right] \leq B\left[\int_{t>k} k^{3/2}P[t > k] + \int_{t>k} t^{3/2}\right] + o(1),\]

where \(B > 0\) is a constant depending only on \(c\) and \(a\). Thus, letting \(k \to \infty\), \(E t^2 < \infty\) regardless of \(m\).

In the alternative case, we may clearly suppose \(3 - \sqrt{5} < c^2 < 1\). Define \(u_m(c)\) to be the first index \(n \geq 1\) for which \(s_n^2 > c^2(n + m)\) where \(m\) is an arbitrary non-negative quantity.

Suppose it has been established for every \(c^2\) in \((3 - \sqrt{5}, 1)\) that \(E u^2_m(c) = \infty\) for all sufficiently large \(m\). Then, for any \(c^2\) in \((3 - \sqrt{5}, 1)\) we may choose \(c^2_o\) likewise in this interval but less than \(c^2\) and be assured of the existence of an integer \(m_o\) such that \(E u^2_m(c_o) = \infty\). Select the integer \(m_1\) so that \(c^2_n > c^2_o(n + m_o)\) for all \(n \geq m_1\). Then by the comparison technique \(E t^2_m = \infty\) for \(m \geq m_1\).

Thus, it suffices to prove the auxiliary proposition involving \(u_m(c)\) and in so doing we denote the latter variable by \(t\).

**Lemma:** For \(0 < c < 1\) and \(m \geq 0\),
\[
\frac{c^2 m}{1-c^2} \leq E t \leq \left[ac(1-c^2)^{-1} + \sqrt{(m-1)(1-c^2)^{-1} + a^2(1-c^2)^{-2}}\right]^2 - m + 1
\]
and thus \( E t = O(m) \).

**Proof:** Choose \( c < c_1 < 1 \) and \( m_1 > 0 \) such that \( c_1^2 n \geq c^2 (n+m) \) for all \( n \geq m_1 \). By the comparison technique and Corollary 2 of [2], \( E t < \infty \). By Theorem 2 of [2], \( E t = E S_t^2 \geq c^2 E(t+m) \) proving the first inequality. On the other hand,

\[
E t = E S_t^2 \leq E \left[ c(t+m-1)^{1/2} + a \right]^2 \leq c^2 E(t+m-1) + 2ac E^{1/2}(t+m-1) + a^2
\]

or

\[
(1-c^2) E(t+m-1) - 2ac E^{1/2}(t+m-1) - (a^2+m-1) \leq 0
\]

yielding the second.

Suppose now that \( E t^2 < \infty \) for all \( m \). By Theorem 3 of [2],

\[
E S_t^4 = 6 E t S_t^2 - 3 E t(t+1) + 4 E S_t \sum_{j=1}^t \gamma_j + E \sum_{j=1}^t s_j
\]

\[
\geq 6 c^2 E t(t+m) - 3 E t(t+1) - 4 a^3 E t |S_t|
\]

(3) \[ \geq (6c^2-3) E t^2 + (6mc^2-3) E t - 4 a^3 c E(t+m-1)^{3/2} - 4 a^4 E t \]

On the other hand,

(4) \[ E S_t^4 \leq E \left[ c(t+m-1)^{1/2} + a \right]^4 = c^4 E(t+m-1)^2 + 4 a c^3 E(t+m-1)^{3/2} \\
+ 6c^2 a^2 E(t+m-1) + 4 ca^3 E(t+m-1)^{1/2} + a^4 \]

whence, combining (3) and (4) and recalling that \( E t = O(m) \).
(6c^2 - 3 - c^4) E t^2 ≤ m^2 c^4 - 2mc^2(3-c^2) E t + 4ac(a^2+c^2) E(t+m-1)^{3/2} + O(m).

Since \( E(t+m-1)^{3/2} \leq 2 E t^{3/2} + 2 m^{3/2} \leq 2 E^{3/4} t^2 + 2 m^{3/2} \) and \( E t > m c^2(1-c^2)^{-1} \) (by the lemma),

(5) \((6c^2 - 3 - c^4) E t^2 \leq m^2 c^4 [1-2(3-c^2)(1-c^2)^{-1}] + 8ac(a^2+c^2)(E^{3/4} t^2 + m^{3/2}) + O(m)

Employing the lemma again, we have \( E t^2 \geq E t \geq m^2 c^4 (1-c^2)^{-2} \rightarrow \infty \) and

(6) \[ 6c^2 - 3 - c^4 \leq O(E^{-1/4} t^2) + O(m^{-1/2}) \].

Hence \( 6c^2 - 3 - c^4 \leq 0 \) which is patently false for \( c^2 \) in \((3-\sqrt{6}, 1)\).
Thus, \( E t^2 = \infty \) for all sufficiently large \( m \) and the theorem is proved.

Theorem 2: Let \( \{X_n\} \) be independent random variables with \( P[|X_n| \leq a < \infty] = E X_n = 0, E X_n^2 = 1 \) for \( n \geq 1 \). If \( t \) designates the smallest integer \( n \geq m \) such that \( |S_n| > c n^{1/\alpha} \), then \( E t^2 < \infty \) for all \( \alpha > 2, c > 0 \) and \( m \geq 1 \).

Proof: For any \( c > 0 \) and \( \alpha > 2 \), if \( m \) is sufficiently large \( c n^{1/\alpha} < 4^{-1} n^{1/2} \) for \( n \geq m \). It follows therefore from the comparison technique and Theorem 1 that \( E t^2 < \infty \) for all sufficiently large \( m \). Consequently, \( E t^2 < \infty \) for all \( m \geq 1, \alpha > 2, c > 0 \).

4. Non-Zero Means. Let the random variables \( \{X_n\} \) of the basic process be independent with \( E X_n = \mu_n, E X_n^2 = 1 + \mu_n^2, n \geq 1 \). If \( S_n = \sum_{i=1}^{n} X_i \) and \( t \) is a stopping variable with \( E t < \infty \), then
\[ E \left( S_t - \sum_{i=1}^{t} \mu_i \right)^2 = E t \]

by Theorem 2 of [2]. If, in addition \( \mu_n = 0 \), \( E S_t = 0 \) by Wald's theorem and the L.H.S. of (7) is just the variance of \( S_t \), say \( \sigma_{S_t}^2 \). On the other hand if \( \mu_n \neq 0 \), this is no longer the case and \( \sigma_{S_t}^2 \) may even be infinite despite the finiteness of (7).

For example, let \( P[X_n = \mu + 1] = P[X_n = \mu - 1] = \frac{1}{2}, \mu \neq 0 \) and define \( t \) as the first index \( n \geq m \) such that \( (S_n - n\mu)^2 > 3n/4 \). According to Theorem 1 of the preceding section, \( E t^2 = \infty \) for all \( m \geq m' \) (and it will now be stipulated that \( m \geq m' \)) while according to (7), \( E(S_t - t\mu)^2 < \infty \). In view of the elementary inequality \( \mu^2 E t^2 \leq 2E(S_t - t\mu)^2 + 2E S_t^2 \), it follows that \( E S_t^2 = \infty \). By Wald's theorem, \( E S_t = \mu E t < \infty \) and thus \( \sigma_{S_t}^2 = \infty \).

Even when both quantities are finite, no general inequality between

\[ E(S_t - \sum_{i=1}^{t} \mu_i)^2 \quad \text{and} \quad \sigma_{S_t}^2 \]

obtains. It is not difficult to verify that

\[ \text{Cov} \left( 2S_t - \sum_{i=1}^{t} \mu_i, \sum_{i=1}^{t} \mu_i \right) \leq 0 \quad \text{is necessary and sufficient for} \quad \sigma_{S_t}^2 \leq E(S_t - \sum_{i=1}^{t} \mu_i)^2 \]

if \( E(\sum_{i=1}^{t} \mu_i)^2 < \infty \), \( E \sum_{i=1}^{t} E|X_i| < \infty \). When \( E X_n = \mu, E X_n^2 = 1 + \mu^2 \) and \( t \) is a stopping variable with \( E t^2 < \infty \), the simple condition \( \mu \text{ Cov} (t, S_t) \leq 0 \) implies \( \sigma_{S_t}^2 \leq E(S_t - t\mu)^2 \). If \( P[X_n = 1] = p = 1 - P[X_n = -b], b > 0 \) and \( t \) denotes the first \( n \geq 1 \) for which \( X_n = 1 \), then \( S_t = -b(t-1) + 1 \). Since \( t \) and \( S_t \) are negatively correlated and \( E t^2 < \infty \), \( \sigma_{S_t}^2 \leq E(S_t - t\mu)^2 \) if \( \mu \geq 0 \), i.e., if \( p \geq b/(b+1) \). Here, this condition is necessary as well.
5. **Martingale Generalizations.** In the following, the basic process \( \{X_n\} \) will be postulated to satisfy \( E|X_n| < \infty, E[X_{n+1} \mid \mathcal{F}_n] = 0, n \geq 1 \) so that
\[
S_n = \sum_{i=1}^{n} X_i
\]
is a martingale.

**Theorem 3:** Let \( \{S_n, n \geq 1\} \) satisfy \( E[X_{n+1} \mid \mathcal{F}_n] = 0, E \sup X_n^2 < \infty \). If \( u_n^2 = E[X_n^2 \mid \mathcal{F}_{n-1}] \), define \( t \) as the first integer \( n \geq m \) for which
\[
S_n^2 > c^2 \sum_{j=1}^{n} u_j^2
\]
where \( 0 < c < 1 \) and \( m = 1, 2, \ldots \). Then
\[
\int_{[t\leq m]} \sum_{j=1}^{n} u_j^2 = o(1) \quad \text{as} \quad n \to \infty.
\]

**Proof:** For any integer \( k \geq m \), set \( t' = \min(t, k) \) and define
\[
z = \sup |X_n|, \quad A_k = \{m < t \leq k\}.
\]
By Theorem 1 of [2]

\[
\int_{[t\leq k]} \sum_{j=1}^{t} u_j^2 + \int_{[t\leq m]} \sum_{j=1}^{t} u_j^2 = E \sum_{j=1}^{t} u_j^2 = E S_t^2 \leq \int_{[A_k]} \sum_{j=1}^{t} u_j^2 + z^2 + \int_{[t\leq k]} c^2 \sum_{j=1}^{t} u_j^2 + o(1).
\]

Thus,
\[
(1-c^2) \int_{[t\leq k]} \sum_{j=1}^{t} u_j^2 + \int_{[t\leq k]} \sum_{j=1}^{t} u_j^2 \leq 2c(\int_{[A_k]} \sum_{j=1}^{t} u_j^2)^{1/2} \left( \int_{[A_k]} \sum_{j=1}^{t} u_j^2 \right)^{1/2} + o(1)
\]

and the conclusion follows.

**Corollary 1:** If further, \( P[ \sum_{j=1}^{t} u_j^2 = \infty ] = 1, P[ t > k ] = o(1) \) and \( E \sum_{j=1}^{t} u_j^2 < \infty \).

**Corollary 2:** If moreover \( P[ u_j^2 > \delta > 0 ] = 1, j \geq 1 \) then \( E t < \infty \).
Corollary 3: If \( \{X_n\} \) are independent with \( \mathbb{E}X_n = 0, \mathbb{E}X_n^2 = \sigma_n^2, \mathbb{E}(\sup_x X_n^2) < \infty, \sum \sigma_n^2 = \infty \) and \( t = \text{1st } \max_{m} \text{ such that } S_n^2 > c^2 \sum_{j=1}^{n} \sigma_j^2, 0 < c < 1, \) then
\[
P(t < \infty) = 1 \text{ and } \mathbb{E}(\sum_{j=1}^{t} \sigma_j^2) < \infty. \text{ If } \sigma_n^2 > \delta > 0, \mathbb{E}t < \infty.
\]

Corollary 3 generalizes corollary 2 of Theorem 2 of [2] wherein \( \sigma_n^2 = 1, n \geq 1. \)

Finally, the method of stopping rules will be utilized to generalize a Kolmogoroff inequality and to derive a result of Doob's [3, p.320] which does not follow from this inequality.

Theorem 4: Let \( \{X_n, n \geq 1\} \) satisfy \( \mathbb{E}X_n^2 < \infty, \mathbb{E}[X_{n+1} | \mathcal{F}_n] = 0 \) and set \( u_n^2 = \mathbb{E}[X_{n+1}^2 | \mathcal{F}_n], z = \sup_{n} |X_n| \). Then, if \( \epsilon > 0 \) for any positive integer \( k, \)
\[
\mathbb{E}(\epsilon+z)^2 \leq \mathbb{E}(\epsilon^2) \leq \epsilon^2 \sum_{j}^{k} u_j^2 \leq \mathbb{E}(\epsilon + z)^2 \leq \epsilon^2 \sum_{j}^{k} u_j^2 \leq \mathbb{E}(\epsilon + z)^2 \leq \epsilon^2 \sum_{j}^{k} u_j^2.
\]

Proof: Let \( t = \text{first } n \geq 1 \text{ such that } S_n^2 > \epsilon^2. \) Set \( t' = \min(t, k). \) Then
\[
\mathbb{E}(\epsilon + z)^2 \geq \mathbb{E}S_{t'}^2 = \mathbb{E} \sum_{j=1}^{t'} u_j^2 \geq \int_{[t > k]} \mathbb{E} \sum_{j=1}^{k} u_j^2 \geq \int_{\max_{n < k} S_n^2 \leq \epsilon^2} \mathbb{E} \sum_{j=1}^{k} u_j^2.
\]

Corollary 1: If moreover \( \mathbb{E}z^2 < \infty, S_n \) diverges a.e. on \( A = \left[ \sum_{j}^{\infty} u_j^2 = \infty \right]. \)

Proof: Let \( t = \text{1st } n \geq m \text{ for which } S_n^2 > \epsilon^2. \) Then for \( k > m \) it follows from the theorem that
\[ E(e + z)^2 \geq \int_{[t \geq k]} \sum_{m}^{k} u_j^2 = \int_{A[t \geq k]} \sum_{m}^{k} u_j^2 \geq \int_{A[t = \infty]} \sum_{m}^{k} u_j^2 \]

whence \( P[A[t = \infty]] = 0 \), i.e., \( \sup_{n \geq m} |S_n - S_{n-1}| > \epsilon \), a.e. in \( A \). Since \( m \) is arbitrary \( S_n \) diverges a.e. in \( A \).

Corollary 2: If, further \( Ez^2 < \infty \) and \( P[\sum_{l=1}^{\infty} u_n^2 = \infty] = 1 \), \( t \) is a bona fide stopping variable.
Footnotes

1. In [2] the terms are used synonymously but it is clearly desirable to make such a distinction.

2. For \( c \geq 1 \), the hypothesis of a uniform bound is superfluous and was not stipulated in [2].

References

