Statistica Sinica **10**(2000), 629-638

ASYMPTOTIC EXPANSIONS FOR THE MOMENTS OF A RANDOMLY STOPPED AVERAGE: EXTENSION AND APPLICATIONS OF A RESULT OF ARAS AND WOODROOFE

Wei Liu and Nan Wang

University of Southampton

Abstract: Aras and Woodroofe (1993) provide asymptotic expansions of the first four moments of $\bar{\mathbf{X}}_t := \mathbf{S}_t/t$ where $t = t_a = \inf\{n \geq 1 : \ Z_n > a\}$, $Z_n = n + \langle \mathbf{c}, \mathbf{S}_n\rangle + \xi_n$, $n = 1, 2, \ldots$. Here $\{S_n\}$ is a driftless random walk in an inner product space $\mathcal{W}$, $\mathbf{c} \in \mathcal{W}$, and $\xi_1, \xi_2, \ldots$ are slowly changing. The first part of this paper supplies similar expansions for stopping time $T = T_a = \inf\{n \geq m : \ Z_n > a\}$ where $m = m_a$ is a random variable. Stopping times of this form arise naturally from the sequential sampling scheme of Liu (1997). The general result is illustrated by an example. The second part of this paper applies Aras and Woodroofe's (1993) result directly to extend Woodroofe's (1977) result on second order expansion of risk from the normal distribution to the bounded density case. Let $Y_1, Y_2, \ldots$ be independent observations from a population with mean μ and variance $\sigma^2 > 0$. The basic problem is to estimate μ by the sample mean $\bar{Y}_n$ given a sample of size n, subject to the loss function $L_n = A\sigma^{2\beta-2}(\bar{Y}_n - \mu)^2 + n, \ A > 0, \beta > 0$. If σ is known, the fixed sample size n that minimizes the risk is given by $n_0 \approx A^{1/2}\sigma^{\beta}$, with the corresponding minimum risk R_{n_0}. However, when σ is unknown, there is no fixed sample size rule that will achieve the risk R_{n_0}. For this case the stopping rule $T = \inf\{n \geq m : \ n > A^{1/2}\hat{\sigma}_n^{\beta}\}$ can be used, and the population mean μ is then estimated by $\bar{Y}_T$. Martinsek (1983) obtained the second order expansion of the risk of this sequential estimation procedure, assuming the initial sample size $m \to \infty$ at a certain rate (but without specifying the form of distribution). If the initial sample size m is assumed to be prefixed, the second order expansion of the risk has been established by Woodroofe (1977) but only for normally distributed Y_i. The present paper provides the second order expansion of the risk under assumptions that m is prefixed and that the Y_i is continuous with a bounded probability density function.

Key words and phrases: Nonlinear renewal theory, risk functions, sequential estimation, stopping times, uniform integrability.

1. Introduction

Let $\mathcal{W}$ denote a finite-dimensional inner product space, with inner product and norm denoted by $\langle \cdot, \cdot\rangle$ and $||\cdot||$; and let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ denote i.i.d., $\mathcal{W}$-valued random vectors with common distribution F. Suppose that F has mean $\mathbf{0}$, covariance operator Σ and high moments as needed. Let $\xi_1, \xi_2, \ldots$ be random variables for which ξ_n is independent of $\mathbf{X}_{n+1}, \mathbf{X}_{n+2}, \ldots$ for all $n = 1, 2, \ldots$. Let $\mathbf{c} \in \mathcal{W}$, and let

$$Z_n = n + \langle \mathbf{c}, \mathbf{S}_n \rangle + \xi_n, \quad n \geq 1,$$
$$t = t_a = \inf\{n \geq 1 : \ Z_n > a\}, \quad a \geq 1,$$

where $\mathbf{S}_n = \mathbf{X}_1 + \cdots + \mathbf{X}_n$ for $n \geq 1$. Aras and Woodroofe (AW (1993)) provide asymptotic expansions as $a \to \infty$ for the first four moments of $\bar{\mathbf{X}}_t := \mathbf{S}_t / t$ and the first two moments of a smooth, suitably bounded function of $\bar{\mathbf{X}}_t$.

The purpose of this paper is twofold. The first is to provide similar asymptotic expansions when the stopping time t is replaced by

$$T = T_a = \inf\{n \geq m : \ Z_n > a\}, \quad a \geq 1,$$

where $m = m_a$ is a random variable satisfying some conditions to be specified. While stopping times of the form t arise naturally from the pure sequential sampling scheme of Anscombe (1953), Robbins (1959) and Chow and Robbins (1965), stopping times of the form T arise from an improved sequential sampling scheme proposed recently by Liu (1997). As an illustration, the general result is applied to the sampling scheme of Liu (1997) for the problem of sequential point estimation. This is contained in Section 2. We use the notation of AW (1993) there and it might be read in conjunction with that work.

The second purpose of this paper is to apply AW's (1993) result directly to extend Woodroofe's (1977) result on second order expansion of risk from the normal distribution to the bounded density case. Let $Y_1, Y_2, \ldots$ be independent observations from a population with mean μ and variance $\sigma^2 > 0$. Given a sample of size n, one wishes to estimate μ by the sample mean $\bar{Y}_n$, subject to the loss function $L_n = A\sigma^{2\beta-2}(\bar{Y}_n - \mu)^2 + n$ for $A > 0$ and $\beta > 0$. For a fixed sample size n, the risk is $R_n = A\sigma^{2\beta} n^{-1} + n$ and is minimized (when σ is known) by using the optimal fixed sample size $n_0 \approx A^{1/2}\sigma^{\beta}$, with the corresponding minimum risk $R_{n_0} = 2A^{1/2}\sigma^{\beta}$. When σ is unknown, the optimal fixed sample size n_0 cannot be used, and there is no fixed sample size rule that will achieve the risk R_{n_0}. For this case the stopping rule

$$T_R = \inf\{n \geq m : \ n > A^{1/2}\hat{\sigma}_n^{\beta}\}, \tag{1.1}$$

where m is the initial sample size and $\hat{\sigma}_n^2 = \sum_{i=1}^{n}(Y_i - \bar{Y}_n)^2/n$, can be used, and the population mean μ is then estimated by $\bar{Y}_{T_R}$. This type of sequential procedure was first proposed by Robbins (1959) in the normal case.

For the general distribution-free case Ghosh and Mukhopadhyay (1979) and Chow and Yu (1981) proved the asymptotic risk efficiency (i.e., $R_{T_R}/R_{n_0} \to 1$ as $A \to \infty$) of the sequential procedure above under some moment assumptions on

Y_1 and when $m \to \infty$ at a certain rate. Chow and Martinsek (1982) proved the stronger result that $R_{T_R} - R_{n_0} = O(1)$ as $A \to \infty$ under similar assumptions. The more elegant second order expansion of R_{T_R} as $A \to \infty$ has been established by Martinsek (1983) under the assumptions:

$$\begin{aligned}&E|Y_1|^{8r} < \infty \text{ for some } r > 1,\\&3 - (Y_1 - \mu)^2/\sigma^2 \text{ is nonlattice,}\\&\text{and } \ \delta A^{1/4} \le m = o(A^{1/2}) \text{ as } A \to \infty \text{ for some } \delta > 0.\end{aligned}$$

AW (1993) point out that the moment assumption above can be relaxed to $E|Y_1|^6 < \infty$. Note, however, all the results above assume that the initial sample size m depends on A and $m = m_A \to \infty$ as $A \to \infty$.

When the initial sample size m is fixed, independent of A, the second order expansion of R_{T_R} has been established by Woodroofe (1977), but only for normally distributed Y_i. Section 3 of this paper establishes the second order expansion of the risk when m is fixed and the Y_i are continuous with a bounded density function, but without specifying the form of the distribution. This result can therefore be viewed as an extension of Woodroofe's (1977) result from the normal distribution to the bounded density case. It is noteworthy that when m is fixed and the Y_i are discrete, then this sequential procedure may not be risk efficient, as demonstrated by Chow and Yu (1981). For a specific two-parameter exponential family of distributions and for a fixed initial sample size m, Bose and Boukai (1993) obtained second order expansion of risk using a stopping rule which differs from (1.1) and only makes sense in this special case.

2. An Extention of a Result of AW

The reader is reminded that the notation of AW (1993) is used throughout this section.

2.1. The extension

The following conditions are needed: for some $3 \le p < \infty$, $0 < \epsilon_1 < 1$ and $0 < \epsilon_2 < \epsilon_0 < 1 < \epsilon_3 < \infty$,

(C1) $E(\mathbf{X}_1) = \mathbf{0}$, $\int_{\mathcal{W}} ||\mathbf{x}||^2 F(d\mathbf{x}) < \infty$ and $\upsilon_p(\mathbf{c}) < \infty$;

(C2) $\left[\left(Z_n - \frac{n}{\epsilon_0}\right)^+\right]^{p+1}$, $n \ge 1$, are uniformly integrable;

(C3) $\sum_{n=1}^{\infty} nP\{\xi_n < -\epsilon_1 n\} < \infty$;

(C4) $\lim_{\delta\to 0}\sup_{n\ge 1} P\{\max_{k\le n\delta} |\xi_{n+k} - \xi_n| > \epsilon\} = 0$, $\forall\ 0 < \epsilon < \infty$;

(C5) there are events A_n, $n = 1, 2, \ldots$, and a $3/2 \le \alpha < \infty$ such that

$\sum_{n=1}^{\infty} nP\left(\cup_{k=n}^{\infty} A'_k\right) < \infty$ and $\max_{k\le n} |\xi_{n+k} I_{A_{n+k}}|^\alpha$, $n \ge 1$, are uniformly integrable;

(C6) $(\mathbf{S}_n^*, \xi_n) \Rightarrow (\mathbf{W}, \xi)$ as $n \to \infty$;

(C7) $\int_{m > \epsilon_3 a} m \, dP \to 0$ as $a \to \infty$;

(C8) $a^p P\{m > \epsilon_2 a\} \to 0$ as $a \to \infty$.

Conditions (C1) and (C3)-(C6) are the same as those of AW (1993) while condition (C2) is slightly stronger, so Theorems 1-4 of AW (1993) still hold under (C1)-(C6) here. Conditions (C7) and (C8) are on the random variable $m = m_a$; if $m = 1$ then (C7) and (C8) are clearly true. The main result of this section is

Theorem 1. *AW's* (1993) *Theorems* 1-4 of AW (1993) *still hold if* t *is replaced by* T *and their* (C1)-(C6) are *replaced by* (C1)-(C8) *above.*

Proof. First, note that

$$a^p P\{t < m\} \to 0 \quad \text{as } a \to \infty. \tag{2.1}$$

This can be seen from $\{t < m\} \subset \{t < \epsilon_2 a\} \cup \{m > \epsilon_2 a\}$, $a^p P\{t < \epsilon_2 a\} \to 0$ as $a \to \infty$ by (C2) and AW's Lemma 1, and (C8). Next, we show that

$$\int_{T > \epsilon a} T \, dP \to 0 \quad \text{as } a \to \infty \quad \text{for some } 1 < \epsilon < \infty. \tag{2.2}$$

Choose $0 < \epsilon_4 < 1$ so that $\epsilon_4 + \epsilon_1 < 1$ and $K_a := [a/(1 - \epsilon_1 - \epsilon_4)] + 1 \ge \epsilon_3 a$. Then, for $n > K_a$, $a - n < -n(\epsilon_1 + \epsilon_4)$ and

$$\begin{aligned} P\{T > n\} &\le P\{\langle \mathbf{c}, \mathbf{S}_n \rangle + \xi_n \le a - n\} + P\{m > n\} \\ &\le P\{\langle \mathbf{c}, \mathbf{S}_n \rangle < -n\epsilon_4\} + P\{\xi_n < -n\epsilon_1\} + P\{m > n\}. \end{aligned}$$

This, together with the inequality of Baum and Katz (1965, Theorem 3), (C3), (C7) and the integral by parts formula, implies

$$\sum_{n > K_a} P\{T > n\} \to 0 \quad \text{as } a \to \infty,$$

which in turn implies (2.2) by the integral by parts formula.

Now we are in the position to prove the theorem. The result corresponding to AW's Theorem 1 follows from

$$\begin{aligned} 0 \le E(T - t) &= \int_{t < m} (T - t) dP \le \int_{t < m} T \, dP \\ &\le \int_{T > \epsilon a} T \, dP + \epsilon a P\{t < m\} \to 0 \quad \text{as } a \to \infty \end{aligned}$$

by (2.1) and (2.2). The result corresponding to AW's Theorem 2 (9) can be seen from

$$\begin{aligned}
&a^2 \left| E\langle \mathbf{b}, \bar{\mathbf{X}}_T\rangle^2 - E\langle \mathbf{b}, \bar{\mathbf{X}}_t\rangle^2 \right| \\
&\le a^2 \left\{ \int_{t<m} \langle \mathbf{b}, \bar{\mathbf{X}}_T\rangle^2 dP + \int_{t<m} \langle \mathbf{b}, \bar{\mathbf{X}}_t\rangle^2 dP \right\} \\
&\le (a^p P\{t<m\})^{\frac{2}{p}} \left(\left\{ E\langle \mathbf{b}, \bar{\mathbf{X}}_T\rangle^{\frac{2p}{p-2}} \right\}^{\frac{p-2}{p}} + \left\{ E\langle \mathbf{b}, \bar{\mathbf{X}}_t\rangle^{\frac{2p}{p-2}} \right\}^{\frac{p-2}{p}} \right) \\
&\to 0 \quad \text{as } a \to \infty
\end{aligned}$$

by Hölder's inequality, (2.1) and the fact that $E\langle \mathbf{b}, \bar{\mathbf{X}}_T\rangle^{\frac{2p}{p-2}}$ and $E\langle \mathbf{b}, \bar{\mathbf{X}}_t\rangle^{\frac{2p}{p-2}}$ are uniformly bounded (see AW's inequality (6)). Other results can be proved similarly.

2.2. An example

The stopping times for both sequential point and interval estimations can often be written as

$$T = \inf\{n \ge m : \hat{\sigma}_n^{2\gamma} < cn\},$$

where $Y_1, Y_2, \ldots$ are i.i.d. observations having mean μ and variance σ^2, $\hat{\sigma}_n^2 = \sum_{i=1}^n (Y_i - \bar{Y_n})^2/n$ is the sample variance, $\bar{Y_n}$ is the sample mean, $0 < \gamma < 2$, and $c > 0$ is a constant allowed to go to zero. For the pure sequential sampling scheme, m is non-random and may depend on c; see e.g. Woodroofe (1982) and Martinsek (1983). Here we consider the sequential sampling scheme of Liu (1997), for which the value of m is random and defined in the following way. Let m_0 be the initial sample size which approaches infinity as $c \to 0$ at rate $O(c^{-b})$ (for some $0 < b < 1$), and so without loss of generality assume $m_0 = C_0^* c^{-b} = C_0 a^b$ for some finite positive constants C_0^* and C_0 ($C_0 = C_0^* \sigma^{-2b\gamma}$), where $a = \sigma^{2\gamma}/c$. Then define $m_i = \max\{(\rho_i/c)\hat{\sigma}_{m_{i-1}}^{2\gamma},\ m_{i-1}\} = \max\{\rho_i a \hat{\sigma}_{m_{i-1}}^{2\gamma}/\sigma^{2\gamma},\ m_{i-1}\}$, $i = 1, \ldots, k$, $m = m_k$, where natural number k and $0 < \rho_1 < \cdots < \rho_k < 1$ are given constants. Let $Z_n = n/\max(\hat{\sigma}_n^{2\gamma}/\sigma^{2\gamma}, C_0^{1/b} n^{1-1/b})$. Then T can be expressed as

$$\begin{aligned}
T &= \inf\{n \ge m :\ Z_n > a\} \\
&= \inf\{n \ge m :\ n + \langle \mathbf{c}, \mathbf{S}_n\rangle + \xi_n > a\}.
\end{aligned}$$

Here $\mathbf{c} = (0, -\gamma)$ and $\mathbf{X}_k = [(Y_k - \mu)/\sigma, (Y_k - \mu)^2/\sigma^2 - 1]$, and conditions (C1)-(C6) are satisfied with $p = 3$ and $\alpha = 3/2$ provided that $E|Y_1|^6 < \infty$ (see AW's Example 2 and Proposition 4). For condition (C7) we have

Lemma. *If $E|Y_1|^8 < \infty$, $\epsilon_3 > \rho_k$ and $1 > b \ge 4/(6-\gamma)$, then*

$$\int_{m_i > \epsilon_3 a} m_i^2 dP \to 0 \quad \text{for all } i = 0, \ldots, k.$$

Proof. We use mathematical induction on $0 \le i \le k$. The result is clearly true for $i = 0$ since $m_0 = O(a^b)$. Assume the result holds for $i = l$, we proceed to prove it for $i = l+1$, where $0 \le l \le k-1$. By noting that

$$\int_{m_{l+1}>\epsilon_3 a} m_{l+1}^2 dP = \Big(\int_{m_{l+1}>\epsilon_3 a, m_l>\epsilon_3 a} + \int_{m_{l+1}>\epsilon_3 a, m_l \le \epsilon_3 a}\Big) m_{l+1}^2 dP$$
$$\le \int_{m_l>\epsilon_3 a} m_{l+1}^2 dP + O(a^2)\int_{m_{l+1}>\epsilon_3 a} (\hat{\sigma}_{m_l}^2/\sigma^2)^{2\gamma} dP$$
$$\le \int_{m_l>\epsilon_3 a} m_l^2 dP + O(a^2)\Big\{\Big(\int_{m_l>\epsilon_3 a} + \int_{m_{l+1}>\epsilon_3 a}\Big)(\hat{\sigma}_{m_l}^2/\sigma^2)^{2\gamma} dP\Big\}$$

and that $\int_{m_l>\epsilon_3 a} m_l^2 dP = o(1)$ by the assumption of induction, it suffices to show that $(*1) := O(a^2)\int_{\hat{\sigma}_{m_j}^2/\sigma^2>1+\delta}(\hat{\sigma}_{m_l}^2/\sigma^2)^{2\gamma} dP = o(1)$ for $\delta > 0$ and $0 \le j \le l$. Let $W_i = (Y_i - \mu)^2/\sigma^2$. Then

$$(*1) \le O(a^2)\int_{\bar{W}_{m_j}>1+\delta} (\bar{W}_{m_l})^{2\gamma} dP$$
$$\le O(a^2)\int_{\bar{W}_{m_j}>1+\delta} |\bar{W}_{m_l} - 1|^{2\gamma} dP + O(a^2) P\{\bar{W}_{m_j} > 1+\delta\}$$

and $P\{\bar{W}_{m_j} > 1+\delta\} \le P\{\sup_{k\ge m_0} |\bar{W}_k - 1| > \delta\} = o(m_0^{-3})$ by the Baum-Katz inequality. It remains to show that $(*2) := \int_{\bar{W}_{m_j}>1+\delta} |\bar{W}_{m_l} - 1|^{2\gamma} dP$ is $o(a^{-2})$. For this we consider the two cases $l = 0$ and $l > 0$ separately. For $l = 0$,

$$(*2) \le m_0^{-\gamma}\Big\{\int |\sqrt{m_0}(\bar{W}_{m_0} - 1)|^4 dP\Big\}^{\gamma/2} (P\{\bar{W}_{m_0} > 1+\delta\})^{\frac{2-\gamma}{2}} = o(a^{-2})$$

since $\int |\sqrt{m_0}(\bar{W}_{m_0} - 1)|^4 dP = O(1)$ by the u.i. of $|\sqrt{n}(\bar{W}_n - 1)|^4$, and $P\{\bar{W}_{m_0} > 1+\delta\} = o(m_0^{-3})$ as before. For $l > 0$,

$$(*2) = \int_{\bar{W}_{m_j}>1+\delta} (\sqrt{a}/m_l)^{2\gamma} |\sum_1^{m_l}(W_i - 1)/\sqrt{a}|^{2\gamma} dP$$
$$\le O(a^{-\gamma})\int_{\bar{W}_{m_j}>1+\delta} |\sum_1^{m_l}(W_i - 1)/\sqrt{a}|^{2\gamma} dP$$
$$+O(a^{\gamma})\int_{m_l<c_0 a} |\sum_1^{m_l}(W_i - 1)/\sqrt{a}|^{2\gamma} dP \qquad (0 < c_0 < \rho_1)$$
$$\le O(a^{-\gamma})\Big\{\int |\sum_1^{m_l}(W_i - 1)/\sqrt{a}|^4 dP\Big\}^{\gamma/2}\Big(P\{\bar{W}_{m_j} > 1+\delta\}\Big)^{(2-\gamma)/2}$$
$$+O(a^{\gamma})\Big\{\int |\sum_1^{m_l}(W_i - 1)/\sqrt{a}|^4 dP\Big\}^{\gamma/2}(P\{m_l < c_0 a\})^{(2-\gamma)/2}$$
$$= o(a^{-2})$$

since $|\sum_1^{m_l}(W_i - 1)/\sqrt{a}|^4$ is u.i. by Chow and Yu's (1981) Lemma 5 (noting that $(m_l/a)^2$ is u.i. from the assumption of induction), $P\{\bar{W}_{m_j} > 1 + \delta\} = o(m_0^{-3})$ as before, and $P\{m_l < c_0 a\} = o(a^{-s})$ for any $s > 0$ (see AW, p. 511, last line). The proof is thus completed.

From the proof it is clear that if $E|Y_1|^{2\beta} < \infty$ for some $\beta \geq 1 + 3/b$, then

$$a^3 P\{m > \epsilon_2 a\} \to 0 \quad \text{as } a \to \infty \quad \text{for } 1 > \epsilon_2 > \rho_k,$$

i.e., condition (C8) holds with $p = 3$. Therefore, under the assumption that

$$E|Y_1|^{2\beta} < \infty \quad \text{for some } \beta > 4 \quad \text{and } 1 > b > \max\{4/(6-\gamma), 3/(\beta - 1)\},$$

Theorem 1 provides expansions for $E(T)$ and the risk $E[c^{-2}\sigma^{4\gamma-2}(\bar{Y}_T - \mu)^2 + T]$.

Finally, we note that if $m = m_0 = C_0 a^b$ (and so is non-random), then (C7) and (C8) are obviously true and the expansions for $E(T)$ and the risk above are supplied by Theorem 1 under the assumption that $E|Y_1|^6 < \infty$ and $0 < b < 1$, which agrees with the result of AW.

3. Second Order Expansion of the Risk for Fixed m

For ease of comparison, the notation of this section agrees largely with that of Martinsek (1983). We also assume that the following slightly more general stopping rule is used in place of T_R in (1.1):

$$t_R = \inf\{n \geq m : \ l_n n > A^{1/2}\hat{\sigma}_n^{\beta}\},$$

where $l_n = 1 + l_0/n + o(1/n)$ as $n \to \infty$. The main result of this section is given by

Theorem 2. *If Y_1 is continuous with a bounded density function, $E|Y_1|^{6r} < \infty$ for some $r > 1$ and the fixed integer $m > 1 + 3\beta$, then as $A \to \infty$,*

$$E(t_R) = A^{1/2}\sigma^{\beta} + \rho - l_0 - \frac{\beta}{2} - \frac{\beta(\beta+2)}{8} Var(W_1^2) + o(1),$$
$$R_{t_R} - R_{n_0} = 2\beta + \beta(\beta+1)\{E(W_1^3)\}^2 + (\beta^2/4 - \beta)Var(W_1^2) + o(1).$$

Here $W_i = (Y_i - \mu)/\sigma, \ i = 1, 2, \ldots,$ and $\rho = E(R)$, where the distribution of R is given by

$$P\{x \leq R \leq x + dx\} = \frac{1}{E(\tau)} P\{\tau + \sum_1^{\tau} \beta(1 - W_i^2)/2 > x\}dx, \qquad 0 < x < \infty,$$

with

$$\tau = \inf\{n \geq 1 : \ n + \sum_1^{n} \beta(1 - W_i^2)/2 > 0\}.$$

Remark. The second order expansion of R_{t_R} is of the same form as Martinsek's (1983) result, and agrees with Woodroofe's (1977) result for the normal distribution. The distribution-free second order expansion of $E(t_R)$ might be useful in other contexts. It is interesting to note that l_0 has no effect on R_{t_R} asymptotically.

Proof. Without loss of generality we set $\mu = 0$ and $\sigma = 1$ throughout the proof. We first show that under the assumptions

$$E|Y_1|^{6r} < \infty \;\text{ for some } r > 1 \;\text{ and} \tag{3.1}$$

$$E(\hat{\sigma}_m^{-\beta s}) < \infty \;\text{ for some } s > 3, \tag{3.2}$$

the conditions (C1)-(C6) of AW (1993) are satisfied with $p = 3$ and $\alpha = 3/2$ and, therefore, the second order expansions of $E(t_R)$ and $R_{t_R} - R_{n_0}$ follow directly from Theorem 1 and Corollary 1 of AW(1993).

For this, we first express t_R in the form (2) of AW (1993). Note that $t_R = \inf\{n \geq 1 : Z_n > a\}$, where $Z_n = l_n n(1/\hat{\sigma}_n)^\beta I_{(n\geq m)}$ and $a = A^{1/2}$. Write $Z_n = n+ < \mathbf{c}, \mathbf{S}_n > +\xi_n, \;\; n \geq 1$, where $\mathbf{c} = (0, \; -\beta/2)$, $\mathbf{X}_i = (Y_i, \; Y_i^2 - 1)$, $\mathbf{S}_n = \sum_{i=1}^n \mathbf{X}_i$, and $\xi_n = Z_n - n- < \mathbf{c}, \mathbf{S}_n >$. Now (C1), (C3)-(C6) of AW (1993) can be established by arguments similar to those of AW's (1993) Example 2 and Proposition 4. It remains to show (C2) and, for this, we only need to consider $n \geq m$ in the sequel. Let $\bar{l} = \sup_n l_n$ and let $0 < \epsilon_0 < 1$ be chosen such that $\epsilon_2 = (\bar{l}\epsilon_0)^{2/\beta} < 1/2$. We show that

$$\sup_{n\geq m} E\left[\left(Z_n - \frac{n}{\epsilon_0}\right)^+\right]^q \;<\; \infty \;\text{ for some }\; q > 3,$$

which is sufficient. Now, by the definitions of Z_n and ϵ_2,

$$\begin{aligned}
& E\Big[\Big(Z_n - \frac{n}{\epsilon_0}\Big)^+\Big]^q = E\Big[\Big(Z_n - \frac{n}{\epsilon_0}\Big)^+ I_{(\hat{\sigma}_n^2<\epsilon_2)}\Big]^q \\
\leq\; & E\left[Z_n^q I_{(\hat{\sigma}_n^2<\epsilon_2)}\right] \;\leq\; \bar{l}^q n^{q(\gamma+1)} E\left\{(n\hat{\sigma}_n^2)^{-q\gamma} I_{(\hat{\sigma}_n^2<\epsilon_2)}\right\} \qquad (\gamma = \beta/2) \\
\leq\; & \bar{l}^q n^{q(\gamma+1)} \left\{E(n\hat{\sigma}_n^2)^{-rq\gamma}\right\}^{1/r} \left\{P(\hat{\sigma}_n^2 < \epsilon_2)\right\}^{1/s} \quad (r > 1, \frac{1}{r} + \frac{1}{s} = 1) \\
\leq\; & \bar{l}^q \left\{E(n\hat{\sigma}_n^2)^{-rq\gamma}\right\}^{1/r} \left\{n^{sq(\gamma+1)} P(\hat{\sigma}_n^2 < \epsilon_2)\right\}^{1/s}.
\end{aligned} \tag{3.3}$$

By noting that $(n+1)\hat{\sigma}_{n+1}^2 > n\hat{\sigma}_n^2$ for all $n \geq 1$, for all $n \geq m$

$$E(n\hat{\sigma}_n^2)^{-rq\gamma} \;\leq\; E(m\hat{\sigma}_m^2)^{-rq\gamma}$$

and is bounded under assumption (3.2) by setting $r > 1$ sufficiently close to 1 and $q > 3$ sufficiently close to 3. Also note that $n^b P(\hat{\sigma}_n^2 < \epsilon_2) \longrightarrow 0$ as $n \to \infty$ for any given $b > 0$; see AW (1993, Example 2). It follows from these observations that (3.3) is uniformly bounded for $n \geq m$. The proof of (C2) is thus completed.

Next we show that assumption (3.2) holds if the Y_i have a bounded probability density function $f(\cdot)$ and $m > 1 + 3\beta$. We first show that

$$P\{(m\hat{\sigma}_m^2)^{-r} > y\} \ \leq \ C_0 y^{-(m-1)/(2r)} \quad \text{for all } y > 0, r > 0,$$

where C_0 is a constant. Letting $B = \sup_x f(x) < \infty$ and applying variable transformation $\mathbf{x} = A^T \mathbf{y}$, where $A^T = (\mathbf{a_1}, \ldots, \mathbf{a_m})$ is an orthogonal matrix with the first row $(1, \ldots, 1)/\sqrt{m}$, we have

$$\begin{aligned} & P\{(m\hat{\sigma}_m^2)^{-r} > y\} \\ = & \int \cdots \int_{\sum_1^m (y_i - \bar{y}_m)^2 < y^{-1/r}} \prod_1^m f(y_i) \ dy_1 \cdots dy_m \\ = & \int \cdots \int_{\sum_2^m x_i^2 < y^{-1/r}} \prod_1^m f(\mathbf{a}_i^T \mathbf{x}) \ dx_1 \cdots dx_m \\ = & \int \cdots \int_{\sum_2^m x_i^2 < y^{-1/r}} \Big\{ \int_{-\infty}^{\infty} \prod_1^m f(\mathbf{a}_i^T \mathbf{x}) dx_1 \Big\} \ dx_2 \cdots dx_m \\ \leq & \int \cdots \int_{\sum_2^m x_i^2 < y^{-1/r}} \sqrt{m} B^{m-1} \ dx_2 \cdots dx_m \\ = & \ C_0 y^{-(m-1)/(2r)}. \end{aligned}$$

So

$$\begin{aligned} E(\sqrt{m}\hat{\sigma}_m)^{-\beta s} = E(m\hat{\sigma}_m^2)^{-\gamma s} \ &= \ \int_0^\infty P\{(m\hat{\sigma}_m^2)^{-\gamma s} > y\} dy \\ &\leq C_0 \int_0^\infty y^{-(m-1)/(\beta s)} dy \ < \ \infty \end{aligned}$$

if $m > 1 + 3\beta$, by setting $s > 3$ sufficiently close to 3. We have therefore established (C1)-(C6) of AW (1993) under the assumptions of Theorem 2.

Acknowledgements

We would like to thank the editor for encouraging us to revise the paper, and an associate editor and a referee for useful comments.

References

Anscombe, F. J. (1953). Sequential estimation. *J. Roy. Statist. Soc. Ser. B* **15**, 1-21.

Aras, G. and Woodroofe, M. (1993). Asymptotic expansions for the moments of a randomly stopped average. *Ann. Statist.* **21**, 503-519.

Baum, L. E. and Katz, M. (1965). Convergence rates in the law of large numbers. *Trans. Amer. Math. Soc.* **120**, 108-123.

Bose, A. and Boukai, B. (1993). Sequential estimation results for a two-parameter exponential family of distributions. *Ann. Statist.* **21**, 484-502.

Chow, Y. S. and Martinsek, A. T. (1982). Bounded regret of a sequential procedure for estimation of the mean. *Ann. Statist.* **10**, 909-914.

Chow, Y. S. and Yu, K. F. (1981). The performance of a sequential procedure for the estimation of the mean. *Ann. Statist.* **9**, 184-189.

Chow, Y. S. and Robins, H. (1965). On the asymptotic theory of fixed width confidence intervals for the mean. *Ann. Math. Statist.* **36**, 457-462.

Ghosh, M. and Mukhopadhyay, N. (1979). Sequential point estimation of the mean when the distribution is unspecified. *Commun. Statist. Theory Methods* **8**, 637-651.

Katz, M. L. (1963). The probability in the tail of a distribution. *Ann. Math. Statist.* **34**, 312-318.

Liu,W. (1997). Improving the fully sequential procedure of Anscombe-Chow-Robins. *Ann. Statist.* **25**, 2164-2171.

Martinsek, A. T. (1983). Second order approximation to the risk of a sequential procedure. *Ann. Statist.* **11**, 827-836. Corrections: *Ann. Statist.* (1986) **14**, 359 and (1991) **19**, 2283.

Robbins, H. (1959). Sequential estimation of the mean of a normal population. In *Probability and Statistics* (Harold Cramer Volume), 235-245. Almquist and Wiksell, Stockholm.

Woodroofe, M. (1977). Second order approximations for sequential point and interval estimation. *Ann. Statist.* **5**, 984-995.

Woodroofe, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis.* Society for Industrial and Applied Mathematics, Philadelphia.

Department of Mathematics, University of Southampton, Southampton, SO17 1BJ, England.

E-mail: wl@maths.soton.ac.uk

(Received August 1998; accepted June 1999)