The Annals of Statistics
1973, Vol. 1, No. 6, 1203-1211

ASYMPTOTIC NORMALITY OF THE STOPPING TIMES OF SOME SEQUENTIAL PROCEDURES

BY P. K. BHATTACHARYA AND ASHIM MALLIK

University of Arizona and University of Minnesota

Two problems of sequential estimation, viz. the estimation of the mean of a normal distribution with unknown variance and the estimation of a binomial proportion are studied as the cost per observation tends to 0. For the first problem the asymptotic distribution of the stopping time of a procedure due to Robbins (1959) is shown to be normal. For the second problem the stopping time of a modification of Wald's (1951) procedure is asymptotically normal when the parameter is different from $\frac{1}{2}$. When the parameter is $\frac{1}{2}$, this stopping time does not enjoy asymptotic normality. The method employed is to first prove the convergence in probability of the stopping time which is then converted to convergence in distribution by using a theorem of Wittenberg (1964). This method also yields a new proof of a theorem of Siegmund (1968).

1. Introduction. Let $X_1, X_2, \cdots$ be a sequence of independent and identically distributed random variables with finite variance and $T_n = X_1 + \cdots + X_n$. In sequential analysis the stopping time τ of a procedure frequently turns out to be the smallest positive integer n for which T_n (or some nice function of T_n such as a quadratic) crosses a boundary $f(n)$. In the asymptotic theory of sequential analysis, instead of a fixed boundary $f(n)$, we have a system of boundaries $f_c(n)$ indexed by a parameter $c \downarrow 0$ and τ_c is the stopping time corresponding to the boundary $f_c(n)$. In the case when $E(X_1) > 0$ and $f_c(n) = c^{-1}n^u$, $0 \leqq u < 1$, the asymptotic distribution of τ_c was derived by Siegmund (1968). His result for special values of u relates to sequential procedures developed by Chow and Robbins (1965) and Darling and Robbins (1967). In this paper we consider two problems of sequential estimation, viz. the estimation of the mean of a normal distribution with unknown variance and the estimation of a binomial proportion, as the cost per observation $c \to 0$. For the first problem we derive the asymptotic distribution of the stopping time of a procedure studied by Robbins (1959) and Starr (1966), and for the second problem we derive the asymptotic distribution of the stopping time of a procedure which (except for some slight modifications introduced by Smith (1971)) belongs to a general class of sequential estimation rules due to Wald (1951).

Our method can be outlined as follows. Because of the nature of the boundaries $f_c(n)$, it is relatively easy to show that there is a function $\varphi(c)$ which tends to ∞ as $c \to 0$ such that $\tau_c/\varphi(c)$ converges in probability to a constant $\alpha > 0$. Let $Y_i = \{X_i - E(X_i)\}/(\text{Var}\,(X_i))^{\frac{1}{2}}$ and $S_n = Y_1 + \cdots + Y_n = \{T_n - nE(X_1)\} \div (\text{Var}\,(X_1))^{\frac{1}{2}}$. Then a theorem due to Wittenberg (1964) applies to S_{τ_c} and $S_{\tau_c - 1}$

Received April 1972.

AMS 1970 *subject classifications.* Primary 60G40; Secondary 62E20.

Key words and phrases. Stopping time, sequential estimation, asymptotic distribution.

and we see that both $S_{\tau_c}/\{\varphi(c)\alpha\}^{\frac{1}{2}}$ and $S_{\tau_c-1}/\{\varphi(c)\alpha\}^{\frac{1}{2}}$ converge in law to a standard normal random variable $N(0, 1)$. Now by definition of τ_c,

$$S_{\tau_c-1}(\text{Var}\,(X_1))^{\frac{1}{2}} < f_c(\tau_c - 1) - (\tau_c - 1)E(X_1)$$

and

$$S_{\tau_c}(\text{Var}\,(X_1))^{\frac{1}{2}} \geqq f_c(\tau_c) - \tau_c E(X_1)\,,$$

and we show that $\{f_c(\tau_c) - f_c(\tau_c - 1)\}/\{\varphi(c)\}^{\frac{1}{2}} = o_p(1)$. This implies that

$$\{f_c(\tau_c) - \tau_c E(X_1)\}/\{\varphi(c)\alpha\,\text{Var}\,(X_1)\}^{\frac{1}{2}} \to_{\mathscr{L}} N(0, 1)$$

from which the asymptotic normality of τ_c is derived. This method not only works for the two problems mentioned above (except in an interesting special case dealt with in Theorem 3(b)), but also yields a new proof of Siegmund's (1968) result.

The key step in our analysis is to convert the convergence in probability of τ_c to convergence in law. We state below a special case of a theorem of Wittenberg (1964) needed for this purpose.

THEOREM 1 (Wittenberg). *Let $Y_1, Y_2, \cdots$ be independent and identically distributed random variables with mean* 0 *and variance* 1 *on some probability space and $S_n = Y_1 + \cdots + Y_n$. If $\tau_1, \tau_2, \cdots$ is a sequence of positive integer-valued random variables on the same probability space such that τ_n/n converges in probability to a positive constant α, then $S_{\tau_n}/(n\alpha)^{\frac{1}{2}}$ converges in law to $N(0, 1)$.*

2. Estimation of a normal mean when the variance is unknown. $X_1, X_2, \cdots$ are independent normal random variables with mean μ and variance σ^2, both unknown. Consider the problem of sequentially estimating μ when the loss incurred in estimating μ by $\hat{\mu}$ after n observations is $|\hat{\mu} - \mu|^2 + cn$. If $\hat{\mu}$ is taken to be the current sample mean $\bar{X}_n = n^{-1}\sum_1^n X_i$ at the stopping time, the risk of stopping at time n becomes $\sigma^2 n^{-1} + cn$. Hence if σ^2 were known, one would use a fixed sample rule using either $[n_0]$ or $[n_0] + 1$ observations where n_0 is the solution in n of the equation $\sigma^2 = cn^2$ and $[a]$ is the largest integer $\leqq a$. When σ^2 is unknown, we can try to imitate this procedure by using the current sample variance $s_n^2 = (n - 1)^{-1}\sum_1^n (X_i - \bar{X}_n)^2$ in place of σ^2 at each stage. This gives rise to the following sequential procedure: "Stop sampling at the first $n \geqq 2$ for which $\sum_1^n (X_i - \bar{X}_n)^2 \leqq c(n - 1)n^2$, and estimate μ by $\bar{X}_n$." Let τ_c denote the stopping time of this procedure. We derive the asymptotic distribution of τ_c as $c \to 0$ in the following theorem.

THEOREM 2. *Suppose $X_1, X_2, \cdots$ are independent normal random variables with mean μ and variance σ^2, and $\bar{X}_n = n^{-1}\sum_1^n X_i$. If τ_c is the first n for which $\sum_1^n (X_i - \bar{X}_n)^2 \leqq c(n - 1)n^2$, then*

$$(\tau_c - c^{-\frac{1}{2}}\sigma)/(\tfrac{1}{2}c^{-\frac{1}{2}}\sigma)^{\frac{1}{2}} \to_{\mathscr{L}} N(0, 1)$$

as $c \downarrow 0$.

REMARK. Robbins (1959) first suggested the above procedure and investigated some of its asymptotic properties as $\sigma \to \infty$. Later, Starr (1966) used Robbins' argument in the context of a somewhat more general loss function and studied various asymptotic properties of the resulting procedure as $\sigma \to \infty$. Here, we are keeping σ^2 fixed and allowing $c \to 0$, but Theorem 2 can be easily converted to an asymptotic result in which c is fixed and $\sigma \to \infty$.

PROOF OF THEOREM 2. We first note that the joint distribution of $\{\sum_1^n (X_i - \bar{X}_n)^2, n = 2, 3, \cdots\}$ is the same as that of $\{\sum_1^{n-1} \xi_i, n = 2, 3, \cdots\}$ where ξ_1/σ^2, $\xi_2/\sigma^2, \cdots$ are independent χ_1^2 random variables (see, e.g., Robbins (1959)). We can, therefore, consider $t_c = \tau_c - 1$ as the first n for which $\sum_1^n \xi_i \leqq cn(n+1)^2$. This reduces the problem to the form described in the introduction with $f_c(n) = cn(n+1)^2$. We first show that

$$c^{\frac{1}{2}} t_c \to_{\text{a.s.}} \sigma \quad \text{as} \quad c \to 0 . \tag{1}$$

To show this, we note that t_c can equivalently be defined as the first n such that $y_n \leqq cg(n)$ where $y_n = n^{-1}\sigma^{-2} \sum_1^n \xi_i$ and $g(n) = \sigma^{-2}(n+1)^2$. Since $y_n \to_{\text{a.s.}} 1$ as $n \to \infty$, Lemma 1 of Chow and Robbins (1965) applies here. Hence $cg(t_c) \to_{\text{a.s.}} 1$, from which (1) follows. In the notation used in the introduction, we can rewrite (1) as

$$t_c/\varphi(c) \to_{\text{a.s.}} \alpha \quad \text{as} \quad c \to 0 \tag{2}$$

where $\varphi(c) = c^{-\frac{1}{2}} \to \infty$ and $\alpha = \sigma > 0$. Let $S_n = \sum_1^n (\xi_i - \sigma^2)/(2^{\frac{1}{2}}\sigma^2)$. Then by virtue of (2), Theorem 1 applies to S_{t_c} and S_{t_c-1} and we have

$$S_{t_c}/(c^{-\frac{1}{2}}\sigma)^{\frac{1}{2}} \to_{\mathscr{L}} N(0, 1) \quad \text{and} \quad S_{t_c-1}/(c^{-\frac{1}{2}}\sigma)^{\frac{1}{2}} \to_{\mathscr{L}} N(0, 1) . \tag{3}$$

But by definition of t_c,

$$S_{t_c-1}/(c^{-\frac{1}{2}}\sigma)^{\frac{1}{2}} > \{c(t_c - 1)t_c^2 - \sigma^2(t_c - 1)\}/(2c^{-\frac{1}{2}}\sigma^5)^{\frac{1}{2}} \tag{4a}$$

and

$$S_{t_c}/(c^{-\frac{1}{2}}\sigma)^{\frac{1}{2}} \leqq \{ct_c(t_c + 1)^2 - \sigma^2 t_c\}/(2c^{-\frac{1}{2}}\sigma^5)^{\frac{1}{2}} , \tag{4b}$$

and using (1) it is easy to see that

$$\text{RHS of (4a)} - \text{RHS of (4b)} = o_p(1) . \tag{5}$$

Combining (3), (4a), (4b) and (5), we have

$$\{ct_c(t_c + 1)^2 - \sigma^2 t_c\}/(2c^{-\frac{1}{2}}\sigma^5)^{\frac{1}{2}} \to_{\mathscr{L}} N(0, 1) . \tag{6}$$

We now write

$$\begin{aligned} &\{ct_c(t_c + 1)^2 - \sigma^2 t_c\}/(2c^{-\frac{1}{2}}\sigma^5)^{\frac{1}{2}} \\ &\quad = (c^{\frac{1}{2}} t_c \sigma^{-1})\{c^{\frac{1}{2}}(t_c + 1)\sigma^{-1} + 1\}(t_c + 1 - c^{-\frac{1}{2}}\sigma)/(2c^{-\frac{1}{2}}\sigma)^{\frac{1}{2}} \\ &\quad = \{1 + o_p(1)\}\{2 + o_p(1)\}(\tau_c - c^{-\frac{1}{2}}\sigma)/(2c^{-\frac{1}{2}}\sigma)^{\frac{1}{2}} \\ &\quad = \{1 + o_p(1)\}(\tau_c - c^{-\frac{1}{2}}\sigma)/(\tfrac{1}{2}c^{-\frac{1}{2}}\sigma)^{\frac{1}{2}} \end{aligned} \tag{7}$$

by using (1). The theorem now follows from (6) and (7).

3. Estimation of a binomial proportion. $X_1, X_2, \cdots$ are independent random variables taking values 0 and 1 with probabilities $1-\theta$ and θ respectively and $T_n = X_1 + \cdots + X_n$. Consider the problem of sequentially estimating θ when the loss incurred in estimating θ by $\hat{\theta}$ after n observations is $|\theta - \hat{\theta}|^2 + cn$. Following a line of argument due to Wald (1951), we see that if we always estimate θ by the current sample mean $n^{-1}T_n$, the risk of stopping at time n becomes $n^{-1}\theta(1-\theta) + cn$. Thus it is advantageous to take one more observation at time n only if $\{n^{-1} - (n+1)^{-1}\}\theta(1-\theta) > c$, i.e., $\theta(1-\theta) > cn(n+1)$. Replacing θ by its current estimate $n^{-1}T_n$ in the last inequality, we obtain a procedure which stops at the first n for which $T_n(n - T_n) \leqq cn^3(n+1)$ and estimates θ by $n^{-1}T_n$. However, this procedure makes no sense because $T_1(1-T_1) = 0$ with probability 1. Even if we modify this procedure by forcing the sampling to continue at least to the kth stage, we still find $T_k(k - T_k) = 0$ with probability $\theta^k + (1-\theta)^k$ even when c becomes very small which makes the procedure unsuitable for small c. To overcome this difficulty, Smith (1971) modified this procedure by using the estimate $(n+1)^{-1}(T_n + \frac{1}{2})$ for θ to determine when $\theta(1-\theta)$ drops below $cn(n+1)$. This gives rise to the following sequential procedure: "Stop sampling at the first n for which $(T_n + \frac{1}{2})(n - T_n + \frac{1}{2}) \leqq cn(n+1)^3$, and estimate θ by $n^{-1}T_n$." Smith (1971) showed that this procedure is asymptotically minimax in the sense of Wald (1951) as $c \to 0$. We now derive the asymptotic distribution of the stopping time τ_c of this procedure in the following theorem.

THEOREM 3. *Suppose $X_1, X_2, \cdots$ are independent random variables taking values 0 and 1 with probabilities $1-\theta$ and θ respectiely, $0 < \theta < 1$, and $T_n = X_1 + \cdots + X_n$. If τ_c is the first n for which $(T_n + \frac{1}{2})(n - T_n + \frac{1}{2}) \leqq cn(n+1)^3$, then as $c \downarrow 0$,*

(a)
$$[\tau_c - c^{-\frac{1}{2}}\{\theta(1-\theta)\}^{\frac{1}{2}}]/[c^{-\frac{1}{4}}|\theta - \tfrac{1}{2}|\{\theta(1-\theta)\}^{-\frac{1}{4}}] \to_{\mathscr{L}} N(0,1) \quad \text{if} \quad \theta \neq \tfrac{1}{2},$$

(b)
$$\begin{aligned} P[\chi_1^2 \leqq t-1] &\leqq \liminf_{c\downarrow 0} P[2\{(4c)^{-\frac{1}{2}} - \tau_c\} \leqq t] \\ &\leqq \limsup_{c\downarrow 0} P[2\{(4c)^{-\frac{1}{2}} - \tau_c\} \leqq t] \\ &\leqq P[\chi_1^2 \leqq t+1] \quad \text{if} \quad \theta = \tfrac{1}{2}. \end{aligned}$$

PROOF. We first show that

$$c^{\frac{1}{2}}\tau_c \to_{\text{a.s.}} \{\theta(1-\theta)\}^{\frac{1}{2}} \quad \text{as} \quad c \to 0. \tag{8}$$

To show this, we note that τ_c can be equivalently defined as the smallest n such that $y_n \leqq cg(n)$ where $y_n = [\{(n+1)^{-1}(T_n + \frac{1}{2})\}/\theta][\{1 - (n+1)^{-1}(T_n + \frac{1}{2})\}/(1-\theta)]$ and $g(n) = n(n+1)\{\theta(1-\theta)\}^{-1}$. Since $y_n \to_{\text{a.s.}} 1$ as $n \to \infty$, Lemma 1 of Chow and Robbins (1965) applies here. Hence $\tau_c \to_{\text{a.s.}} \infty$ and $cg(\tau_c) \to_{\text{a.s.}} 1$, from which (8) follows. In the notation used in the introduction, we now rewrite (8) as

$$\tau_c/\varphi(c) \to_{\text{a.s.}} \alpha \tag{9}$$

where $\varphi(c) = c^{-\frac{1}{2}}$ and $\alpha = \{\theta(1-\theta)\}^{\frac{1}{2}}$. Let $S_n = (T_n - n\theta)/\{\theta(1-\theta)\}^{\frac{1}{2}}$. Then by virtue of (9), Theorem 1 applies to S_{τ_c} and S_{τ_c-1} and we have

$$S_{\tau_c}/\{c^{-1}\theta(1-\theta)\}^{\frac{1}{4}} \to_{\mathscr{L}} N(0,1) \quad \text{and} \tag{10}$$
$$S_{\tau_c-1}/\{c^{-1}\theta(1-\theta)\}^{\frac{1}{4}} \to_{\mathscr{L}} N(0,1)\,.$$

Again, by definition of τ_c,

$$\{\tau_c\theta + \theta^{\frac{1}{2}}(1-\theta)^{\frac{1}{2}}S_{\tau_c} + \tfrac{1}{2}\}\{\tau_c(1-\theta) - \theta^{\frac{1}{2}}(1-\theta)^{\frac{1}{2}}S_{\tau_c} + \tfrac{1}{2}\} \leqq c\tau_c(\tau_c+1)^3 \tag{11a}$$

and

$$\{(\tau_c-1)\theta + \theta^{\frac{1}{2}}(1-\theta)^{\frac{1}{2}}S_{\tau_c-1} + \tfrac{1}{2}\} \times \{(\tau_c-1)(1-\theta) - \theta^{\frac{1}{2}}(1-\theta)^{\frac{1}{2}}S_{\tau_c-1} + \tfrac{1}{2}\} > c(\tau_c-1)\tau_c^{\,3}\,. \tag{11b}$$

At this point we assume that $0 < \theta < \frac{1}{2}$ and define $\beta = \{(1-\theta)/\theta\}^{\frac{1}{2}} - \{\theta/(1-\theta)\}^{\frac{1}{2}}$. Then $\beta > 0$. (For $\frac{1}{2} < \theta < 1$, interchange θ, $1-\theta$ and T_n, $n - T_n$ and follow the same proof.) We now rearrange the terms in (11a) and (11b) to obtain,

$$S_{\tau_c}/\{\varphi(c)\alpha\}^{\frac{1}{2}} + U_c \leqq \tau_c(c\tau_c^{\,2} - \alpha^2)/[\{\varphi(c)\alpha\}^{\frac{1}{2}}\alpha^2\beta] \tag{12a}$$

$$S_{\tau_c-1}/\{\varphi(c)\alpha\}^{\frac{1}{2}} + V_c > \tau_c(c\tau_c^{\,2} - \alpha^2)/[\{\varphi(c)\alpha\}^{\frac{1}{2}}\alpha^2\beta]\,, \tag{12b}$$

where

$$\{\varphi(c)\alpha\}^{\frac{1}{2}}\beta U_c = (2\alpha^2)^{-1} + (4\alpha^2\tau_c)^{-1} - \{\varphi(c)\alpha\}^{-2}(3\tau_c^{\,2} + 3\tau_c + 1) - \tau_c^{\,-1}S_{\tau_c}^2 \tag{13a}$$

$$\{\varphi(c)\alpha\}^{\frac{1}{2}}\beta V_c = (2\alpha^2)^{-1} - 1 + \{4\alpha^2(\tau_c-1)\}^{-1} - (\tau_c-1)^{-1}S_{\tau_c-1}^2\,. \tag{13b}$$

Now $\tau_c/\varphi(c) = \alpha + o_p(1)$ and by (10), $S_{\tau_c}/\{\varphi(c)\}^{\frac{1}{4}}$ and $S_{\tau_c-1}/\{\varphi(c)\}^{\frac{1}{4}}$ are both $O_p(1)$. Hence the RHS of (13a) and (13b) are both $O_p(1)$. Since $\varphi(c) \to \infty$ as $c \to 0$, this implies

$$U_c = o_p(1)\,, \qquad V_c = o_p(1)\,. \tag{14}$$

From (10), (12a), (12b) and (14) we now conclude that

$$\tau_c(c\tau_c^{\,2} - \alpha^2)/[\{\varphi(c)\alpha\}^{\frac{1}{2}}\alpha^2\beta] \to_{\mathscr{L}} N(0,1)\,. \tag{15}$$

We now use (9) to see that

$$\begin{aligned} &\tau_c(c\tau_c^{\,2} - \alpha^2)/[\{\varphi(c)\alpha\}^{\frac{1}{2}}\alpha^2\beta] \\ &\quad = (\tau_c/\varphi(c))\{(\tau_c/\varphi(c)) + \alpha\}\{\tau_c - \varphi(c)\alpha\}/[\{\varphi(c)\alpha\}^{\frac{1}{2}}\alpha^2\beta] \\ &\quad = \{\alpha + o_p(1)\}\{2\alpha + o_p(1)\}\{\tau_c - \varphi(c)\alpha\}/[\{\varphi(c)\alpha\}^{\frac{1}{2}}\alpha^2\beta] \\ &\quad = \{1 + o_p(1)\}\{\tau_c - \varphi(c)\alpha\}/\{\varphi(c)\alpha\beta^2/4\}^{\frac{1}{2}}\,. \end{aligned} \tag{16}$$

The first part of the theorem now follows from (15) and (16).

Now consider the case when $\theta = \frac{1}{2}$. In this case, (11a) and (11b) are rewritten as

$$\tau_c^{\,-1}S_{\tau_c}^2 + 4c(\tau_c+1)^2 \geqq (1 + \tau_c^{\,-1})^2\tau_c(1 - 4c\tau_c^{\,2}) \tag{17a}$$

and

(17 b) $$\tau_c^{-1}S^2_{\tau_c-1} - 4c\tau_c^2 < \tau_c(1 - 4c\tau_c^2) .$$

From (9), (10), (17a) and (17b) we conclude that

(18) $$W_c - 1 < \tau_c(1 - 4c\tau_c^2) \leqq W_c + 1$$

where $W_c \to_{\mathscr{L}} \{N(0, 1)\}^2 = \chi_1^2$. Again, using (9) we see that

(19) $$\begin{aligned}\tau_c(1 - 4c\tau_c^2) &= 2\{(4c)^{-\frac{1}{2}} - \tau_c\}(c^{\frac{1}{2}}\tau_c + 2c\tau_c^2) \\ &= 2\{(4c)^{-\frac{1}{2}} - \tau_c\}\{1 + o_p(1)\} .\end{aligned}$$

The second part of the theorem now follows from (18) and (19).

REMARK. The case $\theta = \frac{1}{2}$ is so different because of the following reason. Since $4(n^{-1}T_n)(1 - n^{-1}T_n) \leqq 1$, τ_c is bounded above by τ_c^* which is the first n for which $4cn^3 \geqq n + 1$. It is easy to see that $\tau_c^* \leqq 2 + [(4c)^{-\frac{1}{2}}]$. Hence, when $\theta = \frac{1}{2}$, $2 + [c^{-\frac{1}{2}}\{\theta(1 - \theta)\}^{\frac{1}{2}}] - \tau_c$ is a nonnegative random variable for all c. However, for other values of θ, $c^{-\frac{1}{2}}\{\theta(1 - \theta)\}^{\frac{1}{2}}$ is much less than τ_c^* and the mass of τ_c is distributed on both sides of $c^{-\frac{1}{2}}\{\theta(1 - \theta)\}^{\frac{1}{2}}$.

4. A theorem of Siegmund.

Let $X_1, X_2, \cdots$ be independent identically distributed random variables with mean $\mu > 0$ and finite variance σ^2, and $T_n = X_1 + \cdots + X_n$. Let τ_c denote the first n for which $T_n \geqq c^{-1}n^u$ where $0 \leqq u < 1$. Such stopping rules arise in the context of some sequential rules studied by Chow and Robbins (1965) and Darling and Robbins (1967). Here we use our method to give an alternative derivation of the asymptotic distribution of τ_c as $c \to 0$ first obtained by Siegmund (1968).

THEOREM 4 (Siegmund). *Suppose $X_1, X_2, \cdots$ are independent and identically distributed random variables with $E(X_1) = \mu > 0$, Var $(X_1) = \sigma^2 < \infty$ and $T_n = X_1 + \cdots + X_n$. If τ_c is the first n for which $T_n \geqq c^{-1}n^u$, $0 \leqq u < 1$, then as $c \downarrow 0$,*

$$\{\mu(1 - u)\sigma^{-1}\lambda_c^{-\frac{1}{2}}\}(\tau_c - \lambda_c) \to_{\mathscr{L}} N(0, 1)$$

where $\lambda_c = (c\mu)^{1/(u-1)}$.

We shall first show that τ_c/λ_c converges in probability to 1. Since the random variables here are unbounded in both directions, Lemma 1 of Chow and Robbins (1965) does not apply here as it did in the other two cases. This and another fact needed in the proof of Theorem 4 are established in the following two lemmas.

LEMMA 1.
$$\tau_c/\lambda_c \to_P 1 \qquad as \quad c \downarrow 0 .$$

PROOF. Fix $\delta > 0$. Then

$$P[\tau_c > \lambda_c(1 + \delta)] \leqq P[\textstyle\sum_1^{\lambda_c(1+\delta)} (X_i - \mu) < c^{-1}\{\lambda_c(1 + \delta)\}^u - \lambda_c(1 + \delta)\mu]$$

which is easily seen to be $O(c^{1/(1-u)})$ by the Tchebychev inequality. On the other hand

$$\begin{aligned}P[\tau_c \leqq \lambda_c(1-\delta)] &= P[\textstyle\sum_1^n (X_i - \mu) \\ &\geqq c^{-1}n^u - n\mu \quad \text{for some} \quad n \leqq \lambda_c(1-\delta)] \\ &= P[\max_{1\leqq n\leqq \lambda_c(1-\delta)} (c^{-1}n^u - n\mu)^{-1} \textstyle\sum_1^n (X_i - \mu) \geqq 1]\end{aligned}$$

since $c^{-1}n^u - n\mu > 0$ for $1 \leqq n \leqq \lambda_c(1-\delta)$. Now for c small enough so that $\lambda_c \geqq \{\delta + u^{1/(1-u)} - 1\}^{-1}$, the numbers $(c^{-1}n^u - n\mu)^{-1}$, $1 \leqq n \leqq \lambda_c(1-\delta)$ are non-increasing. We can thus apply the Hájek–Rényi inequality (1955) to get

$$\begin{aligned}P[\tau_c \leqq \lambda_c(1-\delta)] &\leqq \sigma^2 c^2 \textstyle\sum_{n=1}^{\lambda_c(1-\delta)} (n^u - nc\mu)^{-2} \\ &= \sigma^2 c^2 \textstyle\sum_{n=1}^{\lambda_c(1-\delta)} n^{-2u}\{1 - (n/\lambda_c)^{1-u}\}^{-2} \\ &\leqq \sigma^2 c^2 \{1 - (1-\delta)^{1-u}\}^{-2} \textstyle\sum_{n=1}^{\lambda_c(1-\delta)} n^{-2u} .\end{aligned} \tag{20}$$

We now consider three cases.

Case (i). $u > \frac{1}{2}$. Here $\sum_{n=1}^{\lambda_c(1-\delta)} n^{-2u} < \sum_{n=1}^{\infty} n^{-2u} < \infty$. Hence $c^2 \sum_{n=1}^{\lambda_c(1-\delta)} n^{-2u} = O(c^2)$.

Case (ii). $u = \frac{1}{2}$. Here $\sum_{n=1}^{\lambda_c(1-\delta)} n^{-2u} < 1 + \log\{\lambda_c(1-\delta)\}$. Hence $c^2 \sum_{n=1}^{\lambda_c(1-\delta)} n^{-2u} = O(c^2 \log c)$.

Case (iii). $u < \frac{1}{2}$. Here $\sum_{n=1}^{\lambda_c(1-\delta)} n^{-2u} < 1 + (1-2u)^{-1}[\{\lambda_c(1-\delta)\}^{1-2u} - 1]$. Hence $c^2 \sum_{n=1}^{\lambda_c(1-\delta)} n^{-2u} = O(c^{1/(1-u)})$.

Thus in all cases the RHS of (20) goes to 0 as $c \to 0$ and that concludes the proof.

LEMMA 2. *Suppose* $Z_n = 1 + o_p(1)$, *and* a *is a constant. Then*

$$Z_n^a - 1 = (Z_n - 1)\{a + o_p(1)\} .$$

PROOF. By stochastic Taylor expansion.

PROOF OF THEOREM 4. Let $S_n = (T_n - n\mu)/\sigma$. Since $\tau_c/\lambda_c \to 1$ in probability and $\lambda_c \to \infty$ as $c \to 0$, Theorem 1 applies on S_{τ_c} and $S_{\tau_c - 1}$ and we have

$$S_{\tau_c}/\lambda_c^{\frac{1}{2}} \to_{\mathscr{L}} N(0,1) \quad \text{and} \quad S_{\tau_c-1}/\lambda_c^{\frac{1}{2}} \to_{\mathscr{L}} N(0,1) . \tag{21}$$

Again, by definition of τ_c,

$$S_{\tau_c}/\lambda_c^{\frac{1}{2}} \geqq (c^{-1}\tau_c^u - \tau_c\mu)/(\sigma\lambda_c^{\frac{1}{2}}) \tag{22a}$$

and

$$S_{\tau_c-1}/\lambda_c^{\frac{1}{2}} < (c^{-1}(\tau_c - 1)^u - (\tau_c - 1)\mu)/(\sigma\lambda_c^{\frac{1}{2}}) . \tag{22b}$$

Now since $\tau_c^{-1} = o_p(1)$, we use Lemma 2 to get

$$\begin{aligned}\text{RHS of (22a)} - \text{RHS of (22b)} &= [c^{-1}\{\tau_c^u - (\tau_c - 1)^u\} - \mu]/(\sigma\lambda_c^{\frac{1}{2}}) \\ &= c^{-1}\tau_c^u\{1 - (1 - \tau_c^{-1})^u\}/(\sigma\lambda_c^{\frac{1}{2}}) + o(1) \\ &= c^{-1}\tau_c^u[\tau_c^{-1}\{u + o_p(1)\}]/(\sigma\lambda_c^{\frac{1}{2}}) + o(1) \\ &= u[c\lambda_c^{1-u}\{1 + o_p(1)\}]^{-1}\{1 + o_p(1)\}/(\sigma\lambda_c^{\frac{1}{2}}) + o(1) \\ &= u\sigma^{-1}c^{-1}\lambda_c^{u-\frac{3}{2}}\{1 + o_p(1)\} + o(1) = O_p(c^{\frac{1}{2}(1-u)}) .\end{aligned} \tag{23}$$

By (22a), (22b) and (23), we have

$$(\mu\tau_c - c^{-1}\tau_c^{\,u})/(\sigma\lambda_c^{\frac{1}{2}}) \to_{\mathscr{L}} N(0, 1)\,. \tag{24}$$

Finally, using Lemma 1 and Lemma 2 again, we have

$$\begin{aligned}(\mu\tau_c - c^{-1}\tau_c^{\,u}) &= \mu\tau_c\{1 - (\lambda_c/\tau_c)^{1-u}\} = \mu\tau_c\{1 - (\lambda_c/\tau_c)\}\{1 - u + o_p(1)\} \\ &= \mu(1-u)(\tau_c - \lambda_c)\{1 + o_p(1)\}\,,\end{aligned} \tag{25}$$

and the theorem follows from (24) and (25).

REMARK. In the framework described at the beginning of this section, let t_c be the first n for which $T_n \leqq cn^u$, $u > 1$. What can we say about the asymptotic distribution of t_c as $c \downarrow 0$? In order to make the probability of early crossing negligible as $c \to 0$ in this case, we must have $P(X_1 > 0) = 1$. Under this condition we tried to look for some simple relationship between τ_c and t_c hoping that Theorem 4 will give the answer to the above question as an immediate corollary. The obvious way to connect the two problems is to reflect the boundary cn^u, $u > 1$, as well as the sample path $\{(n, T_n),\ n = 1, 2, \cdots\}$ across the equiangular line. In this way the problem is transformed to one in which a renewal process crosses the boundary $c^{-1/u}s^{1/u}$. If X_i are exponential random variables, the renewal process is a Poisson process and in that case the asymptotic distribution of t_c can be obtained from Theorem 4 with a little effort. This argument also extends to the case where the X_i are gamma random variables with any degrees of freedom. However, we could see no way to make this argument work in general because the renewal process $\{(T_n, n),\ n = 1, 2, \cdots\}$ does not have stationary and independent increments in general. However, we can apply the method used in this paper directly to this problem to get the result,

$$\{\mu(u-1)\sigma^{-1}\lambda_c^{-\frac{1}{2}}\}(t_c - \lambda_c) \to_{\mathscr{L}} N(0, 1)$$

as $c \downarrow 0$, where $\lambda_c = (\mu/c)^{1/(u-1)}$. Except for some small differences in the boundary, Theorem 2 can now be regarded as a special case of this result.

REFERENCES

[1] CHOW, Y. S. and ROBBINS, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.* **36** 457–462.

[2] DARLING, D. A. and ROBBINS, H. (1967). Iterated logarithm inequalities. *Proc. Nat. Acad. Sci. U.S.A.* **57** 1188–1192.

[3] HÁJEK, J. and RÉNYI, A. (1955). Generalization of an inequality of Kolmogorov. *Acta Math. Acad. Sci. Hungar.* **6** 281–283.

[4] ROBBINS, H. (1959). Sequential estimation of the mean of a normal population. *Probability and Statistics* (Harald Cramér Volume), Almquist and Wiksell, Upsala.

[5] SIEGMUND, D. (1968). On the asymptotic normality of one-sided stopping rules. *Ann Math. Statist.* **39** 1493–1497.

[6] SMITH, ROBERT P. (1971). Contributions in sequential analysis. Ph. D. dissertation, Univ. of Arizona.

[7] STARR, N. (1966). On the asymptotic efficiency of a sequential procedure for estimating the mean. *Ann. Math. Statist.* **37** 1173–1185.

[8] WALD, ABRAHAM (1951). Asymptotic minimax solutions of sequential point estimation problems. *Proc. Second Berkeley Symp. Math. Statist. Prob.* Univ. of California Press.

[9] WITTENBERG, HELEN (1964). Limiting distributions of random sums of independent random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **1** 7–18.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ARIZONA
TUCSON, ARIZONA 85721

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF PITTSBURGH
PITTSBURGH, PENNSYLVANIA 15213