On asymptotic normality of sequential estimators for branching processes with immigration

Yongcheng Qi*

Department of Mathematics and Statistics, University of Minnesota Duluth, Campus Center 140, 1117 University Drive, Duluth, MN 55812, USA

Received 8 July 2004; received in revised form 18 October 2005; accepted 22 October 2006

Available online 22 December 2006

Abstract


© 2006 Elsevier B.V. All rights reserved.

MSC: Primary 62L12; secondary 60J80

Keywords: Sequential estimator; Modified sequential estimator; Branching process; Stopping time; Asymptotic normality

1. Introduction

The study of Galton–Watson processes with immigration has drawn much attention since the work by Heyde and Seneta (1972, 1974), and early study on the estimation of means \( m \) of offspring numbers and immigration rates dates back to Smoluchowski (1916). Based only on the generation sizes, the maximum likelihood approach in the parametric models is, in general, too complicated to be useful, as noted by Heyde and Seneta (1972). However, if the numbers of immigrants are also observed, the approach yields very useful results under some parametric models. See, e.g., Bhat and Adke (1981), Venkataraman (1982) and Venkataraman and Nanthi (1982).

Many estimators of the offspring mean have been proposed and studied in Heyde and Seneta (1971), Heyde (1970), Quine (1976), Klimko and Nelson (1978), Seneta (1970) and Wei and Winnicki (1987) under subcritical \( (m < 1) \), critical \( (m = 1) \) or supercritical \( (m > 1) \) cases.

In an attempt to provide a unified estimator, Wei and Winnicki (1990) proposed the conditional weighted least squares method. It turns out that the limiting distribution for the estimator in the critical case, drastically different from that of the other two cases, is nonnormal.

* Fax: +1 218 726 8399.
E-mail address: yqi@d.umn.edu.

0378-3758/ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2006.10.007
Later, Sriram et al. (1991) proposed a sequential estimator based on full information on both the generation sizes and the immigration process. It is shown that this sequential estimator is consistent and asymptotically normal in the subcritical case and the critical case. Recently, Shete and Sriram (1998) modified the sequential estimator and constructed a so-called fixed-precision estimator and proved that the modified estimator is unbiased in general and is asymptotically efficient as sequential estimator in the subcritical case and in the critical case. But some interesting questions remain unsolved. Whether or not the sequential estimator and the modified sequential estimator are asymptotically normal in the supercritical case are still open problems.

Recently, Qi and Reeves (2002) proposed a class of two-stage sequential estimators. These two-stage estimators are asymptotically normal in all three cases, and therefore, they can be used to construct the confidence interval for the mean parameter \( m \) without requirement of any prior information about \( m \).

In the present paper we will study the limiting distributions of sequential estimators and modified sequential estimators. The contribution of the paper is to prove the asymptotic normality for the sequential estimators in Sriram et al. (1991) and the modified sequential estimators in Shete and Sriram (1998) in the supercritical case. It is worth mentioning that the modified sequential estimators seem not asymptotically normal in the supercritical case in Shete and Sriram (1998) as are shown in the simulation study of that paper. However, we will use some different random normalization constants for the modified sequential estimators in the present paper. That gives a unified normal limit for the modified sequential estimator in all three cases. Even so, we are unable to show in this paper that both the sequential and modified sequential estimators are uniformly normal in the supercritical case when the \( m \) is limited to a bounded interval. That might be true as suggested by the result of Shiryaev and Spokoiny (1997) for AR(1) processes when a standard normal error is assumed.

The paper is organized as follows. In Section 2, we introduce the sequential estimation methods, including the sequential estimators by Sriram et al. (1991), the modified sequential estimators by Shete and Sriram (1998) and the two-stage sequential estimators by Qi and Reeves (2002). In Section 3 we show that the sequential estimators are asymptotically normal in the supercritical case. In Section 4 we focus on study of the limiting distributions for the modified sequential estimators and prove that these estimators always have normal limits if (random) normalization constants are appropriately chosen.

2. Sequential estimation

Define the following branching process with immigration:

\[
Z_n = \sum_{k=1}^{Z_{n-1}} \xi_{n-1,k} + Y_n, \quad n \geq 1, \tag{2.1}
\]

where \( Z_n \) denotes the size of the \( n \)th generation of a population, \( Y_n \) the number of immigrants in the \( n \)th generation, and \( \xi_{n-1,k} \) the number of offspring of the \( k \)th individual belonging to the \( (n - 1) \)th generation. The initial population size \( Z_0 \) is a random variable independent of \( \{Z_n\} \) and \( \{Y_n\} \). \( \{\xi_{n-1,k}, n \geq 1, k \geq 1\} \) and \( \{Y_n, n \geq 1\} \) are an independent array and sequence of independent and identically distributed (i.i.d) integer-valued random variables with unknown means \( m \) and \( \lambda \), and variances \( \sigma^2 \in (0, \infty) \) and \( \sigma_Y^2 \in (0, \infty) \), respectively.

Throughout this paper we assume that both \( \{Z_n\} \) and \( \{Y_n\} \) are observed. An estimator for the offspring mean is given by

\[
\hat{m}_n = \frac{\sum_{i=1}^{n} (Z_i - Y_i)}{\sum_{i=1}^{n} Z_{i-1}}. \tag{2.2}
\]

In order to achieve some better asymptotic properties for the estimation, Sriram et al. (1991) defined the stopping rule by

\[
N_c = \inf\{n \geq 1 : \sum_{i=1}^{n} Z_{i-1} \geq c \sigma^2\}, \tag{2.3}
\]

where \( c > 0 \) is some constant chosen appropriately. The sequential estimator of \( m \) is then given by \( \hat{m}_{N_c} \). They put the variance \( \sigma^2 \) in the definition of \( N_c \) in their paper in order to be able to prove the uniformly asymptotic normality and
to control the expected mean squared error of estimation. Such a stopping rule $N_c$ is well defined for all $c > 0$ if $\sigma^2$ is known.

Shete and Sriram (1998) proposed a modified sequential estimator defined by

$$
\hat{m}(c) = \frac{\sum_{j=1}^{N_c-1} (Z_j - Y_j) + \theta_c (Z_{N_c} - Y_{N_c})}{c\sigma^2},
$$

where $\theta_c \in (0,1]$ is a correction multiplier uniquely defined by

$$
\sum_{j=1}^{N_c-1} Z_{j-1} + \theta_c Z_{N_c-1} = c\sigma^2.
$$

It is shown in Shete and Sriram (1998) that the modified estimator is a fixed precision estimator, that is, it is unbiased and has bounded mean squared error for all $m > 0$. Just like the sequential estimator, the modified estimator is proved to be asymptotically normal in case $m \in (0,1]$.

In order to define the two-stage sequential estimators based on $N_c$, let $G = G(c, N_c)$ be an integer-valued function of $c$ and $N_c$ satisfying that as $c \to \infty$,

$$
\frac{G(c, N_c) - N_c}{c^2} \to 0 \quad \text{a.s. for some } \alpha \in (0, \frac{1}{2}) \quad \text{if } m \in (0,1]
$$

or

$$
G(c, N_c) - N_c \to \infty \quad \text{a.s. if } m > 1.
$$

Then $\hat{m}_G$ is the two-stage sequential estimator of Qi and Reeves (2002). In practice, the function $G$ satisfying both (2.5) and (2.6) can be easily selected free of $m$. Many examples of such choices are given in Qi and Reeves (2002).

Qi and Reeves (2002) proved the following theorem:

**Theorem 2.1.** $\hat{m}_G$ converges almost surely to $m$ as $c$ tends to infinity. If $G = G(c, N_c)$ satisfies (2.5) and (2.6), then

$$
\left( \sum_{i=1}^{G} Z_{i-1} \right)^{1/2} \frac{\hat{m}_G - m}{\sigma} \xrightarrow{d} N(0, 1).
$$

For simplicity, we assume in the following sections that $\sigma^2$ is known in the definition (2.3). When $\sigma^2$ is unknown, we can absorb the $\sigma^2$ into the constant $c$ and instead define

$$
N(c) = \inf \left\{ n \geq 1 : \sum_{i=1}^{n} Z_{i-1} \geq c \right\},
$$

and define the sequential estimator, the modified sequential estimator and the two-stage sequential estimators, correspondingly by the new stopping rule $N(c)$ instead of $N_c$. For example, we can use $\hat{m}_{N(c)}$ as a sequential estimator of $m$. Then our results in the sequel are still valid. Moreover, the variance $\sigma^2$ can be estimated. Define the following estimator for $\sigma^2$:

$$
\hat{\sigma}_{n}^2 = \frac{1}{n} \sum_{i=1}^{n} \frac{(Z_i - \hat{m}_n Z_{i-1} - Y_i)^2}{Z_{i-1}}
$$

as in Shete and Sriram (1998). It is proved in Proposition 4.1 of Shete and Sriram (1998) that $\hat{\sigma}_{n}^2 \to \sigma^2$ a.s. as $n \to \infty$ provided that $E_{\xi}^4 < \infty$. When $m > 1$, the strong consistency of $\hat{\sigma}_{n}^2$ can be proved in a similar fashion. Since both $N(c)$ and $G$ tend to infinity almost surely as $c \to \infty$, it can be seen that both $\hat{\sigma}_{N(c)}^2$ and $\hat{\sigma}_G^2$ are strongly consistent. This fact is important if one wishes to construct a confidence interval for $m$ by using either of the three sequential estimators when the variance $\sigma^2$ is unknown.
3. Asymptotic normality of sequential estimators

We first give our theorem about the limiting distribution of the sequential estimators and then provide its proof. Although we deal with the asymptotic normality of the sequential estimators only in the supercritical case in this paper, for completeness we state the theorem for all three cases.

**Theorem 3.1.** For any \( m \in (0, \infty) \),

\[
\left( \sum_{i=1}^{N_c} Z_{i-1} \right)^{1/2} \frac{\hat{m}_{N_c} - m}{\sigma} \xrightarrow{d} N(0, 1). \tag{3.1}
\]

**Proof.** The theorem was proved in Sriram et al. (1991) when \( m \in (0, 1] \). Therefore, we need only to show (3.1) for the supercritical case \( m > 1 \).

Define

\[ W = \lim_{n \to \infty} \frac{Z_n}{m^n} \tag{3.2} \]

Then Seneta (1970) proved that

\[ W \text{ is continuous and } 0 < W < \infty \text{ a.s.} \tag{3.3} \]

Therefore, it is concluded that

\[ \lim_{n \to \infty} \frac{\sum_{j=1}^{n} Z_{j-1}}{mn} = \frac{W}{m-1} =: V \text{ a.s.} \tag{3.4} \]

Suppose we are able to show that

\[ \lim_{c \to \infty} P \left( N_c = \left\lfloor \log \sigma^2 - \log V \right\rfloor + 1 \right) = 1, \tag{3.5} \]

where \( \lfloor x \rfloor \) denotes the largest integer less than \( x \). Now we can select a sequence of integers \( \{k(c)\} \) depending only on \( c \) and satisfying

\[ \lim_{c \to \infty} k(c) = \infty \text{ and } \lim_{c \to \infty} c/m^{k(c)} = \infty. \tag{3.6} \]

Set \( s(c) = c/m^{k(c)} \). Then from (3.5)

\[ \lim_{c \to \infty} P(N_c = N_{s(c)} + k(c)) = 1. \tag{3.7} \]

Therefore, showing (3.1) is equivalent to proving

\[
\left( \sum_{i=1}^{N_{s(c)}+k(c)} Z_{i-1} \right)^{1/2} \frac{\hat{m}_{N_{s(c)}+k(c)} - m}{\sigma} \xrightarrow{d} N(0, 1),
\]

which is valid by applying Theorem 2.1 to the two-stage estimator \( \hat{m}_{N_{s(c)}+k(c)} \) with \( G = N_{s(c)} + k(c) \). The two-stage sequential estimator is based on the stopping rule \( N_{s(c)} = \inf\{n \geq 1 : \sum_{i=1}^{n} Z_{i-1} \geq s(c)\sigma^2\} \). Eq. (3.7) simply implies that the stopping rule \( N_c \) is asymptotically equivalent to a two-stage sequential estimator. Therefore, the remaining task is to prove (3.5).

First rewrite (3.4) as

\[
\sum_{j=1}^{n} Z_{j-1} = V m^n (1 + \delta_n),
\]

where \( \lim_{n \to \infty} \delta_n = 0 \text{ a.s.} \).
Note that for any fixed \( |x| < 1 \)
\[
t(c, x) := \inf \{ n \geq 1 : Vm^n(1 + x) \geq c \sigma^2 \} = \left[ \frac{\log c \sigma^2 - \log(1 + x) - \log V}{\log m} \right] + 1.
\]

Let \( \Omega = \{ \omega : \lim_{n \to \infty} \delta_n = 0 \} \). Then \( P(\Omega) = 1 \). Now fix \( \delta > 0 \). For any \( \omega \in \Omega \), if \( n \) is large enough, then
\[
Vm^n(1 - \delta) < \sum_{j=1}^{n} Z_{j-1} < Vm^n(1 + \delta).
\]

Therefore, \( t(c, \delta) \leq N_c \leq t(c, -\delta) \) if \( c \) is large. This implies
\[
\lim_{c \to \infty} P(t(c, \delta) \leq N_c \leq t(c, -\delta)) = 1.
\]

Note that \( t(c, 0) = [(\log c \sigma^2 - \log V) / \log m] + 1 \) and \( t(c, \delta) \leq t(c, 0) \leq t(c, -\delta) \). Thus
\[
\lim_{c \to \infty} P(|N_c - t(c, 0)| \leq t(c, -\delta) - t(c, \delta)) = 1.
\]

If we can show that
\[
\lim_{\delta \to 0} \sup_{c} P(t(c, -\delta) - t(c, \delta) \geq 1) = 0,
\]
then for any \( \delta > 0 \)
\[
\lim_{c \to \infty} P \left( N_c \neq \left\lfloor \frac{\log c - \log V}{\log m} \right\rfloor + 1 \right)
\]
\[
= \lim_{c \to \infty} P(N_c \neq t(c, 0))
\]
\[
= \lim_{c \to \infty} P(|N_c - t(c, 0)| \geq 1)
\]
\[
= \lim_{c \to \infty} P(|N_c - t(c, 0)| \geq 1, |N_c - t(c, 0)| \leq t(c, -\delta) - t(c, \delta))
\]
\[
\leq \lim_{c \to \infty} P(t(c, -\delta) - t(c, \delta) \geq 1)
\]
\[
\leq \sup_{c} P(t(c, -\delta) - t(c, \delta) \geq 1)
\]
which tends to 0 as \( \delta \to 0 \), and (3.5) is proved.

Since
\[
\frac{\log c \sigma^2 - \log(1 - \delta) - \log V}{\log m} - \frac{\log c \sigma^2 - \log(1 + \delta) - \log V}{\log m} = \frac{\log(1 + \delta) - \log(1 - \delta)}{\log m} = : r_\delta,
\]
which tends to 0 as \( \delta \to 0 \), we have
\[
P(t(c, -\delta) - t(c, \delta) \geq 1)
\]
\[
= P \left( \bigcup_{k=1}^{\infty} \left\{ \frac{\log c \sigma^2 - \log(1 + \delta) - \log V}{\log m} \leq k \leq \frac{\log c \sigma^2 - \log(1 - \delta) - \log V}{\log m} \right\} \right)
\]
\[
\leq \sum_{k} P(k - r_\delta < \frac{\log c \sigma^2 - \log V}{\log m} < k + r_\delta).
\]

Denote \( H = \log V / \log m \) and \( v(c) = \log c \sigma^2 / \log m \). Then \( H \) has a continuous distribution function. Let \( Q(H, s) \) denote the concentration function of \( H \), that is
\[
Q(H, s) = \sup_{x} P(x \leq H \leq x + s).
\]
Then \( \lim_{s \to 0} Q(H, s) = 0 \). Therefore, we have for any \( r_\delta < \frac{1}{2} \) and positive integer \( N \),

\[
\sup_c P(t(c, -\delta) - t(c, \delta) \geq 1) \\
\leq \sup_c \sum_k P(k - r_\delta < v(c) - H < k + r_\delta) \\
\leq \sup_c \sum_k P(v(c) - k - r_\delta < H < v(c) - k + r_\delta) \\
\leq P(|H| > N) + (2N + 1)Q(H, 2r_\delta),
\]

which tends to 0 by first letting \( r_\delta \) tend to 0 and then \( N \) to infinity. This proves (3.8). \( \Box \)

4. Asymptotics of the modified sequential estimator

First we state a theorem on asymptotic normality of the modified sequential estimator.

**Theorem 4.1.** For \( m \in (0, \infty) \)

\[
\frac{c\sigma^2}{\sqrt{\theta^2cZ_{Nc-1} + \sum_{j=1}^{N_c-1} Z_{j-1}}} \frac{\hat{m}(c) - m}{\sigma} \xrightarrow{d} N(0, 1). \tag{4.1}
\]

**Proof.** It is shown in Sriram et al. (1991) that in the case \( m \leq 1 \)

\[
\frac{Z_{Nc-1}}{\sum_{i=1}^{Nc-1} Z_{i-1}} \to 0 \quad \text{and} \quad \frac{\sum_{i=1}^{N_c} Z_{i-1}}{c\sigma^2} \to 1 \quad \text{a.s. as} \quad c \to \infty,
\]

(see, e.g., Eqs. (4.5) and (4.7) in the aforementioned paper.) Thus it is seen that

\[
\frac{c\sigma^2}{\sqrt{\theta^2cZ_{Nc-1} + \sum_{j=1}^{N_c-1} Z_{j-1}}} \frac{\hat{m}(c) - m}{\sigma} = (1 + o_p(1)) \frac{\sum_{i=1}^{N_c} Z_{i-1}}{c\sigma^2} \frac{\hat{m}(c) - m}{\sigma}.
\]

Hence, the theorem follows from Theorem 4.1 of Shete and Sriram (1998).

Next, assume \( m > 1 \). We need to introduce some notation as in Qi and Reeves (2002).

Let \( \{\xi_i\} \) be i.i.d. random variables distributed the same as \( \xi_{1,1} \) and independent of \( Z_0 \) and \( \{Y_n\} \). For any nonnegative integer \( t \) define

\[
S_1(t) = Z'_0(t) = t \quad \text{and} \quad Z'_1(t) = \sum_{i=1}^{Z'_0(t)} \xi_i + Y_1.
\]

For \( n \geq 2 \) define

\[
S_n(t) = \sum_{i=1}^{n} Z'_{i-1}(t) \quad \text{and} \quad Z'_n(t) = \sum_{i=S_{n-1}(t)+1}^{S_n(t)} \xi_i + Y_n.
\]

Then one can easily show that the sequences \( \{Z_{n-1}, Y_n, n \geq 1\} \) and \( \{Z'_{n-1}(Z_0), Y_n, n \geq 1\} \) have the same joint distribution by proving that \( \{Z_{n-1}\} \) and \( \{Z'_{n-1}(Z_0)\} \), given \( \{Y_n, n \geq 1\} \), have the same joint distribution. Thus, without loss of generality, we can simply assume

\[
\hat{m}_n = \frac{\sum_{i=1}^{n} (Z'_i - Y_i)}{\sum_{i=1}^{n} Z'_{i-1}}
\]
and
\[ N_c = \inf \{ n \geq 1 : S_n \geq c\sigma^2 \}, \]
where \( Z_i' = Z_i'(Z_0) \) and \( S_n = S_n(Z_0) \) with \( Z_0 \) being the initial population size. Since \( S_n \geq \sum_{i=1}^{n-1} Y_i \), we have \( S_n \to \infty \) a.s. by the strong law of large numbers. Thus, \( P(N_c < \infty) = 1 \) and \( N_c \to \infty \) a.s. as \( c \to \infty \). Therefore, \( N_c \) is well defined.

Accordingly, we have
\[
\hat{m}_n - m = \frac{1}{S_n} \sum_{i=1}^{S_n} (\xi_i - m),
\]
and
\[
\hat{m}(c) = \frac{\sum_{j=1}^{N_c-1} (Z_j' - Y_j) + \theta_c(Z_{N_c}' - Y_{N_c})}{c\sigma^2}
\]
with
\[
\theta_c = \frac{c\sigma^2 - \sum_{j=1}^{N_c-1} Z_j'}{Z_{N_c}'}.
\]
Therefore,
\[
\hat{m}(c) - m = \frac{\sum_{i=1}^{S_{N_c} - 1} (\xi_i - m) + \theta_c \sum_{i=S_{N_c} - 1}^{S_{N_c}} (\xi_i - m)}{c\sigma^2}.
\]
It suffices to show that
\[
\frac{\sum_{i=1}^{S_{N_c} - 1} (\xi_i - m) + \theta_c \sum_{i=S_{N_c} - 1}^{S_{N_c}} (\xi_i - m)}{c\sigma^2} \to N(0, 1)
\]
since \( S_{N_c} - 1 = \sum_{j=1}^{N_c-1} Z_j' - 1 \).

Following the notation in the proof of Theorem 3.1, there exist integers \( \{ k(c) \} \) satisfying (3.6) such that (3.7) holds with \( s(c) = c/mk(c) \). Therefore,
\[
\frac{\sum_{i=1}^{S_{N_c} - 1} (\xi_i - m) + \theta_c \sum_{i=S_{N_c} - 1}^{S_{N_c}} (\xi_i - m)}{c\sigma^2} \to N(0, 1)
\]
and
\[
\frac{\sum_{i=1}^{S_{N_c} + k(c) - 1} (\xi_i - m) + \theta_c \sum_{i=S_{N_c} + k(c) - 1}^{S_{N_c} + k(c)} (\xi_i - m)}{c\sigma^2} \to N(0, 1)
\]
have the same asymptotic distribution function.

Let
\[
\theta'_c = \frac{c\sigma^2 - \sum_{j=1}^{k(c) - 1} m_j Z_j'_{N(c)} - 1}{m^{k(c)} Z_j'_{N(c)} - 1}.
\]
In view of (3.2) and (3.4) we have
\[
\sum_{j=1}^{N(c)} Z_j'_{N(c)} \to 0, \quad \frac{Z_j'_{N(c)} - 1}{m^{k(c)} Z_j'_{N(c)} - 1} \to 1 \quad \text{and} \quad \sum_{j=N(c)+1}^{N_c} Z_j' \to 1, \quad \sum_{j=1}^{k(c) - 1} m_j Z_j'_{N(c)} - 1 \to 1,
\]
from which we conclude that
\[ \theta_c - \theta'_c \xrightarrow{p} 0 \quad \text{as} \quad c \to \infty. \]

Note that
\[
S_{N(c)+k(c)} \left( \sum_{i=S_{N(c)+k(c)}-1+1}^{N(c)+k(c)} (\xi_i - m) \right) = \sum_{j=1}^{N(c)+k(c)-1} Z'_{j-1}(\hat{m}_{N_{i(c)}+k(c)} - m) - \sum_{j=1}^{N(c)+k(c)-1} Z'_{j-1}(\hat{m}_{N_{i(c)+k(c)-1}} - m).
\]

From Theorems 2.1 and 3.1, we have
\[
\hat{m}_{N(c)+k(c)} - m = o_p \left( \left( \sum_{j=1}^{N(c)+k(c)-1} Z'_{j-1} \right)^{-1/2} \right)
\]
and
\[
\hat{m}_{N(c)+k(c)-1} - m = o_p \left( \left( \sum_{j=1}^{N(c)+k(c)-1} Z'_{j-1} \right)^{-1/2} \right)
\]
yielding
\[
S_{N(c)+k(c)} \left( \sum_{i=S_{N(c)+k(c)}-1+1}^{N(c)+k(c)} (\xi_i - m) \right) = o_p \left( \left( \sum_{j=1}^{N(c)+k(c)-1} Z'_{j-1} \right)^{-1/2} \right).
\]

And moreover,
\[
S_{N(c)} \left( \sum_{i=1}^{N(c)} (\xi_i - m) \right) = o_p \left( \left( \sum_{j=1}^{N(c)} Z'_{j-1} \right)^{-1/2} \right) = o_p \left( \left( \sum_{j=1}^{N(c)+k(c)-1} Z'_{j-1} \right)^{-1/2} \right).
\]

Therefore, it is concluded from the above arguments that
\[
\sum_{i=1}^{S_{N(c)+k(c)-1}} (\xi_i - m) + \theta_c \sum_{i=S_{N(c)+k(c)}-1+1}^{S_{N(c)+k(c)}} (\xi_i - m) = \sum_{i=S_{N(c)+k(c)}+1}^{S_{N(c)+k(c)-1}} (\xi_i - m) + \theta'_c \sum_{i=S_{N(c)+k(c)}+1}^{S_{N(c)+k(c)-1}} (\xi_i - m) + o_p \left( \left( \sum_{j=1}^{N(c)+k(c)-1} Z'_{j-1} \right)^{-1/2} \right)
\]
and from (3.2) that
\[
0^2 Z'_{N_{i(c)+k(c)-1}} + \sum_{j=1}^{N_{i(c)+k(c)-1}} Z'_{j-1} = (1 + o_p(1)) \left(0^2 m^{k(c)} Z'_{N(c)-1} + \sum_{j=1}^{k(c)-1} m^i Z'_{N(c)-1} \right).
\]
Thus, to finish the proof of the theorem, it suffices to show that
\[
\sum_{i=S_{N(c)}(c)+1}^{S_{N(c)}(c)+k(c)-1} (\xi_i - m) + \theta \sum_{i=S_{N(c)}(c)+1}^{S_{N(c)}(c)+k(c)+1} (\xi_i - m) \quad \overset{d}{\rightarrow} \quad N(0, 1).
\] (4.3)

Let
\[
A_c = \frac{\sum_{i=S_{N(c)}(c)+1}^{S_{N(c)}(c)+k(c)-1} (\xi_i - m)}{\sigma S_{N(c)}(c) - S_{N(c)}(c)}
\]
and
\[
B_c = \frac{\sum_{i=S_{N(c)}(c)+1}^{S_{N(c)}(c)+k(c)+1} (\xi_i - m)}{\sigma Z_{N(c)}(c) - 1}
\]

Denote the sigma-fields \(F_n = \sigma(\xi_j, 1 \leq j \leq n)\) for \(n \geq 1\) and the stopping-sigma-field \(F_N\) for any stopping time \(N\) by
\[
F_N = \{ A : A \cap \{ N = n \} \in F_n \text{ for all } n \geq 1 \}.
\]

We want to show that conditional on \(F_{N(c)}\),
\[
(A_c, B_c) \overset{d}{\rightarrow} (U_1, U_2),
\] (4.4)
where \(U_1\) and \(U_2\) are two independent random variables with standard normal distribution.

We need only to prove (4.4) along any given subsequence. We will consider a given \(\{c_n\}\) with \(c_n \rightarrow \infty\).

Let \(H_n = (R_n, M_n, Q_n)\) be distributed the same as \((S_{N(c_n)}(c_n), N_{N(c_n)}(c_n), Z_{N(c_n)}(c_n))\) and independent of \(\{\xi_j\}\) and \(\{Y_j\}\). Then from Lemma 4.3 in the appendix
\[
A_{c_n} = \frac{\sum_{i=S_{N(c_n)}(c_n)+1}^{S_{N(c_n)}(c_n)+k(c_n)-1} (\xi_i - m)}{\sigma S_{N(c_n)}(c_n) - S_{N(c_n)}(c_n)} \overset{d}{=} \frac{1}{\sigma} \sum_{i=1}^{S_{k(c_n)}(c_n) - 1} (\xi_i - m) + \frac{S_{k(c_n)}(c_n) - 1}{S_{N(c_n)}(c_n)} \overset{d}{\rightarrow} N(0, 1).
\] (4.5)

It is easily seen from (3.4) that as \(n \rightarrow \infty\)
\[
\frac{S_{N(c_n)}(c_n)+k(c_n)-1}{m^{k(c_n)-1} S_{N(c_n)}(c_n)} \rightarrow 1 \text{ a.s.}
\]

Since \(S_{k(c_n)}(c_n) - 1(Q_n)\) and \(S_{N(c_n)}(c_n)+k(c_n)-1 - S_{N(c_n)}(c_n)\) have the same distribution, we have
\[
\frac{S_{k(c_n)}(c_n) - 1(Q_n)}{m^{k(c_n)-1} R_n} \rightarrow 1 \text{ in probability.}
\]

Thus, in view of Lemmas 4.3 and 4.4 in the appendix, conditional on \(F_{N(c_n)}\), \(A_{c_n} \overset{d}{\rightarrow} N(0, 1)\).

We can also prove that conditional on \(F_{N(c_n)}\), \(\{\xi S_{N(c_n)}(c_n) + j, j \geq 1\}\) are independent and identically distributed the same as \(\xi_1\) and that \(\{S_{N(c_n)}(c_n)+k(c_n)-1+j, j \geq 1\}\) and \((S_{N(c_n)}(c_n)+k(c_n)-1, S_{N(c_n)}(c_n), Z_{N(c_n)}(c_n)+k(c_n)-1)\) are conditionally independent. Now note that
\[
A_{c_n} = \frac{\sum_{j=1}^{S_{N(c_n)}(c_n)+k(c_n)-1} (\xi S_{N(c_n)}(c_n)+j - m)}{\sigma S_{N(c_n)}(c_n)+k(c_n)-1 - S_{N(c_n)}(c_n)}
\]
and
\[ B_{cn} = \frac{\sum_{j=1}^{k(cn)} (\xi_{N(cn)} + j - m)}{\sigma \sqrt{\sum_{j=1}^{k(cn)} m j Z'}_{N(cn) - 1}}. \]

Thus, conditional on \( \mathcal{F}_{N(cn)} \), \( A_{cn} \) and \( B_{cn} \) are independent and \( B_{cn} \xrightarrow{d} N(0, 1) \) by applying the central limit theorem. That proves (4.4).

Finally, note that with probability 1,
\[ \frac{Z'_{N(c)} + k(c) - 1}{m k(c) Z'_{N(c) - 1}} \rightarrow 1 \]

and
\[ \frac{S_{N(c)} + k(c) - 1 - S_{N(c)}}{\sum_{j=1}^{k(c) - 1} m j Z'}_{N(c) - 1} \rightarrow 1. \]

If we set
\[ A'_{c} = A_{c} \times \frac{S_{N(c) + k(c) - 1} - S_{N(c)}}{\sqrt{\sum_{j=1}^{k(c) - 1} m j Z'}_{N(c) - 1}} \]

and
\[ B'_{c} = B_{c} \times \frac{Z'_{N(c) + k(c) - 1}}{m k(c) Z'_{N(c) - 1}}, \]

then, given \( \mathcal{F}_{N(c)} \), \( A'_{c} \) and \( B'_{c} \) are asymptotically independent and converge in distribution to \( N(0, 1) \). Moreover, note that
\[ \frac{\sum_{i=S_{N(c)} + k(c) - 1}^{S_{N(c)} + k(c) - 1 + 1} (\xi_{i} - m) + \sum_{i=S_{N(c)} + k(c) - 1 + 1}^{S_{N(c)} + k(c) - 1} (\xi_{i} - m)}{\sigma \sqrt{\sum_{j=1}^{k(c) - 1} m j Z'}_{N(c) - 1}} = \frac{\sum_{j=1}^{k(c) - 1} m j A'_{c}}{\sqrt{\sum_{j=1}^{k(c) - 1} m j Z'}_{N(c) - 1}} + \frac{\sum_{j=1}^{k(c) - 1} m j B'_{c}}{\sigma \sqrt{\sum_{j=1}^{k(c) - 1} m j Z'}_{N(c) - 1}}. \]

Since the coefficients of \( A'_{c} \) and \( B'_{c} \) are \( \mathcal{F}_{N(c)} \)-measurable, (4.3) follows. \( \square \)

**Corollary 4.2.** If \( m \in (0, 1] \), then
\[ \left( \sum_{j=1}^{N_{c}} Z_{j-1} \right)^{1/2} \frac{\hat{m}(c) - m}{\sigma} \xrightarrow{d} N(0, 1) \]

and
\[ c^{-1/2} \left( \frac{\hat{m}(c) - m}{\sigma} \right) \xrightarrow{d} N(0, 1). \]

The corollary follows from Theorem 4.1 with an observation that \( Z_{i} / \sum_{j=1}^{n} Z_{j-1} \rightarrow 0 \) a.s. as \( n \rightarrow \infty \) (see, eg., Sriram et al., 1991).
Appendix

The following two lemmas are from Qi and Reeves (2002).

**Lemma 4.3.** For $c > 0$, $\{\xi_{N_c+j}, j \geq 1\}$ and $(S_{N_c}, Z'_{N_c}, N_c)$ are independent, and moreover, $\{\xi_{N_c+j}, j \geq 1\}$ is a sequence of i.i.d. random variables distributed as $\xi_1$.

**Lemma 4.4.** Assume that $\{\eta_j, j \geq 1\}$ is a sequence of i.i.d. random variables with $E\eta_1 = 0$ and $E\eta_1^2 = 1$ and $\{H_n\}$ is a sequence of random vectors. For each $n \geq 1$, $\{\eta_j\}$ and $H_n$ are independent. If $\{v_n\}$ and $\{u_n\}$ are two sequences of integer-valued random variables such that $v_n$ is measurable with respect to $\sigma(H_n)$ and $u_n$ is the function of $\{\eta_j\}$ and $H_n$, namely $u_n = f_n(H_n, \eta_j, j \geq 1)$, with

$$
\frac{u_n}{v_n} \rightarrow 1 \text{ in probability, (4.6)}
$$

then

$$
P \left( \sum_{j=1}^{u_n} \eta_j \leq x | H_n \right) \rightarrow \Phi(x) \text{ in probability (4.7)}
$$

for any $x \in \mathbb{R}$, where $\Phi(x)$ is the standard normal distribution function.

Acknowledgments

The author would like to thank Professor Richard Green, an associate editor and a referee for their helpful comments. The research was supported in part by NSF grant DMS 0604176.

References


