

Boundary Crossing Probability in the Autoregressive Process†

Kiheon Choi ¹ and Jaechang Lee ²

ABSTRACT

The limiting distribution of the excess over the boundary is determined for the autoregressive process.

KEYWORDS: Local limit theorem, conditional probability, stopping time.

1. INTRODUCTION

The quadratic forms are of importance in many applications of the theory of stochastic processes. The quadratic forms can be transformed to weighted sums of squares of independent identically distributed normal variates. In many applications, these weights are or approximate the eigenvalues of a Toeplitz matrix. Choi(1991) found the asymptotic density of quadratic forms. In Section 2, we study the problem of finding the asymptotic density of a conditional quadratic form. Section 3 is devoted to the problem of finding the asymptotic joint density for linear and quadratic forms. The remainder of the paper uses the results of these sections. In Section 4, finally, we study the problem of the limiting distribution of the excess over the boundary for the sum of the autoregressive process.

¹Department of Statistics, Duksung Womens University, Seoul, 132-714, Korea.

²Department of Statistics, Korea University, Seoul, 136-701, Korea.

† This research was supported by the Korea Science Foundation.

2. ASYMPTOTIC DENSITY OF A CONDITIONAL QUADRATIC FORM

Consider the first order autoregressive model

$$X_n = \theta + \rho X_{n-1} + \epsilon_n, \quad n \in \mathcal{Z} \quad (1)$$

where ϵ_n , $n \in \mathcal{Z}$, are i.i.d. normally distributed random variables with mean 0 and variance 1. Here X_0 is a constant and $\rho \in (-1, 1)$. The stochastic process scheme is often assumed to correspond to a stationary process. Let

$$S_n = \sum_{i=1}^n X_{i-1} \epsilon_i.$$

Then

$$\begin{aligned} S_n &= \sum_{k=1}^n \left[X_0 \rho^{k-1} + \theta \left(\frac{1 - \rho^{k-1}}{1 - \rho} \right) \right] \epsilon_k + \sum_{k=1}^n \sum_{j=1}^{k-1} \rho^{k-j-1} \epsilon_j \epsilon_k \\ &= \mathbf{b}' \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix} + \frac{1}{2} (\epsilon_1, \dots, \epsilon_n) \mathbf{B}_n \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}, \end{aligned}$$

where \mathbf{B}_n is the same as in Choi(1991) and the \mathbf{b} is a vector for which

$$b_j = X_0 \rho^{j-1} + \theta \left(\frac{1 - \rho^{j-1}}{1 - \rho} \right), \quad \forall j = 1, \dots, n.$$

To find the conditional distribution of S_n given $\epsilon_{n-m+1}, \dots, \epsilon_n$ write S_n in the form

$$\begin{aligned} S_n &= S_{n-m} + X_0 \sum_{k=n-m+1}^n \rho^{k-1} \epsilon_k + \sum_{k=n-m+1}^n \left(\frac{1 - \rho^{k-1}}{1 - \rho} \right) \theta \epsilon_k \\ &\quad + \sum_{k=n-m+1}^n \sum_{j=1}^{n-m} \rho^{k-j-1} \epsilon_j \epsilon_k + \sum_{k=n-m+1}^n \sum_{j=n-m+1}^{k-1} \rho^{k-j-1} \epsilon_j \epsilon_k \\ &= S_{n-m} + X_0 \sum_{k=n-m+1}^n \rho^{k-1} \epsilon_k + \sum_{k=n-m+1}^n \left(\frac{1 - \rho^{k-1}}{1 - \rho} \right) \theta \epsilon_k \\ &\quad + \sum_{j=1}^{n-m} \sum_{k=n-m+1}^n \rho^{k-j-1} \epsilon_k \epsilon_j + \sum_{k=n-m+1}^n \sum_{j=n-m+1}^{k-1} \rho^{k-j-1} \epsilon_j \epsilon_k \\ &= \mathbf{A}' \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{n-m} \end{pmatrix} + \frac{1}{2} (\epsilon_1, \dots, \epsilon_{n-m}) \mathbf{B}_{n-m} \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{n-m} \end{pmatrix} + C \end{aligned}$$

where

$$A_j = X_0 \rho^{j-1} + \frac{1 - \rho^{j-1}}{1 - \rho} \theta + \sum_{k=n-m+1}^n \rho^{k-j-1} \epsilon_k, \quad \forall j = 1, \dots, n - m$$

$$C = \sum_{k=n-m+1}^n \sum_{j=n-m+1}^{k-1} \rho^{k-j-1} \epsilon_j \epsilon_k + X_0 \sum_{k=n-m+1}^n \rho^{k-1} \epsilon_k + \sum_{k=n-m+1}^n \left(\frac{1 - \rho^{k-1}}{1 - \rho} \right) \theta \epsilon_k.$$

Let f_n denote the density of S_n and $f_{n,m}$ denote the conditional density of S_n given $\epsilon_{n-m+1}, \dots, \epsilon_n$. Then

$$f_{n,m}(s | \epsilon_{n-m+1}, \dots, \epsilon_n) = f_{n-m}(s - C; \rho, \mathbf{A})$$

for all s and $\epsilon_{n-m+1}, \dots, \epsilon_n$, where \mathbf{A} and C are as above (function of $\epsilon_{n-m+1}, \dots, \epsilon_n$).

Theorem 2.1. Let $\tilde{h}_{n,m}(\cdot | s)$ denote a conditional density of $\epsilon_{n-m+1}, \dots, \epsilon_n$, given $S_n = s$ and let $h_m(\cdot)$ denote the density of $\epsilon_1, \dots, \epsilon_m$. If $c > 0$, then

$$\lim_{n \rightarrow \infty} \tilde{h}_{n,m}(t_1, \dots, t_m | s) = h_m(t_1, \dots, t_m),$$

for all t_1, \dots, t_m , uniformly for all s such that

$$\left| \frac{s - \mu_n}{\sigma_n} \right| \leq c.$$

Where μ_n and σ_n is the means and variances of S_n .

Proof. Let f_n denote the density of S_n and let $f_{n,m}$ denote the conditional density of S_n , given $\epsilon_{n-m+1}, \dots, \epsilon_n$. Then, for fixed t_1, \dots, t_m ,

$$\tilde{h}_{n,m}(t_1, \dots, t_m | s) = \frac{f_{n,m}(s | t_1, \dots, t_m)}{f_n(s)} h_m(t_1, \dots, t_m).$$

From Choi(1991), $f_n(s)$ and $f_{n,m}(s | t_1, \dots, t_m)$ may be approximated by normal densities with means

$$\mu_n = \frac{1}{2} \sum_{k=1}^n \lambda_{nk} = \frac{1}{2} \text{tr}(\mathbf{B}_n) = 0$$

$$\mu'_n = \frac{1}{2} \sum_{k=1}^{n-m} \lambda_{(n-m)k} + C = \frac{1}{2} \text{tr}(\mathbf{B}_{n-m}) + C = C$$

and variances

$$\sigma_n^2 = \sum_{k=1}^n \beta_k^2 + \frac{1}{2} \sum_{k=1}^n \lambda_{nk}^2$$

$$\sigma_n'^2 = \sum_{k=1}^{n-m} \alpha_k^2 + \frac{1}{2} \sum_{k=1}^{n-m} \lambda_{(n-m)k}^2,$$

respectively, where α_k , $k = 1, 2, \dots, n - m$, are as in Choi(1991) with \mathbf{b} replaced by \mathbf{A} . Now, by Theorem 3.1 in Choi(1991) applied to both the numerator and denominator,

$$\begin{aligned} \frac{f_{n,m}(s \mid t_1, \dots, t_m)}{f_n(s)} &= \frac{\frac{1}{\sigma_n' \sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{s-\mu_n'}{\sigma_n'}\right)^2\right\} + o\left(\frac{1}{\sigma_n'}\right)}{\frac{1}{\sigma_n \sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{s-\mu_n}{\sigma_n}\right)^2\right\} + o\left(\frac{1}{\sigma_n}\right)} \\ &= \left| \frac{\sigma_n}{\sigma_n'} \right| \exp\left\{-\frac{1}{2}\left(\frac{s-\mu_n'}{\sigma_n'}\right)^2 + \frac{1}{2}\left(\frac{s-\mu_n}{\sigma_n}\right)^2\right\} (1 + o(1)) \\ &= \left| \frac{\sigma_n}{\sigma_n'} \right| \exp\left\{\frac{1}{2}\left(\frac{s-\mu_n}{\sigma_n}\right)^2 \left[1 - \frac{\sigma_n^2}{\sigma_n'^2}\right] \right. \\ &\quad \left. - \frac{\sigma_n^2}{\sigma_n'^2} \left(\frac{s-\mu_n}{\sigma_n}\right) \left(\frac{\mu_n - \mu_n'}{\sigma_n}\right) - \frac{1}{2}\left(\frac{\mu_n - \mu_n'}{\sigma_n}\right)^2 \frac{\sigma_n^2}{\sigma_n'^2}\right\} (1 + o(1)). \end{aligned}$$

Since \mathbf{C} is an orthogonal matrix,

$$\sigma_n^2 = \sum_{j=1}^n b_j^2 + \frac{1}{2} \sum_{k=1}^n \lambda_{nk}^2.$$

Here

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n b_j^2 = \left(\frac{\theta}{1-\rho}\right)^2$$

and

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \lambda_{nk}^2 = \frac{2}{(1-\rho^2)^3} \{3\rho^3 - 3\rho^2 + 1\}.$$

So

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sigma_n^2 = \sigma^2 \in (0, \infty) \quad (\text{say})$$

Similarly,

$$\sigma_n'^2 = \sum_{j=1}^{n-m} A_j^2 + \frac{1}{2} \sum_{k=1}^{n-m} \lambda_{(n-m)k}^2$$

and

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{n-m} \lambda_{(n-m)k}^2 = \frac{2}{(1-\rho^2)^3} \{3\rho^3 - 3\rho^2 + 1\}.$$

Let $a_j = A_j$ for $j = 1, 2, \dots, n-m$ and $a_j = 0$ for $j = n-m+1, \dots, n$.
Then

$$\begin{aligned} \left| \frac{\sigma_n'^2}{\sigma_n^2} - 1 \right| &\leq \frac{1}{\sigma_n^2} \left| \|\mathbf{a}\|^2 - \|\mathbf{b}\|^2 \right| + o(1) \\ &\leq \frac{\|\mathbf{a}\| + \|\mathbf{b}\|}{\sigma_n^2} \cdot \|\mathbf{a} - \mathbf{b}\| + o(1). \end{aligned}$$

Now $\|\mathbf{a}\| + \|\mathbf{b}\| = O(\sqrt{n})$; and $\|\mathbf{b} - \mathbf{a}\| = O(1)$. So

$$\lim_{n \rightarrow \infty} \left| \frac{\sigma_n'^2}{\sigma_n^2} - 1 \right| = 0.$$

Next

$$\mu'_n - \mu_n = C = C(t_1, \dots, t_n) = O(1).$$

So

$$\lim_{n \rightarrow \infty} \frac{\mu'_n - \mu_n}{\sigma_n} = 0.$$

The theorem follows easily.

Theorem 2.2. Let $0 < \delta < 1$ and $c > 0$, then there is a constant $B = B(\delta, c)$ for which

$$\tilde{h}_{n,m}(t_1, \dots, t_m | s) \leq B h_m(t_1, \dots, t_m),$$

for all t_1, \dots, t_m , for all $m < n\delta$, and for all s such that

$$|s - \mu_n| \leq c\sqrt{n}.$$

Proof. Let f_n denote the density of S_n and let $f_{n,m}$ denote the conditional density of S_n , given $\epsilon_{n-m+1}, \dots, \epsilon_n$. Then, for fixed t_1, \dots, t_m ,

$$\tilde{h}_{n,m}(t_1, \dots, t_m | s) = \frac{f_{n,m}(s | t_1, \dots, t_m)}{f_n(s)} h_m(t_1, \dots, t_m).$$

From Choi(1991), $f_n(s)$ and $f_{n,m}(s | t_1, \dots, t_m)$ may be approximated by

normal densities with means

$$\begin{aligned}\mu_n &= \frac{1}{2} \sum_{k=1}^n \lambda_{nk} = \frac{1}{2} \operatorname{tr}(\mathbf{B}_n) = 0 \\ \mu'_n &= \frac{1}{2} \sum_{k=1}^{n-m} \lambda_{(n-m)k} + C = \frac{1}{2} \operatorname{tr}(\mathbf{B}_{n-m}) + C = C\end{aligned}$$

and variances

$$\begin{aligned}\sigma_n^2 &= \sum_{k=1}^n \beta_k^2 + \frac{1}{2} \sum_{k=1}^n \lambda_{nk}^2 \\ \sigma'_n{}^2 &= \sum_{k=1}^{n-m} \alpha_k^2 + \frac{1}{2} \sum_{k=1}^{n-m} \lambda_{(n-m)k}^2,\end{aligned}$$

respectively, as in the proof of Theorem 2.1. So

$$f_{n,m}(s|t_1, \dots, t_m) \leq \frac{1}{\sqrt{2\pi} \sigma'_n} (1 + o(1))$$

and

$$f_n(s) \geq \frac{1}{\sqrt{2\pi} \sigma_n} e^{-\frac{1}{2}c^2n/\sigma_n^2} (1 + o(1)).$$

Thus

$$\frac{f_{n,m}(s|t_1, \dots, t_m)}{f_n(s)} \leq \frac{\sigma'_n}{\sigma_n} e^{\frac{1}{2}c^2n/\sigma_n^2} (1 + o(1)),$$

which is bounded under the conditions.

Theorem 2.3. Let $h_{n,m}(\cdot|s)$ denote a conditional density of $\epsilon_1, \dots, \epsilon_m$, given $S_n = s$ and let $0 < \delta < 1$ and $c > 0$, then there is a constant $C = C(\delta, c)$ for which

$$h_{n,m}(t_1, \dots, t_m|s) \leq C h_m(t_1, \dots, t_m),$$

for all t_1, \dots, t_m , for all $m < n\delta$, and for all s such that

$$|s - \mu_n| \leq c\sqrt{n}.$$

Proof. Let f_n denote the density of S_n and let $f_{n,m}^*$ denote the conditional density of S_n , given $\epsilon_1, \dots, \epsilon_m$. Then, for fixed t_1, \dots, t_m ,

$$h_{n,m}(t_1, \dots, t_m | s) = \frac{f_{n,m}^*(s | t_1, \dots, t_m)}{f_n(s)} h_m(t_1, \dots, t_m).$$

From Choi(1991), $f_n(s)$ and $f_{n,m}^*(s | t_1, \dots, t_m)$ may be approximated by normal densities with means

$$\begin{aligned} \mu_n &= \frac{1}{2} \sum_{k=1}^n \lambda_{nk} = \frac{1}{2} \text{tr}(\mathbf{B}_n) = 0 \\ \mu_n^* &= \frac{1}{2} \sum_{k=m}^n \lambda_{(n-m)k} + C' = \frac{1}{2} \text{tr}(\mathbf{B}_{n-m}) + C' = C' \end{aligned}$$

and variances

$$\begin{aligned} \sigma_n^2 &= \sum_{k=1}^n \beta_k^2 + \frac{1}{2} \sum_{k=1}^n \lambda_{nk}^2 \\ \sigma_n^{*2} &= \sum_{k=m}^n \gamma_k^2 + \frac{1}{2} \sum_{k=m}^n \lambda_{(n-m)k}^2 \end{aligned}$$

respectively, as in the proof of Theorem 2.1. So

$$f_{n,m}^*(s|t_1, \dots, t_m) \leq \frac{1}{\sqrt{2\pi} \sigma_n^*} (1 + o(1))$$

and

$$f_n(s) \geq \frac{1}{\sqrt{2\pi} \sigma_n} e^{-\frac{1}{2}c^2n/\sigma_n^2} (1 + o(1)).$$

Thus

$$\frac{f_{n,m}^*(s|t_1, \dots, t_m)}{f_n(s)} \leq \frac{\sigma_n^*}{\sigma_n} e^{\frac{1}{2}c^2n/\sigma_n^2} (1 + o(1)),$$

which is bounded under the conditions.

3. ASYMPTOTIC JOINT DENSITY FOR LINEAR AND QUADRATIC FORMS

Observe that

$$\begin{aligned} X_{n-m} &= \rho^{n-m} X_0 + \sum_{j=1}^{n-m} \rho^{n-m-j} \epsilon_j \\ &= c + \mathbf{a}'\epsilon, \quad (\text{say}) \end{aligned}$$

and

$$S_n = \mathbf{b}'\epsilon + \frac{1}{2} \epsilon' \mathbf{B}_n \epsilon,$$

where $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ are i.i.d. standard normal random variables and \mathbf{B}_n and $\mathbf{b} = \mathbf{b}_n$ are defined in Section 2.

Lemma 3.1. The joint characteristic function of X_{n-m} and S_n is given by

$$\begin{aligned} \varphi_n(s, t) &:= E[\exp(isX_{n-m} + itS_n)] \\ &= |\mathbf{I} - it\mathbf{B}_n|^{-\frac{1}{2}} \exp\left\{ics - \frac{1}{2}(\mathbf{a}s + \mathbf{b}t)'(\mathbf{I} - it\mathbf{B}_n)^{-1}(\mathbf{a}s + \mathbf{b}t)\right\}. \end{aligned}$$

Proof. We introduce, as in the proof of Lemma 3.2 in Choi(1991), the orthogonal matrix \mathbf{C} for which $\mathbf{C}'\mathbf{D}\mathbf{C} = \mathbf{B}_n$. Then

$$\begin{aligned} |\mathbf{I} - it\mathbf{B}_n| &= |\mathbf{I} - it\mathbf{D}| \\ &= (1 - it\lambda_{n1}) \cdots (1 - it\lambda_{nn}) \end{aligned}$$

where the $\lambda_{n1}, \dots, \lambda_{nn}$ are the eigenvalues of the matrix \mathbf{B}_n . Let $\mathbf{Z} = \mathbf{C}\epsilon$, $\alpha = \mathbf{C}\mathbf{a}$ and $\beta = \mathbf{C}\mathbf{b}$. Then it is easily seen that Z_1, \dots, Z_n are independent normal random variables. Moreover,

$$\begin{aligned} (\mathbf{a}s + \mathbf{b}t)'(\mathbf{I} - it\mathbf{B}_n)^{-1}(\mathbf{a}s + \mathbf{b}t) &= (\mathbf{a}s + \mathbf{b}t)'[\mathbf{C}'(\mathbf{I} - it\mathbf{D})^{-1}\mathbf{C}](\mathbf{a}s + \mathbf{b}t) \\ &= (\alpha s + \beta t)'(\mathbf{I} - it\mathbf{D})^{-1}(\alpha s + \beta t). \end{aligned}$$

so that

$$(\mathbf{a}s + \mathbf{b}t)'(\mathbf{I} - it\mathbf{B}_n)^{-1}(\mathbf{a}s + \mathbf{b}t) = \sum_{k=1}^n \frac{(\alpha_k s + \beta_k t)^2}{1 - it\lambda_{nk}}.$$

Now

$$\begin{aligned} E\{e^{isX_{n-m} + itS_n}\} &= E\left\{\exp\left[ics + \sum_{k=1}^n [i(s\alpha_k + t\beta_k)Z_k + \frac{1}{2}it\lambda_{nk}Z_k^2]\right]\right\} \\ &= e^{ics} \prod_{k=1}^n E\left\{\exp\left[i(s\alpha_k + t\beta_k)Z_k + \frac{1}{2}it\lambda_{nk}Z_k^2\right]\right\}. \end{aligned}$$

Hence, by using Lemma 3.1 in Choi(1991),

$$\begin{aligned} \varphi_n(s, t) &= \left[\prod_{k=1}^n (1 - it\lambda_{nk})\right]^{-\frac{1}{2}} \exp\left\{ics - \frac{1}{2} \sum_{k=1}^n \frac{(s\alpha_k + t\beta_k)^2}{1 - it\lambda_{nk}}\right\} \\ &= |\mathbf{I} - it\mathbf{B}_n|^{-\frac{1}{2}} \exp\left\{ics - \frac{1}{2}(\mathbf{a}s + \mathbf{b}t)'(\mathbf{I} - it\mathbf{B}_n)^{-1}(\mathbf{a}s + \mathbf{b}t)\right\}. \end{aligned}$$

Theorem 3.1. Let $f_n(x|y)$ denote a conditional density of X_{n-m} , given $S_n = y$ and let $f(x)$ denote the normal density. If $c > 0$, then

$$\lim_{n \rightarrow \infty} f_n(x|y) = f(x), \quad \forall x,$$

provided

$$|y - \mu_n| \leq c\sqrt{n}.$$

Proof. The characteristic function of X_{n-m} and S_n/\sqrt{n} is $\varphi_n(s, t/\sqrt{n})$. As in the proof of Theorem 3.1 in Choi(1991), there is no loss of generality in supposing that $M \geq 1$ and that $(n + \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2)/\sigma_n^2 \leq M$ for all n . Let $\delta = 1/8M^3 \leq 1/8M$. If $|t| \leq \delta\sqrt{n}$, then

$$\begin{aligned} \log \varphi_n(s, t/\sqrt{n}) &= -\frac{1}{2} \sum_{k=1}^n \log \left(1 - it \frac{\lambda_{nk}}{\sqrt{n}} \right) + ics - \frac{1}{2} \sum_{k=1}^n \left\{ \frac{(s\alpha_k + t\beta_k/\sqrt{n})^2}{1 - it \frac{\lambda_{nk}}{\sqrt{n}}} \right\} \\ &= \frac{1}{2} \sum_{p=2}^{\infty} \frac{1}{p} \left\{ \sum_{k=1}^n \left(\frac{\lambda_{nk}}{\sqrt{n}} \right)^p \right\} (it)^p + ics - \frac{1}{2} \sum_{p=0}^{\infty} \left\{ \sum_{k=1}^n (s\alpha_k + t\beta_k/\sqrt{n})^2 \left(\frac{\lambda_{nk}}{\sqrt{n}} \right)^p \right\} (it)^p \\ &= ics - \frac{1}{2} c_{1n} s^2 - \frac{1}{2} c_{2n} t^2 + R_n(s, t), \quad -\infty < t < \infty, \end{aligned}$$

where

$$\begin{aligned} c_{1n} &= \sum_{k=1}^n \alpha_k^2 \\ c_{2n} &= \frac{1}{n} \left\{ \sum_{k=1}^n \beta_k^2 + \frac{1}{2} \sum_{k=1}^n \lambda_{nk}^2 \right\}, \end{aligned}$$

and

$$\begin{aligned} |R_n(s, t)| &\leq \frac{1}{2} \sum_{p=3}^{\infty} \frac{nM^p}{(\sqrt{n})^p} |t|^p + \frac{1}{2} \frac{t^2}{n} \|\mathbf{b}\|^2 \sum_{p=1}^{\infty} \left| \frac{Mt}{\sqrt{n}} \right|^p \\ &\quad + \frac{1}{2} s^2 \|\mathbf{a}\|^2 \sum_{p=1}^{\infty} \left| \frac{Mt}{\sqrt{n}} \right|^p + \left(\frac{st}{\sqrt{n}} \right) |\mathbf{a}'\mathbf{b}| \left\{ \sum_{p=0}^{\infty} \left| \frac{Mt}{\sqrt{n}} \right|^p + 1 \right\} \end{aligned}$$

for all $s, |t| \leq \delta\sqrt{n}$ and all $n \geq 1$, by Taylor's Theorem applied to the logarithm. Here R_n tends to zero as n tends to infinity, for all s and t , since

$$\begin{aligned} |R_n| &\leq \frac{1}{2} \frac{n|Mt|^3}{n^{3/2}} \sum_{p=3}^{\infty} \left| \frac{Mt}{\sqrt{n}} \right|^{p-3} + \left\{ \frac{1}{2} \frac{M|t|^3}{n^{3/2}} \|\mathbf{b}\|^2 + \frac{1}{2} \frac{Ms^2|t|}{n^{1/2}} \|\mathbf{a}\|^2 \right. \\ &\quad \left. + \frac{Ms|t|^2}{n} |\mathbf{a}'\mathbf{b}| \right\} \sum_{p=1}^{\infty} \left| \frac{Mt}{\sqrt{n}} \right|^{p-1} + \frac{st}{\sqrt{n}} |\mathbf{a}'\mathbf{b}| \end{aligned}$$

$$\begin{aligned} &\leq \frac{2|Mt|^3}{3\sqrt{n}} + \frac{2M|t|^3}{3n^{3/2}} \|\mathbf{b}\|^2 + \frac{2Ms^2|t|}{3\sqrt{n}} \|\mathbf{a}\|^2 \\ &\quad + \frac{4Ms|t|^2}{3n} |\mathbf{a}'\mathbf{b}| + \frac{4st}{3\sqrt{n}} |\mathbf{a}'\mathbf{b}| \end{aligned}$$

and

$$\begin{aligned} \lim_{n \rightarrow \infty} c_{1n} &= \frac{1}{1-\rho^2} = c_1 \\ \lim_{n \rightarrow \infty} c_{2n} &= \frac{1}{2} \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} g^2(x) dx \right] = \frac{1}{(1-\rho^2)^3} [3\rho^3 - 3\rho^2 + 1] = c_2. \end{aligned}$$

Hence

$$\lim_{n \rightarrow \infty} \varphi_n(s, t/\sqrt{n}) = e^{-\frac{1}{2}c_1s^2 - \frac{1}{2}c_2t^2}, \quad \forall s, t;$$

and, therefore the distribution function of X_{n-m} and S_n converges to the joint normal.

It is clear that X_{n-m} and S_n have a density $f_n(x, y)$ for all $n \geq 1$ and that φ_n is integrable with respect to Lebesgue measure for all $n \geq 3$. So,

$$\begin{aligned} f_n(x, y) &= \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-isx-ity} \varphi_n(s, t/\sqrt{n}) ds dt, \\ f(x, y) &= \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-isx-ity} \{e^{-\frac{1}{2}c_1s^2 - \frac{1}{2}c_2t^2}\} ds dt, \end{aligned}$$

and

$$\begin{aligned} |f_n(x, y) - f(x, y)| &= \left| \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-isx-ity} [\varphi_n(s, t/\sqrt{n}) - e^{-\frac{1}{2}c_1s^2 - \frac{1}{2}c_2t^2}] ds dt \right| \\ &\leq \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\varphi_n(s, t/\sqrt{n}) - e^{-\frac{1}{2}c_1s^2 - \frac{1}{2}c_2t^2}| ds dt \end{aligned}$$

for all $-\infty < x, y < \infty$ and all $n \geq 3$.

If n is sufficiently large and $|t| \leq \delta\sqrt{n}$, then $|R_n| \leq \frac{1}{4}(c_1s^2 + c_2t^2)$ and, therefore

$$\varphi_n(s, t/\sqrt{n}) \leq e^{-\frac{1}{4}(c_1s^2 + c_2t^2)}.$$

So

$$\int_{-\delta\sqrt{n}}^{\delta\sqrt{n}} \int_{-\delta\sqrt{n}}^{\delta\sqrt{n}} |\varphi_n(s, t/\sqrt{n}) - e^{-\frac{1}{2}c_1s^2 - \frac{1}{2}c_2t^2}| ds dt \rightarrow 0$$

by the dominated convergence theorem. It remains to show that

$$\lim_{n \rightarrow \infty} \int_{|t| \geq \delta \sqrt{n}} \int |\varphi_n(s, t/\sqrt{n})| ds dt \rightarrow 0$$

as $n \rightarrow \infty$. Now

$$|\varphi_n(s, t/\sqrt{n})| = \prod_{k=1}^n |1 - it\lambda_{nk}/\sqrt{n}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}As^2 - Bts/\sqrt{n} - \frac{1}{2}Ct^2/n\right\}$$

where

$$A = \sum_{k=1}^n \frac{\alpha_k^2}{1 + t^2\lambda_{nk}^2/n}, \quad B = \sum_{k=1}^n \frac{\alpha_k\beta_k}{1 + t^2\lambda_{nk}^2/n}, \quad C = \sum_{k=1}^n \frac{\beta_k^2}{1 + t^2\lambda_{nk}^2/n}.$$

It follows that

$$\begin{aligned} |\varphi_n(s, t/\sqrt{n})| &= \prod_{k=1}^n |1 + t^2\lambda_{nk}^2/n|^{-\frac{1}{4}} \exp\left\{-\frac{1}{2}As^2 - Bts/\sqrt{n} - \frac{1}{2}Ct^2/n\right\} \\ &\leq \left(\frac{1}{1 + t^2/4n}\right)^{\frac{1}{4}N_n} \exp\left\{-\frac{1}{2}[A(s + Bt/A/\sqrt{n})^2 + (C - B^2/A)(t^2/n)]\right\} \end{aligned}$$

where

$$N_n = \#\left\{k : \lambda_{nk}^2 \geq \frac{1}{4}\right\}$$

and

$$\lim_{n \rightarrow \infty} \frac{1}{n} N_n = \frac{1}{2\pi} \text{meas}\left[x|g(x) \geq \frac{1}{2}\right] > 0.$$

First, for fixed t , we look at that

$$\int |\varphi_n(s, t/\sqrt{n})| ds \leq \left(\frac{1}{1 + t^2/4n}\right)^{\frac{1}{4}N_n} \sqrt{\frac{2\pi}{A}} \exp\left(-\frac{1}{2}(t^2/n)(C - B^2/A)\right).$$

Now, $C - B^2/A$ is nonnegative by the Cauchy-Schwarz inequality to $CA - B^2$ and $A \geq \frac{\sum_{k=1}^n \alpha_k^2}{1 + Mt^2/4n}$. So

$$\begin{aligned} &\int_{|t| \geq \delta \sqrt{n}} \int_{-\infty}^{\infty} |\varphi_n(s, t/\sqrt{n})| ds dt \\ &\leq \sqrt{2\pi} \int_{|t| \geq \delta \sqrt{n}} \left(\frac{1}{1 + t^2/4n}\right)^{\frac{1}{4}N_n} \sqrt{\frac{1 + Mt^2/4n}{\sum_{k=1}^n \alpha_k^2}} dt \\ &\leq \sqrt{2\pi} \left(\frac{1}{1 + \delta^2/4}\right)^{\frac{1}{4}N_n} \frac{1}{\sqrt{\sum_{k=1}^n \alpha_k^2}} \left\{ \int_{|t| \geq \delta \sqrt{n}} \left(\frac{1}{1 + t^2/4n}\right) dt \right\} \end{aligned}$$

$$+ \frac{1}{2\sqrt{n}} \int_{|t| \geq \delta\sqrt{n}} |t| e^{-\frac{1}{2}(\frac{c-B^2/A}{n})t^2} dt \Big\} \\ \longrightarrow 0, \quad \text{as } n \rightarrow \infty.$$

Therefore

$$f_n(x, y) = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-isx-ity} \varphi_n(s, t/\sqrt{n}) ds dt \\ \rightarrow \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-isx-ity} \varphi(s, t) ds dt = f(x, y).$$

Also, by using Theorem 3.1 in Choi(1991), it is clear that S_n has density $f_n(y)$ for all $n \geq 1$ and that $\varphi_n(0, t/\sqrt{n})$ is integrable with respect to Lebeque measure for all $n \geq 3$. So

$$f_n(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ity} \varphi_n(0, t/\sqrt{n}) dt \\ \rightarrow \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ity} \varphi(0, t) dt = f(y)$$

as $n \rightarrow \infty$ and $f_n(y)$ is bounded away from zero when $|y - \mu_n| \leq c\sqrt{n}$. The theorem follows.

4. MAIN RESULTS

Let $\{\epsilon_k : k \in \mathcal{Z}\}$ be any sequence of i.i.d. normal random variables with mean 0 and variance one; let $\{X_k : k \in \mathcal{Z}^+\}$ be a stochastic process satisfying the autoregression equation

$$X_k = \rho X_{k-1} + \epsilon_k, \quad k \in \mathcal{Z}^+$$

where X_0 is arbitrary and $\rho \in (-1, 1)$ is unknown; let

$$S_n = \sum_{k=1}^n X_{k-1} \epsilon_k$$

and for $a > 0$, $\Delta > 0$, let

$$t_a = \inf \{n \geq 1 : S_n + n\Delta > a\}$$

be the stopping time. Next, define

$$R_a = S_{t_a} + t_a\Delta - a.$$

Thus, R_a is the excess of the boundary a at the time which it first crosses a .

The goal is to find the limiting distribution of R_a . The approach is to compute the conditional probability of crossing the boundary a , looking backward along the sequence S_{n-1}, \dots, S_{n-k} . we give approximation for the conditional probabilities

$$\Psi_a(n, r) = \Pr(t_a \geq n | S_n = a + r)$$

where $a, r > 0, n = 1, 2, \dots$. We use these approximations to obtain the density of R_a .

Lemma 4.1. As $a \rightarrow \infty$,

$$a^{-1}t_a \rightarrow \frac{1}{\Delta}, \quad \text{w.p.1.}$$

Proof. Observe first that $t_a \rightarrow \infty$ as $a \rightarrow \infty$. If $\mathcal{F}_n = \sigma(\epsilon_1, \dots, \epsilon_n)$, then X_k is \mathcal{F}_n -measurable for all $k \leq n$ and, therefore, $X_{n-1}\epsilon_n$ is \mathcal{F}_n -measurable for all n . So, since $E(X_{n-1}\epsilon_n | \mathcal{F}_{n-1}) = 0$, $\{S_n = \sum_{k=1}^n X_{k-1}\epsilon_k, \mathcal{F}_n, n \geq 1\}$ is a martingale. The lemma then follows from the law of large numbers for martingales (See, Hall(1980)), since

$$\Delta \leftarrow \frac{S_{t_a-1} + (t_a - 1)\Delta}{t_a} \leq \frac{a}{t_a} \leq \frac{S_{t_a} + t_a\Delta}{t_a} \rightarrow \Delta, \quad \text{w.p.1.}$$

The main results of this paper gives an approximation to the conditional probability that $t_a \geq n$, given that $S_n + n\Delta = a + r$. That is, an approximation to

$$\Phi_a(n, r) = \Pr(t_a \geq n | S_n = a - n\Delta + r),$$

for $a, r > 0$ and $n = 1, 2, \dots$. Now

$$\begin{aligned} \Phi_a(n, r) &= \Pr(t_a \geq n | S_n = a - n\Delta + r) \\ &= \Pr(S_k + k\Delta \leq a, \forall k < n | S_n = a - n\Delta + r) \\ &= \Pr(S_n - S_{n-k} \geq r - k\Delta, \forall k < n | S_n = a - n\Delta + r) \\ &= \Pr(S_n - S_{n-k} \geq r - k\Delta, \forall k \leq \frac{1}{2}m | S_n = a - n\Delta + r) - \gamma \end{aligned}$$

where

$$\begin{aligned} \gamma_a &= \Pr(S_n - S_{n-k} \geq r - k\Delta, \forall k \leq \frac{1}{2}m, \text{ and} \\ &S_n - S_{n-k} < r - k\Delta, \exists k \in (\frac{1}{2}m, n) | S_n = a - n\Delta + r). \end{aligned}$$

Let

$$\Phi(r) = \Pr(S_{-k} \geq r - k\Delta, \forall k \geq 1)$$

where

$$S_{-k} = \sum_{i=-k+1}^0 Y_{i-1} \epsilon_i.$$

and

$$Y_i = \sum_{j=0}^{\infty} \rho^j \epsilon_{i-j}, \quad \forall i \in \mathbf{Z}.$$

Theorem 4.1. If $n = n_a \rightarrow \infty$ as $a \rightarrow \infty$, in such a way that

$$\frac{a - \mu_n}{\sigma_n} = O(1),$$

then

$$\lim \Phi_a(n, r) = \Phi(r),$$

for all $r > 0$.

Proof. First observe that for all $k \leq \frac{1}{2}m$, for large m and n

$$\begin{aligned} S_n - S_{n-k} &= \sum_{j=n-k+1}^n X_{j-1} \epsilon_j \\ &= \sum_{j=n-k}^{n-1} \sum_{i=n-m+1}^j \rho^{j-i} \epsilon_i \epsilon_{j+1} + X_{n-m} \sum_{j=n-k}^{n-1} \rho^{j-(n-m)} \epsilon_{j+1} \\ &= S_{n,m,k} + S'_{n,m,k} \quad (\text{say}). \end{aligned}$$

Note

$$S_{n,m,k} = \sum_{j=m-k}^{m-1} \sum_{i=1}^j \rho^{j-i} \epsilon_{n-m+i} \epsilon_{n-m+j+1}.$$

Letting $j' = j - m + 1$ and $i' = i - m$, we find that

$$\begin{aligned} S_{n,m,k} &= \sum_{j=m-k}^{m-1} \sum_{i=1}^j \rho^{j-i} \epsilon_{n-m+i} \epsilon_{n-m+j+1} \\ &\stackrel{\text{in dist}}{=} \sum_{j=m-k}^{m-1} \sum_{i=1}^j \rho^{j-i} \epsilon_{i-m} \epsilon_{j-m+1} \end{aligned}$$

$$= \sum_{j'=-k+1}^0 \sum_{i'=1-m}^{j'-1} \rho^{j'-i'-1} \epsilon_{i'} \epsilon_{j'} = T_{m,k} \quad (\text{say})$$

Next observe that for all $\epsilon > 0$,

$$\begin{aligned} \Pr(S_n - S_{n-k} \geq r - k\Delta, \forall k \leq \frac{1}{2}m | S_n = a - n\Delta + r) \\ \leq \Pr(S_{n,m,k} \geq r - \epsilon - k\Delta, \forall k \leq \frac{1}{2}m | S_n = a - n\Delta + r) \\ + \Pr(|S'_{n,m,k}| > \epsilon, \exists k \leq \frac{1}{2}m | S_n = a - n\Delta + r) \end{aligned}$$

and

$$\begin{aligned} \Pr(S_n - S_{n-k} \geq r - k\Delta, \forall k \leq \frac{1}{2}m | S_n = a - n\Delta + r) \\ \geq \Pr(S_{n,m,k} \geq r + \epsilon - k\Delta, \forall k \leq \frac{1}{2}m | S_n = a - n\Delta + r) \\ - \Pr(|S'_{n,m,k}| > \epsilon, \exists k \leq \frac{1}{2}m | S_n = a - n\Delta + r). \end{aligned}$$

Now, let $s = r \pm \epsilon$ and

$$B^* = \left\{ \epsilon \in \mathbf{R}^m : \sum_{j=m-k}^{m-1} \sum_{i=1}^j \rho^{j-i} \epsilon_i \epsilon_{j+1} \geq s - k\Delta, \quad \forall k \leq \frac{1}{2}m \right\}.$$

Then

$$\begin{aligned} \Pr(S_{n,m,k} \geq s - k\Delta, \forall k \leq \frac{1}{2}m | S_n = a - n\Delta + r) \\ = \int \cdots \int_{B^*} \tilde{h}_{n,m}(\cdot | S_n = a - n\Delta + r) d\epsilon_{n-m+1}, \dots, d\epsilon_n \\ \rightarrow \int \cdots \int_{B^*} h_m(\epsilon_{-m+1}, \dots, \epsilon_0) d\epsilon_{-m+1}, \dots, d\epsilon_0 \\ = \Pr(T_{m,k} \geq s - k\Delta, \forall k \leq \frac{1}{2}m), \end{aligned}$$

by Theorem 2.1 in Section 2 and Scheffé's Theorem (See, Lehmann(1959), p351).
Moreover

$$\begin{aligned} \Pr(T_{m,k} \geq s - k\Delta, \forall k \leq \frac{1}{2}m) \\ = \Pr\left(\sum_{j'=-k+1}^0 \sum_{i'=-m+1}^{j'-1} \rho^{j'-i'-1} \epsilon_{i'} \epsilon_{j'} \geq s - k\Delta, \forall k \leq \frac{1}{2}m \right) \\ \rightarrow \Pr\left(\sum_{j=-k+1}^0 Y_{j-1} \epsilon_j \geq s - k\Delta, \forall k \geq 1 \right) \end{aligned}$$

$$= \Pr(T_k \geq s - k\Delta, \forall k \geq 1) \quad \text{as } m \rightarrow \infty$$

where $T_k = S_{-k}$.

On the other hand

$$\begin{aligned} S'_{n,m,k} &= X_{n-m} \sum_{j=n-k}^{n-1} \rho^{j-(n-m)} \epsilon_{j+1} \\ &= \rho^{\frac{1}{2}m} X_{n-m} \sum_{j=n-k}^{n-1} \rho^{j-(n-\frac{1}{2}m)} \epsilon_{j+1}, \quad \text{for } k \leq \frac{1}{2}m. \end{aligned}$$

So

$$\begin{aligned} &\Pr(|S'_{n,m,k}| \geq \varepsilon, \exists k \leq \frac{1}{2}m \mid S_n = a - n\Delta + r) \\ &\leq \Pr(|X_{n-m}| \geq \sqrt{\frac{\varepsilon}{\rho^{\frac{1}{2}m}}} \mid S_n = a - n\Delta + r) \\ &\quad + \Pr\left(\left| \sum_{j=n-k}^{n-1} \rho^{j-(n-\frac{1}{2}m)} \epsilon_{j+1} \right| > \sqrt{\frac{\varepsilon}{\rho^{\frac{1}{2}m}}} \mid S_n = a - n\Delta + r\right) \\ &= (I) + (II), \quad (\text{say}). \end{aligned}$$

Using Theorem 3.1 in Section 3 and Scheffé's Theorem, it may be shown that the first term (I) tends to zero as $m \rightarrow \infty$ for fixed $\varepsilon > 0$. To see how observe that

$$\int |f_n(x \mid y) - f(x)| dx \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

So

$$\lim_{c \rightarrow \infty} \sup_n \int_{|x| > c} f_n(x \mid y) dx = 0.$$

Therefore

$$\lim_{c \rightarrow \infty} \sup_n \Pr\{|X_{n-m}| > c \mid S_n = a - n\Delta + r\} = 0.$$

Similarly, by using Theorem 2.2 in Section 2 and Scheffé's Theorem. The second term (II) tends to zero.

Finally, we show that

$$\gamma \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

Now, for any δ , $0 < \delta < \frac{1}{2}$,

$$\begin{aligned}
 \gamma &= \Pr(S_n - S_{n-k} \geq r - k\Delta, \forall k \leq \frac{1}{2}m, S_n - S_{n-k} < r - k\Delta, \\
 &\quad \exists k \in (\frac{1}{2}m, n) | S_n = a - n\Delta + r) \\
 &\leq \Pr(S_n - S_{n-k} < r - k\Delta, \exists k \in (\frac{1}{2}m, n) | S_n = a - n\Delta + r) \\
 &\leq \Pr(S_n - S_{n-k} < r - k\Delta, \exists k \in (\frac{1}{2}m, n\delta] | S_n = a - n\Delta + r) \\
 &\quad + \Pr(S_n - S_{n-k} < r - k\Delta, \exists k \in (n\delta, n) | S_n = a - n\Delta + r) \\
 &= E_{a,1}(m, \delta) + E_{a,2}(\delta). \quad (\text{say})
 \end{aligned}$$

The second of these terms is bounded. Indeed, replacing $n - k$ by k , we find that

$$\begin{aligned}
 E_{a,2}(\delta) &= \Pr(S_n - S_{n-k} < r - k\Delta, \exists k \in (n\delta, n) | S_n = a - n\Delta + r) \\
 &= \Pr(S_k > a - k\Delta, \exists k \in (0, n(1 - \delta)) | S_n = a - n\Delta + r) \\
 &\leq C \Pr(S_k > a - k\Delta, \exists k \in (0, n(1 - \delta))) \\
 &= C \Pr(t_a \leq n(1 - \delta)). \quad (*)
 \end{aligned}$$

by Theorem 2.3 in Section 2; and last terms in (*) approaches zero as $n \rightarrow \infty$, since by Lemma 4.1 and $n/a \rightarrow 1/\Delta$.

Let

$$D' = \left\{ \epsilon \in R^m : \sum_{j=n-k+1}^n \sum_{i=1}^{j-1} \rho^{j-i-1} \epsilon_i \epsilon_j < r - k\Delta, \exists k \in (\frac{1}{2}m, n\delta] \right\}.$$

Then

$$\begin{aligned}
 E_{a,1}(m, \delta) &= \Pr(S_n - S_{n-k} < r - k\Delta, \exists k \in (\frac{1}{2}m, n\delta] | S_n = a - n\Delta + r) \\
 &= \int_{D'} \tilde{h}_{n,m}(t|s) dt \\
 &\leq B \int_{D'} h_m(t) dt \\
 &\leq B \Pr(\sum_{i=1}^k X_{n-i} \epsilon_{n-i+1} < r - k\Delta, \exists k \in (\frac{1}{2}m, n\delta]) \\
 &\leq B \Pr(\sum_{i=1}^k X_{n-i} \epsilon_{n-i+1} < r - k\Delta, \exists k \in (\frac{1}{2}m, \infty)),
 \end{aligned}$$

by using Theorem 2.2 in Section 2 and tends to zero as $m \rightarrow \infty$.

The theorem now follows by letting $a \rightarrow \infty$, $m \rightarrow \infty$, and $\epsilon \rightarrow 0$, in that order.

Theorem 4.2. Let $t_a^* = (t_a - a/\Delta)/\sqrt{a/\Delta}$. Then t_a^* and R_a are asymptotically independent as $a \rightarrow \infty$; the limiting distribution of t_a^* is normal with mean 0 and variance σ_n^2/Δ^2 ; and the limiting distribution of R_a has density

$$h^*(r) = \frac{1}{\Delta} \Phi(r), \quad r > 0.$$

Proof. The proof of theorem follows from Theorem 2 in Woodroffe(1982). Let g_a denote the joint density of t_a and R_a ; that is

$$g_a(n, r) = \frac{d}{dr} \Pr\{t_a = n, R_a \leq r\}.$$

Then

$$\begin{aligned} g_a(n, r) &= \frac{1}{\sigma_n} \phi\left(\frac{a + r - n\Delta}{\sigma_n}\right) \Phi_a(n, r) \\ &\sim \frac{1}{\sigma_n} \phi\left(\frac{a - n\Delta}{\sigma_n}\right) \Phi(r) \end{aligned}$$

under the conditions of the theorem. The theorem follows easily.

REFERENCES

- (1) Choi, K. (1989). Local limit theorem for the distribution of S_n . *Sequential Analysis*, Vol 8, 395-403.
- (2) Choi, K. (1991). Asymptotic density of quadratic forms. *The Korean Journal of Applied Statistics*, Vol. 4, No. 2, 149-156.
- (3) Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton.
- (4) Graybill, F.A. (1983). *Matrices with Applications in Statistics*. 2nd ed., Wadsworth .
- (5) Grenander, U., Szegö, G. (1958). *Toeplitz Forms and their applications*. University of California Press.
- (6) Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application*. Academic Press, New York.
- (7) Lehmann, E. (1959). *Testing Statistical Hypotheses*. Wiley, New York.
- (8) Woodroffe, M. (1982). Non linear renewal theory in sequential analysis. *Soc. Indust. Appl. Math.*, Philadelphia.