Journal of the Korean
Statistical Society
Vol. 21, No. 2, 1992

Extreme Value of Moving Average Processes with Negative Binomial Noise Distribution

You-Sung Park[1]

ABSTRACT

In this paper, we investigate the limiting distribution of $M_n = \max(X_1, X_2, \cdots, X_n)$ in the infinite moving average processes $\{X_t = \sum c_i Z_{t-i}\}$ generated from i.i.d. negative binomial variables Z_i's. While no limit result is possible, nonetheless asymptotic bounds are derived. We also present the tail behavior of X_t, i.e., weighted sum of i.i.d. random variables. This continues a study made by Rootzen(1986) for discrete innovation sequences.

1. INTRODUCTION

A general form of a stationary moving average processes $\{X_t\}$ can be expressed by

$$X_t = \sum_{i=-\infty}^{\infty} c_i Z_{t-i}, \quad t = 0, \pm 1, \pm 2, \cdots \tag{1.1}$$

where Z_t is an i.i.d. sequence and the c_i is appropriately decreasing to zero to guarantee convergence in (1.1). Such processes $\{X_t\}$ of corse include the ARMA processes used in time series analysis. The extremal behavior of $\{X_t\}$ depends on both the weight $\{c_i\}$ and the tail of the marginal d.f. of the noise variables $\{Z_t\}$

[1] Institute of Statistics, Korea University, Seoul, 136-701, Korea.

in an intricate and interesting way. For example, in the case of the Z_t having a distribution with slowly decaying to zero, it happens typically that one large innovation is responsible for a large X_t value. The X_t value becomes a partial maximum when that large innovation also.

Let $M_n = \max(X_1, \cdots, X_n)$ where $\{X_i\}$ are generated in (1.1). David and Resnick(1985) are concerned with noise variables which have a regularly varying tail

$$P(Z_0 > x) = x^{-q} L(x)$$

where $q > 0$ and L is slowly varying at infinity. With assumption that all c_i are nonnegative for convenience and letting $c_+ = \max\{c_i\}$, they proved that for $x > 0$

$$P\{a_n^{-1}(M_n - b_n) \leq x\} \xrightarrow{d} \exp(-x^{-q}) \text{ as } n \to \infty \tag{1.2}$$

with $a_n = \gamma_n c_+$ and $b_n = 0$ such that

$$P(Z_0 < \gamma_n) \leq 1 - n^{-1} \leq P(Z_0 \leq \gamma_n).$$

An intuitive interpretation of (1.2) is that the maximum M_n asymptotically is achieved when the largest Z_t is multiplied by the largest weihgt c_i, i.e., the limiting behavior of M_n can be investigated by pretending $M_n = \max(c_+ Z_1, \cdots, c_+ Z_n)$.

The other class of moving averages, which has been studied by Rootzen(1986) is specified by

$$P(Z_0 > x) \sim k x^q e^{-x^p} \text{ as } x \to \infty$$

where $k, p > 0$ and q is constant.

He showed that the normalized M_n with the center constant b_n of order $(\log n)^{1/p}$ and the scale constant a_n of order $(\log n)^{1/p-1}$ converges to $\exp(-e^{-x})$. This implies that for $p > 1$ the distribution of M_n becomes more and more concentrated as $n \to \infty$, and that it becomes spread out for $0 < p < 1$.

2. DEFINITIONS AND PRELIMINARY RESULTS

Let F belong to the class G of distribution functions whose support consists of all sufficiently large positive integers.

Definition 2.1. Define a continuous distribution function F_c associated with each $F \in G$. Namely, define for any real number x

$$h_c(x) = h([x]) + (x - [x])(h([x+1]) - h([x])) \tag{2.1}$$

where $[x]$ is the largest integer not exceeding x and $h(n) = -\log(1 - F(n))$ for integers n.

Then set $F_c(x) = 1 - \exp\{-h_c(x)\}$. F_c has the properties that $F_c(x) = F(n)$ for integer n and $F(x) \leq F_c(x) \leq F(x+1)$. Furthermore, F_c is strictly increasing and for all sufficiently large n there exists an unique β_n such that

$$1 - F_c(\beta_n) = n^{-1} \tag{2.2}$$

Let $\{X_t\}$ be a sequence of stationary processes with the marginal distribution function F and let

$$F_{1,2,\ldots,k}(u_n) = P(X_1 \leq u_n, X_2 \leq u_n, \cdots, X_k \leq u_n)$$

where u_n is a sequence of constants such that $u_n \to \infty$ as $n \to \infty$. Now, define, so called, mixing conditions as follows:

Definition 2.2. The condition $D(u_n)$ is said to hold if for any integers $1 \leq i_1 < \cdots < i_p < j_1 < \cdots < j_q \leq n$ for which $j_1 - i_p \geq m$, we have

$$|F_{i_1,\cdots,i_p,j_1,\cdots,j_p}(u_n) - F_{i_1,\cdots,i_p}(u_n)F_{j_1,\cdots,j_p}(u_n)| \leq \alpha_{m(n)}$$

where $\alpha_{m(n)} \to 0$ as $n \to \infty$ for some sequence $m(n) = o(n)$.

Defition 2.3. The condition $D'(u_n)$ is said to hold for the stationary sequence $\{X_n\}$ if

$$\lim_{n\to\infty} \sup n \sum_{i=2}^{[n/k]} P(X_1 > u_n, X_i > u_n) \to 0 \text{ as } k \to \infty.$$

The following two lemmas are useful for moving average processes with tail behavior (1.3), which essentially shows the mixing conditions $D(u_n)$ and $D'(u_n)$ conditions with $u_n = a_n x + b_n$.

Lemma 2.4 (Rootzen). Suppose that the moving average processes $\{X_t = \sum c_i Z_{t-i}\}$ with $c_i = O(|i|^{-\eta})$ for some $\eta > 1, a_n^{-1} = O((\log n)^a)$ for some a and

$E(Z^2) < \infty$, then for each $\epsilon, \delta > 0$

$$nP(a_n^{-1}|\sum_{n\delta}^{\infty} c_i Z_i| > \epsilon) \to 0 \text{ and}$$
$$nP(a_n|\sum_{-\infty}^{-n\delta} c_i Z_i > \epsilon) \to 0 \text{ as } n \to \infty.$$

Lemma 2.5 (Rootzen). Suppose that for some constant $\gamma \in (0, 1]$, and writing $n' = [n^\gamma]$, it holds for $u_n = a_n x + b_n$ for any x that

$$n\sum_{t=1}^{2n'} P(X_0 + X_t > 2u_n) \to 0 \text{ as } n \to \infty \tag{2.3}$$
$$n^2 P(a_n^{-1}|\sum_{n'+1}^{\infty} c_i Z_i| > 1) \to 0,$$
$$n^2 P(a_n^{-1}|\sum_{-\infty}^{-n'-1} c_i Z_i| > \epsilon) \to 0 \text{ as } n \to \infty \tag{2.4}$$

and that

$$P(\sum_{-\infty}^{n'} c_i Z_i > u_n) = O(n^{-1}),\ P(\sum_{-n'}^{\infty} c_i Z_i > u_n) = O(n^{-1}). \tag{2.5}$$

Then $D'(u_n)$ holds.

In this paper, the moving average processes defined in (1.1) with a discrete innovation distribution will be considered. As easily seen, for example, when $c_o = 1, c_i = 0$ for $i \neq 0$ then X_t is obviously discrete. This makes difficulty to derive asymptotic distribution of M_n as indicated in McCormick and Park(1992). We first investigate the asymptotic tail behaviour of $X_0 = \sum c_i Z_i$ and establish a limiting stability for maxima generated by the moving average processes with negative binomial innovation distribution.

3. NEGATIVE BINOMIAL MOVING AVERAGE

Let $\{X_t = \sum c_i Z_{t-i}\}, t = 0, \pm 1, \pm 2, \cdots$, be a moving average processes with $\{c_i\}$ given constants and with the noise sequence $\{Z_t\}$ consisting of i.i.d. random variables whose probability mass function is given as

$$P(Z_i = z) = \binom{\beta + z - 1}{z}(1-\theta)^\beta \theta^z, \quad z = 0, 1, 2, \cdots, \beta \geq 1, 0 < \theta < 1. \tag{3.1}$$

We assume that all c_i are nonnegative and that $c_i = O(i^d)$ for some $d > 1$. Define a continuous d.f. F_c associated with (3.1) as in Definition 2.1 so that

$$1 - F(x+1) \leq 1 - F_c(x) \leq 1 - F(x), \; x \geq 0. \tag{3.2}$$

where $F(\cdot)$ is the distribution function of (3.1). Note that since

$$P(Z_i > z)/P(Z_i = z) \to \frac{\theta}{1-\theta} \text{ as } z \to \infty \text{ and } x! \sim (\frac{x}{e})^x (2\pi x)^{1/2},$$

$$1 - F_c(z) = P_c(Z_i > z) \sim k(\beta + z - 1)^{\beta - 1} e^{z \log \theta} \text{ as } z \to \infty \tag{3.3}$$

where $k = \theta(1-\theta)^{\beta-1}/(\beta-1)!$ and $A(z) \sim B(z)$ implies $A(z)/B(z) \to 1$ as $z \to \infty$.

Proposition 3.1. Suppose that a random variable X is stochastically larger than a random variable Y, i.e., $P(X > x) \geq P(Y > x)$. Then, if X_i are independent and Y_i are independent,

$$P[\sum_{-n}^{n} c_i X_i > z] \geq P[\sum_{-n}^{n} c_i Y_i > z] \text{ for } c_i \geq 0, \; i = 1, 2, \cdots, n.$$

In addition, if X_i are nonnegative r.v., then

$$P[\sum_{-\infty}^{\infty} c_i X_i > z] \geq P[\sum_{-\infty}^{\infty} c_i Y_i > z].$$

Proof. Let $F(\cdot), G(\cdot)$ be distribution functions of X and Y, respectively. Then, we have

$$P[\sum_{i=-n}^{n} c_i X_i < z] = \int \cdots \int F[c_o^{-1}(z - \sum_{\substack{i=-n \\ i \neq 0}}^{n} c_i x_i)] dF^{n-1}(x)$$

$$\leq \int \cdots \int G[c_o^{-1}(z - \sum_{\substack{i=-n \\ i \neq 0}}^{n} c_i x_i)] dG^{n-1}(x) = P[\sum_{i=-n}^{n} c_i Y_i < z]$$

where $dF^{n-1}(x) = dF(x_2) \cdot dF(x_3) \cdots dF(x_n)$.
Now, if X_i are nonnegative, then

$$P[\sum_{-\infty}^{\infty} c_i X_i > z] \geq P[\sum_{i=-n}^{n} c_i Y_i > z] \text{ for any fixed } n.$$

Taking $n \to \infty$, we have

$$P[\sum_{-\infty}^{\infty} c_i X_i > z] \geq P[\sum_{-\infty}^{\infty} c_i Y_i > z].$$

The following two propositions will be used for the tail behavior of $\sum c_i Z_i$ and lead to boundaries of extreme value for negative binomial noise distribution.

Proposition 3.2. Suppose that the random variable Y_1 satisfies

$$P(Y_1 > z) \sim k^m \Gamma^m(\beta)\Gamma^{-1}(m\beta)(-\log\theta)^{m-1}(z+\beta-1)^{m\beta-1}\exp(z\log\theta)$$
$$\text{for } m \geq 1 \tag{3.4}$$

as $z \to \infty$ where $k = \theta(1-\theta)^{\beta-1}/(\beta-1)!$ and Y_1 is independent of a random variable Y_2 which satisfies

$$E(\exp(\alpha Y_2)) < \infty \text{ for some } \alpha > -\log\theta.$$

Then

$$(Y_1 + Y_2 > z)$$
$$\sim k^m \Gamma^m(\beta)\Gamma^{-1}(m\beta)(-\log\theta)^{m-1}(z+\beta-1)^{m\beta-1}\exp(z\log\theta)E(\exp(-Y_2\log\theta)). \tag{3.5}$$

Proof. The proof follows immediately from the same argument as Lemma 7.1 (i) in Rootzen(1986) by choosing γ to be $(-\log\theta)/\alpha < \gamma < 1$.

Remark. The relation (3.5) is uniform in $Y_2 \in \{Y : E(\exp(\alpha Y) \leq c\}$ for fixed $Y_1, c > 0$ and $\alpha > -\log\theta$.

Proposition 3.3. Suppose that independent random variables $Y_1, \cdots, Y_m$ are according to (3.3) with the same parameter θ and its p.d.f $f(z) \sim -(\log\theta)k\exp\{(\beta - 1)\log(\beta + z - 1) + z\log\theta\}$ as $z \to \infty$. Then

$$P(\sum_{i=1}^{m} Y_i > z) \sim k^m \Gamma^m(\beta)\Gamma^{-1}(m\beta)(-\log\theta)^{m-1}(z+\beta-1)^{m\beta-1}\exp(z\log\theta)$$
$$\text{as } z \to \infty. \tag{3.6}$$

Proof. It is easily seen that

$$P(Y_1+Y_2>z)=\int_0^z P(Y_1>z-x)dF_2(x)+P(Y_2>z).$$

Hence, by assumptions and letting

$h(i,\beta,z,\theta)=k^i(\beta+z-1)^{i\beta-1}e^{z\log\theta}$, we have

$$P(Y_1+Y_2>z)=h(1,\beta,z,\theta)\int_0^z \frac{P(Y_1>z-x)}{h(1,\beta,z,\theta)}dF_2(x)+O(z^{\beta-1}e^{z\log\theta})$$

$$=h(1,\beta,z,\theta)\int_0^z(1-\frac{x}{z+\beta-1})^{\beta-1}e^{x\log\theta}dF_2(x)+O(z^{\beta-1}e^{z\log\theta-1})$$

$$\sim k\log\theta^{-1}h(1,\beta,z,\theta)\int_0^z(1-\frac{x}{z+\beta-1})^{\beta-1}(x+\beta-1)^{\beta-1}dx$$

$$=k^2\log\theta^{-1}(z+\beta-1)^{\beta}e^{z\log\theta}\int_0^{z(z+\beta-1)^{-1}}(1-y)^{\beta-1}(y(z+\beta-1)+\beta-1)^{\beta-1}dy$$

$$\sim\log\theta^{-1}h(2,\beta,z,\theta)\int_0^1(1-y)^{\beta-1}y^{\beta-1}dy$$

$$=k^2\Gamma^2(\beta)\Gamma^{-1}(2\beta)(-\log\theta)(z+\beta-1)^{2\beta-1}\exp(z\log\theta). \tag{3.7}$$

Now, let $Y=Y_1+Y_2$. Then we have by (3.7) and the same argument above that

$$P(Y+Y_3>z)=P(Y_1+Y_2+Y_3>z)\sim(\log\theta^{-1})^2h(3,\beta,z,\theta)\Gamma^3(\beta)\Gamma^{-1}(3\beta).$$

Repeating this procedure to evaluate that the tail of $\sum_{i=1}^{m}Y_i$ yields (3.6).

Lemma 3.4. Suppose that independent random variables $\{Z_t\}$ satisfy (3.1). Then, for sufficiently large z,

$$k^*(z/c_+ +\beta-1)^{m\beta-1}e^{z\log\theta/c_+}E\exp\{\sum_{i\notin\Lambda}c_iZ_i(-\log\theta/c_+)\}$$

$$\le P\{\sum c_iZ_i>z\}\le$$

$$k^*\{(z-\sum c_i)/c_+ +\beta-1)^{m\beta-1}e^{(z-\sum c_i)\log\theta/c_+}E\exp\{\sum_{i\notin\Lambda}c_iZ_i(-\log\theta/c_+)\}$$

where $k^*=\{\theta(1-\theta)^{\beta-1}/(\beta-1)!\}^m\Gamma^m(\beta)\Gamma^{-1}(m\beta)(-\log\theta)^{m-1}$, $c_+=\max_i\{c_i\}$, $\Lambda=\{i|c_i=c_+\}$ and m is the size of Λ.

Proof. Let $P_c(\cdot)$ be a probability measure corresponding to the d.f. F_c in (3.3).

We may without loss of generality assume that $c_+ = 1$ since

$$P_c(\sum c_i Z_i > z) = P_c(\sum c_i Z_i / c_+ > z/c_+).$$

Let $\underline{c} = \max\{c_i | i \notin \Lambda\} < 1$. Then, by Lemma 7.2 of Rootzen(1986), one can easily show that

$$E \exp\{\alpha \sum_{i \notin \Lambda} c_i Z_i\} < \infty \text{ for any } \alpha < (-\log\theta)/\underline{c}$$

with $\Psi(s) = E(e^{sZ})$ for $0 \le s < -\log\theta$. Now, assume for a moment that

$$F_c'(x) = f_c(x) \sim -(\log\theta) k(\beta + x - 1)^{\beta-1} e^{x\log\theta} \text{ as } x \to \infty. \tag{3.8}$$

Then, applying proposition 3.2 and 3.3, it follows that for $\alpha \in (-\log\theta, (-\log\theta)/\underline{c})$

$$P_c(\sum c_i Z_i > z) \sim k^*(z/c_+ + \beta - 1)^{m\beta-1} e^{z\log\theta/c_+} E\exp\{\sum_{i\notin\Lambda} c_i Z_i(-\log\theta/c_+)\}.$$

The claim then follows since

$$P_c(\sum c_i Z_i > z) \le P(\sum c_i Z_i > z) \le P_c(\sum c_i Z_i > z - \sum c_i) \tag{3.9}$$

by inequality (3.2) and Proposition 3.1.

To end this proof, we must show (3.8). By definition F_c and h_c in (2.1) and (2.2), we have $1 - F_c(x) = \exp(-h_c(x))$ so that

$$f_c(x) = h_c'(x)\exp(-h_c(x)).$$

It suffices to show that $h_c'(x) \sim -\log\theta$ since, by (3.3),

$$\exp(-h_c(x)) \sim k(\beta + x - 1)^{\beta-1} e^{x\log\theta} \text{ as } x \to \infty.$$

By definition of $h_c(x)$,

$$h_c'(x) = \lim_{t\to 0} \frac{h_c(x+t) - h_c(x)}{t}$$

$$= \lim_{t\to\infty} t^{-1}\{h([x+t]) - h([x]) + (x+t-[x+t])(h([x+t+1]) - h([x+t]))$$

$$-(x-[x])(h([x+1]) - h([x])\}$$

$$= h([x+1]) - h([x])$$
$$= -\log(\frac{1-F([x+1])}{1-F([x])})$$
$$\sim -\log(\frac{\theta(1-\theta)^{-1}P(Z_t=[x]+1)}{\theta(1-\theta)^{-1}P(Z_t=[x])})$$
$$\sim -\log\theta \text{ as } x \to \infty.$$

This completes the proof.

Theorem 3.5. Suppose that $X_t = \sum c_i Z_{i-t}$ be a random variable for all c_i positive with $c_i = O(i^{-d})$ for some $d > 1$ such that for some $\epsilon > 0, nP(x+\alpha_n - \epsilon < X_t \leq x + \alpha_n + \epsilon) \to 0$ as $n \to \infty$ where $\{Z_i\}$ are i.i.d. negative binomial r.v.'s with parameter θ. Then,

$$\lim_{n\to\infty} \sup P(M_n \leq x + \alpha_n) \leq \exp(-e^{(\log\theta)x/c_+})$$

and

$$\lim_{n\to\infty} \inf P(M_n \leq x + \alpha_n) \geq \exp(-e^{(\log\theta)(x-\sum c_i)/c_+})$$

where $M_n = \max(X_1, \cdots, X_n)$,

$$\alpha_n = \frac{c_+}{-\log n}\{\log n + \log T + (m\beta - 1)\log(\frac{\log n}{-\log\theta} + \beta - 1)\}$$

where $T = \{\frac{\theta(1-\theta)^{\beta-1}}{(\beta-1)!}\}^m \Gamma^m(\beta)\Gamma^{-1}(m\beta)(-\log\theta)^{m-1} E\exp\{\sum_{i\notin\Lambda} c_i Z_i(-\log\theta$
$/c_+)\}$.

Proof. First, assume that the conditions $D(u_n), D'(u_n)$ with $u_n = x + \alpha_n$ hold. Here, α_n is such that

$$n(1 - F_c(\alpha_n)) = nP_c(\sum c_i Z_i > \alpha_n) \to 1 \text{ as } n \to \infty.$$

The result then follows immediately by the inequality (3.9) and the same arguments of Theorem 2.1 in McCormick and Park(1992) since, for any fixed y,

$$\lim_{z\to\infty} \frac{1-F_c(z)}{1-F_c(z+y)} = \lim_{z\to\infty} \frac{T(z/c_+ + \beta - 1)^{m\beta-1} e^{z\log\theta/c_+}}{T((z+y)/c_+ + \beta - 1)^{m\beta-1} e^{(z+y)\log\theta/c_+}}$$

$$= \exp(-y \log \theta / c_+).$$

To show that the condition $D(u_n)$ holds for $u_n = x + \alpha_n$, let $1 \le i_1 < \cdots < i_p < j_1 < \cdots < j_q \le n$ be integers with $j_1 - i_p \ge 2n\nu$ for fixed $\nu > 0$ and $\underline{X}_i = (X_{i_1}, \cdots, X_{i_p}), \underline{X}_j = (X_{j_1}, \cdots, X_{j_p})$. We have by Rootzen(1986) that for any $\epsilon > 0$ and $u_n = x + \alpha_n$

$$\sup_{i,j} |P(\underline{X}_i \le u_n, \underline{X}_j \le u_n) - P(\underline{X}_i \le u_n) P(\underline{X}_j \le u_n)|$$

$$\le nP(u_n - 2\epsilon < X_0 \le u_n + 2\epsilon) + 2nP(|\sum_{n\nu}^{\infty} c_i Z_i| > \epsilon) + 2nP(|\sum_{\infty}^{-n\nu} c_i Z_i| > \epsilon).$$

The first and last two terms tend to zero by assumption and Lemma 2.4, respectively.

To prove that the condition $D'(u_n)$ holds for $u_n = x + \alpha_n$, we will use Lemma 2.5. However, (2.3) and (2.4) can be easily shown by the same arguments of Theorem 7.4 in Rootzen(1986) with $(-\log\theta)/2 < \alpha < (-\log\theta)/c^*$, $c^* = \max_{n>0} \max_i (c_i + c_{i-n}), \gamma > 0$ with $1 + \gamma < 2\alpha(-\log\theta)^{-1}$ and $u_n \sim c_+(-\log\theta)^{-1} \log n$. Finally (2.5) can be done by the same methods as in Lemma 3.4 using uniformity of (3.5) and the inequality (3.9). Thus it follows that

$$P(\sum_{-\infty}^{n'} c_i Z_i > u_n) \le P_c(\sum_{-\infty}^{n'} c_i Z_i > u_n - \sum c_i) \text{ for any fixed } n'$$

and

$$P_c(\sum_{-\infty}^{n'} c_i Z_i > u_n - \sum c_i)$$

$$\sim T\{(u_n - \sum c_i)/c_+ + \beta - 1)^{m\beta - 1} e^{(u_n - \sum c_i)\log\theta / c_+} \text{ as } n \to \infty$$

which, by the choice of $u_n = x + \alpha_n$, proves the first part of (2.5). The second part is the same. Hence the proof follows.

Collorary 3.6. Theorem 3.5 holds without the assumption $nP(x + \alpha_n - \epsilon < X_t \le x + \alpha_n + \epsilon) \to 0$ as $n \to \infty$ for the $2m + 1$ dependent discrete valued process $\{X_t\}$ defined as

$$X_t = \sum_{i=-m}^{m} c_i Z_{t-i}.$$

Proof. The proof follows immediately when we choose sufficiently small $\epsilon > 0$ so that $P(x + \alpha_n - \epsilon < X_t \le x + \alpha_n + \epsilon) \to 0$, which is always possible since Z_t is an integer valued random variables.

REFERENCES

(1) Anderson,C.W.(1970). Extreme value theory for a class of discrete distributions with applications to some stochastic processes. *Journal of Applied Probability.* 7,99-113.

(2) Davis,R. and Resnick,S.I.(1985). Limiting theory for moving averages of random variables with regularly varying tail probabilities. *Annals of Probability.* 13,179-195.

(3) Gnedenko,B.V.(1943). Sur la distribution limit du terme maximum d'une serie aleatoire. *Annals of Mathematics.* 44,423-453.

(4) Leadbetter,M.R.,Lindgren,G. and Roozen,H.(1983). *Extremes and Related Properties of Random Sequences and Processes,* Springer-Verlag, New York.

(5) McCormick,W.P.and Park,Y.S.(1992). Asymptotic analysis of extremes from AR(1) poisson processes. To appear in *Journal of Applied Probability.*

(6) Rootzen,H.(1986). Extreme value theory for moving average processes. *Annals of Probability.* 14,612-652.