Asymptotics of an Extreme-value Estimator for a First-Order Bifurcating Autoregressive Process with an Unknown Location Parameter

Andrew Bartlett

Southern Illinois University Edwardsville, Edwardsville, IL 62025

Abstract

Asymptotics of an alternative extreme-value estimator for the autocorrelation parameter in a first-order bifurcating autoregressive (BAR) process with non-gaussian innovations are derived. This contrasts with traditional estimators whose asymptotic behavior depends on the central part of the innovation distribution. Within any BAR model, the main concern is addressing the complex dependency between generations. The inability of traditional methods to handle this dependency motivated an alternative procedure. Our interest lies in investigating through simulation how significant the dependency between generations really is. Additionally, we will investigate the asymptotic properties of our extreme value estimator associated with the BAR(1) model with non-gaussian innovations. Finally, the implications of our extreme value approach are discussed with an extensive simulation study that not only asses the reliability of our proposed estimate but presents the findings for a new estimator of an unknown location parameter θ and its implications.

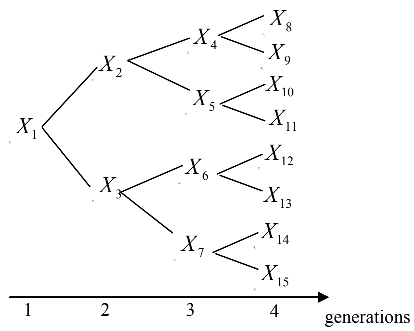
1 Introduction

The bifurcating autoregressive processes (BAR, for short) are widely known for analyzing cell lineage data, where each individual in one generation gives rise to two offspring. In addition to their biological importance, BAR models are essential in developing theoretical methods for the analysis of dependent data. Methods of statistical analysis of cell lineage data, which explicitly considered the dependence between the cells, was first introduced by Cowan and Staudte (1986). The Cowan and Staudte model regards each line of descent as an first-order autoregressive AR(1) process. More precisely, the first-order bifurcating autoregressive process, BAR(1), is defined by the equation

$$X_t = \phi X_{\lfloor t/2 \rfloor} + \epsilon_t, \qquad \text{for } 2 \leq t \leq 2^k - 1, \tag{1.1}$$

where k represents the number of generations, $|\phi| < 1$, and $\lfloor \cdot \rfloor$ denotes the greatest integer function, so that one can write recursively $X_2 = \phi X_1 + \epsilon_2, X_3 = \phi X_1 + \epsilon_3, X_4 = \phi X_2 + \epsilon_4$, etc. For illustration, the data structure with four generations is given below.

Figure 1: Illustration with four Generations in a Binary Tree



Over the past two decades there have been many extensions of bifurcating autoregressive models. For example, referring to the variables X_{2t} and X_{2t+1} as the daughters of X_t, in computing their values these daughters receive innovations ϵ_{2t} and ϵ_{2t+1}, respectively. Some BAR models allow for the innovation pair $(\epsilon_{2t}, \epsilon_{2t+1})$ to be dependent while other models require an iid structure for the innovations. In this paper we consider the latter situation, but examine through simulation how our estimation procedure performs with dependent innovations. The motivation from a biological rationale for exploring the situation when the innovation pair $(\epsilon_{2t}, \epsilon_{2t+1})$ are dependent is that the sister cells grow in a similar environment, particularly early in their lives, and hence one expects the correlation between sisters to at least exist. However, other more distant relatives, for example cousins, share less of their environment and it seems reasonable to suppose that their environmental effects are independent.

Whether the innovation pair is assumed to be independent or dependent, a common theme in almost all BAR models is the use of traditional methods to obtain estimators whose asymptotic behavior depends on the central part of the innovation distribution. For example, Huggins and Basawa (1999, 2000) proposed a bifurcating autoregressive moving average or BARMA (p, q) to account for this dependence and studied maximum likelihood estimation for a Gaussian BAR model of order p. Zhou and Basawa (2005b) introduced non-Gaussian BAR(1) with Exponential type innovations and established the exact and asymptotic distributions of a maximum likeihood estimator.

While the papers mentioned above made some remarkable breakthroughs, there were some limitations to their approach. For example, the inability to handle the complex dependency between generations, an issue to our knowledge that has yet to be resolved. With this being said, our interest lies in providing some insight through simulation how significant the dependency between generations really is. In conjunction, we will investigate the asymptotic properties of our extreme value estimator associated with the BAR(1) model with non-gaussian innovations.

A main advantage of our estimation procedure is that it relies heavily upon the large innovations, and because of this, we were able to investigate the complex dependency found between generations. The inability of traditional methods to handle this dependency motivated our alternative approach. Furthermore, we are able to consider the infinite variance case $0 < \beta < 2$, while papers mentioned above exclusively considered the finite variance case. Within this setup, we propose an estimate for the correlation parameter ϕ and a unknown location parameter θ under a stationary model where we assume the innovation sequence $(\epsilon_{2t}, \epsilon_{2t+1}), t \geq 1$, is a sequence of independently and identically distributed positive bivariate random variables. We then derive the limiting laws for our estimators $(\hat{\phi}, \hat{\theta})$ through the use of point processes and regular variation that removes the complexity and difficulty presented in Theorem 2 of Zhou and Basawa (2005b) under the specified bivariate exponential innovation distribution. Here the process is extended, first by considering alternative estimates for ϕ rather than the typical least-square or maximum likelihood estimate such that innovations $\{Z_t\}$ follow an non-gaussian distribution F and secondly assuming F to be regularly varying at both endpoints θ and infinity. The rest of the paper is organized as follows: asymptotic limit results for the autocorrelation parameter ϕ and location parameter θ are presented in Section 2, Section 3 contains an extensive simulation study that assess the reliability and performance of our proposed estimates and Section 4 verifies that our extreme value theory method produces the same limit law as Theorem 2 in Zhou and Basawa (2005b), Additionally, Section 5 presents detailed proofs for some of our main results found in Section 2 under iid assumption on the innovations.

2 Asymptotic Results

Suppose $\{X_t\}$ is a sequence of observations of some characteristic of individual t. Beginning with the positive starting value X_1, the first order bifurcating autoregressive process, BAR(1), with a positive left endpoint θ is defined as,

$$X_t = \phi X_{\lfloor t/2 \rfloor} + \epsilon_t, \quad \text{for } 2 \leq t \leq 2^{(k+1)} - 1, \tag{2.1}$$

where $0 < \phi < 1$ and $\lfloor x \rfloor$ denotes the largest integer less than or equal to x. We assume that the innovations $\{\epsilon_t\}$ are such that the offspring $(\epsilon_{2t}, \epsilon_{2t+1}), t \geq 1$ are a sequence of independently and identically distributed nonnegative bivariate random vectors with $(\epsilon_{2t}, \epsilon_{2t+1}) \sim F_1$. Additionally, $\{\epsilon_t\}$ is assumed to have the same marginal distribution F such that $\theta = \inf\{x : F(x) > 0\}$ and $\sup\{x : F(x) < 1\} = \infty$. Lastly, we assume F is regularly varying with index α at its positive left endpoint θ, abbreviated $F \in RV_\alpha$, and $\bar{F} = 1 - F$ is regularly varying with index $-\beta$ at infinity, its right endpoint. That is, there exists $\beta > 0$ such that

$$\lim_{t \to \infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = x^{-\beta}, \text{ for all } x > 0. \tag{2.2}$$

In this paper we will consider the situation where k generations evolved and the offspring from that generation are included. That is,

$$1 + 2 + 2^2 + \ldots + 2^k = 2^{k+1} - 1.$$

In order to notationally clarify the transition from the k^{th} generation to the t^{th} individual in the BAR(1) process $\{X_t\}$, let

$$n = 2^{k+1} - 1.$$

That is, the total number of individuals in this process including the offspring during the k^{th} generation is n. Now by defining

$$t = \begin{cases} 2\tilde{t} & \text{if } t \text{ is even (male);} \\ 2\tilde{t} + 1 & \text{if } t \text{ is odd (female),} \end{cases}$$

observe that we can now express the process in (2.1) from the offspring perspective,

$$\begin{cases} X_{2\tilde{t}} = \phi X_{\tilde{t}} + \epsilon_{2\tilde{t}}, & \text{for } 1 \le \tilde{t} \le \lfloor n/2 \rfloor; \\ X_{2\tilde{t}+1} = \phi X_{\tilde{t}} + \epsilon_{2\tilde{t}+1}, & \text{for } 1 \le \tilde{t} \le \lfloor n/2 \rfloor. \end{cases} \tag{2.3}$$

The process in (2.3) allows for a better interpretation of the bifurcating autoregressive process, where the model parameter ϕ represents the strength of correlation between the mother and its offspring under the assumption that the variance is finite. Now denoting $x \wedge y = \min(x, y)$, we have

$$Y_{\tilde{t}} = \epsilon_{2\tilde{t}} \wedge \epsilon_{2\tilde{t}+1} = X_{2\tilde{t}} \wedge X_{2\tilde{t}+1} - \phi X_{\tilde{t}}.$$

Then $Y_{\tilde{t}} \sim G(x) = 2F(x) - F_1(x, x)$, with $G(\theta) = 0$. Finally, the first-order autoregressive process, AR(1), is

$$X_t^* = \phi X_{t-1}^* + \epsilon_t^*, \tag{2.4}$$

where $\{\epsilon_t^*\}$ is a sequence of i.i.d. random variables with the same marginal distribution as $\{\epsilon_t\}$.

By assuming (2.2), we are considering a time series with heavy-tailed errors and within certain time series applications a better model is achieved. Thus, our goal is to capitalize on the behavior of extreme value estimators over traditional estimators when $0 < \beta < 2$. This contrasts with estimators whose asymptotic behavior depends on the central part of the innovation distribution when a second or higher moment is finite. Since the estimate for θ depends on the estimate for the autocorrelation coefficient, we begin studying the asymptotic properties of $\hat{\phi}_n$ and then move onto asymptotic properties for $\hat{\theta}_n$. The motivation for the natural estimator $\hat{\phi}_n$, comes from the observation when $X_{\lfloor t/2 \rfloor}$ is large, equation (2.1) implies

$$0 \le \phi \le X_t / X_{\lfloor t/2 \rfloor}. \tag{2.5}$$

Therefore, by minimizing the ratio in (2.5) we expect

$$\hat{\phi}_n = \bigwedge_{t=2}^{n} \frac{X_t}{X_{\lfloor t/2 \rfloor}}$$

to be a reasonable estimator for ϕ. We now turn to showing that $b_n(\hat{\phi}_n - \phi)$ converges in distribution where

$$b_n = F^{\leftarrow}(1 - \frac{1}{n}) := \inf\{x : F(x) \ge (1 - 1/n)\}.$$

Let $\bar{F}(x) = 1 - F(x)$. Then we have

$$\lim_{n \to \infty} \frac{\bar{F}(b_n x)}{\bar{F}(b_n)} = x^{-\beta}, \text{ for all } x > 0. \tag{2.6}$$

Now define $a_n = G^{\leftarrow}(1/n) - \theta$. Then

$$\lim_{n \to \infty} \frac{G(\theta + a_n y)}{G(\theta + a_n)} = y^{\alpha}, \text{ for all } y > 0. \tag{2.7}$$

First observe that the stationary solution is

$$X_t = \epsilon_t + \phi\epsilon_{\lfloor t/2\rfloor} + \phi^2\epsilon_{\lfloor t/2^2\rfloor} + \ldots + \phi^m\epsilon_{\lfloor t/2^m\rfloor} + \ldots.$$

Thus we begin with a truncation of X_t's by defining

$$X_t^{(m)} = \sum_{j=0}^{m} \phi^j \epsilon_{\lfloor t/2^j\rfloor}, \quad m \geq 1, t \geq 2^m$$

as an approximation to X_t. Furthermore,

$$X_1^{*(m)} = \sum_{j=0}^{m} \phi^j \epsilon_{2^j} \stackrel{d}{=} X_t^{(m)}, \quad m \geq 1 \quad X_1^* = X_1^{*(\infty)}. \tag{2.8}$$

Now using the fact that $Y_{\tilde{t}} = \epsilon_{2\tilde{t}} \wedge \epsilon_{2\tilde{t}+1}$, we can determine the necessary point process, since

$$\begin{aligned} P[b_n(\hat{\phi}_n - \phi) > x] &= P\left[\bigwedge_{t=2}^{n}\left(\frac{X_t - \phi X_{\lfloor t/2\rfloor}}{b_n^{-1}X_{\lfloor t/2\rfloor}}\right) > x\right] \\ &= P\left[\bigwedge_{t=2}^{n}\left(\frac{\epsilon_t}{b_n^{-1}X_{\lfloor t/2\rfloor}}\right) > x\right] \\ &= P\left[\bigwedge_{\tilde{t}=1}^{\lfloor n/2\rfloor}\left(\frac{\epsilon_{2\tilde{t}} \wedge \epsilon_{2\tilde{t}+1}}{b_n^{-1}X_{\tilde{t}}}\right) > x\right] \\ &= P\left[\bigwedge_{\tilde{t}=1}^{\lfloor n/2\rfloor}\left(\frac{Y_{\tilde{t}}}{b_n^{-1}X_{\tilde{t}}}\right) > x\right]. \end{aligned}$$

Thus we define the following point process:

$$\mathcal{I}_n = \sum_{\tilde{t}=1}^{\lfloor n/2\rfloor} \varepsilon_{(Y_{\tilde{t}}, b_n^{-1}X_{\tilde{t}})} \quad \text{and} \quad \mathcal{I}_n^{(m)} = \sum_{\tilde{t}=1}^{\lfloor n/2\rfloor} \varepsilon_{(Y_{\tilde{t}}, b_n^{-1}X_{\tilde{t}}^{(m)})}.$$

Observe that the point process $\mathcal{I}_n$ consists of two independent components, where the first component consists of the marks for the minimum of the offspring individuals $Y_{\tilde{t}}$ and the second component consists of the points from the parents $b_n^{-1}X_{\tilde{t}}$. Since we are looking at the first order bifurcating process from the natural perspective of (2.3), we will let $t = \tilde{t}$, so that $1 \leq t \leq \lfloor n/2\rfloor$.

Now we consider establishing convergence of the point process $\mathcal{I}_n^{(m)}$ by first defining rectangles

$$R_i = [a_i, b_i] \times [a_i^{'}, b_i^{'}], 1 \leq i \leq q. \tag{2.9}$$

Thus, we need to show for any $q \geq 1$ that the $q-$dimensional distribution converges. That is,

$$(\mathcal{I}_n^{(m)}(R_1), \ldots, \mathcal{I}_n^{(m)}(R_q)) \stackrel{d}{\to} Pois(\lambda_1) \times \ldots \times Pois(\lambda_q) \quad \text{as } n \to \infty, \tag{2.10}$$

where $\lambda_i \equiv \lambda_i^{(m)} = \lim_{n\to\infty} E[\mathcal{I}_n^{(m)}(R_i)]$ and $Pois(\lambda)$ denotes a Poisson distribution with parameter λ, while $X \times Y$ means that X and Y are independent.

We will prove (2.10) for the case of $q = 2$. That is, we will show that if $a_1^{'} < b_1^{'} < a_2^{'} < b_2^{'}$ so that $R_1 \cap R_2 = \emptyset$, then

$$(\mathcal{I}_n^{(m)}(R_1), \mathcal{I}_n^{(m)}(R_2)) \Rightarrow Pois(\lambda_1) \times Pois(\lambda_2) \quad \text{as } n \to \infty.$$

The proof of the other cases is similar, thus omitted. Now suppose that we have constructed arbitrary blocks $Q_{l,s}$ in such a way that $X_t^{(m)}$ and $X_{t'}^{(m)}$ are independent for all t if $t \in Q_{l,s}$ and $t' \in Q_{l',s'}$ such that $(l, s) \neq (l', s')$. Under this assumption. we proceed with the argument and define the following indicator function's

$$I_{(l,s)}^i = \begin{cases} 1 & \text{if } \sum_{t \in Q_{l,s}} \mathbb{1}[(Y_t, b_n^{-1} X_t^{(m)}) \in R_i] \geq 1 \text{ for } i = 1, 2, \\ 0 & \text{otherwise }, \end{cases} \tag{2.11}$$

and

$$\tilde{I}_{(l,s)}^i = \begin{cases} 1 & \text{if } \sum_{t \in Q_{l,s}} \mathbb{1}[(Y_t, b_n^{-1} X_t^{(m)}) \in R_i] \geq 2 \text{ for } i = 1, 2, \\ 0 & \text{otherwise }. \end{cases} \tag{2.12}$$

Now let $I_{(l,s)} = (I_{(l,s)}^1, I_{(l,s)}^2)$ and $\tilde{I}_{(l,s)} = (\tilde{I}_{(l,s)}^1, \tilde{I}_{(l,s)}^2)$.
Now observe that $\{I_{(l,s)} : 1 \leq s \leq \lfloor 2^l/r \rfloor, l = l_0, \ldots, k\}$ are i.i.d. Bernoulli random vectors. The following lemma shows if there exists at least two different indices in the same interval $Q_{l,s}$ such that the points in $\mathcal{I}_n^{(m)}$ fall within either region R_1, R_2 or both for some (l, s) then the event is negligible as n tends to infinity.
First observe since $X_t^{(m)}$ and $X_{t'}^{(m)}$ are independent if $t \in Q_{l,s}$ and $t' \in Q_{l',s'}$ with $(l, s) \neq (l', s')$ then $\sum_{t \in Q_{l,s}} \mathbb{1}[(Y_t, b_n^{-1} X_t^{(m)}) \in R_i]$ and $\sum_{t' \in Q_{l',s'}} \mathbb{1}[(Y_{t'}, b_n^{-1} X_{t'}^{(m)}) \in R_i]$ are independent for $i = 1, 2$.

Lemma 2.1. *Suppose $t \in Q_{l,s}$ and $t' \in Q_{l',s'}$ such that $(l, s) \neq (l', s')$. Under the assumptions that $0 < \phi < 1, \theta > 0, \bar{F} \in RV_{-\beta}$ we have*

$$P\left[\sum_{l=l_0}^{k} \sum_{s=1}^{\lfloor 2^l/r \rfloor} \tilde{I}_{(l,s)} \neq (0, 0)\right] \to 0, \ \ as\ k \to \infty.$$

Remark 1. By definition as k tends to infinity $n = 2^{k+1} - 1$ tends to infinity and therefore we can speak of regular variation in terms of k.
The next lemma shows the probability that there exists at least one index in the same interval $Q_{l,s}$ for each component of $I_{(l,s)}$, such that the point or points in $\mathcal{I}_n^{(m)}$ fall within region R_i is zero for $i = 1, 2$ and some (l, s).

Lemma 2.2. *Under the conditions specified in Lemma 2.1 for some $l_0 \leq l \leq k, 1 \leq s \leq \lfloor 2^l/r \rfloor$, we have*

$$P[I_{(l,s)} = (1, 1)] = o\left(\frac{r}{2^{k+1}}\right).$$

The following lemma calculates the probability that there exists at least one index in the same interval $Q_{l,s}$, such that the point or points in $\mathcal{I}_n^{(m)}$ fall within exactly one region for some (l, s).

Lemma 2.3. *Let $Y_1 \sim G$. Then under the conditions specified in Lemma 2.1 for $r = o(2^k)$ and some $l_0 \le l \le k, 1 \le s \le \lfloor 2^l/r \rfloor$, we have*

$$\frac{2^{k+1}}{r} P[I_{(l,s)} = (j, 1-j)] \sim \gamma_{2-j}, \quad \text{for } j = 0, 1, \quad \text{and} \quad k \text{ large}$$

where $\gamma_i = (G(b_i) - G(a_i)) \cdot \kappa_i$ and $\kappa_i = (b_i^{'-\beta} - a_i^{'-\beta})(1 - \phi^{(m\beta)})/(1 - \phi^\beta)$ for $i = 1, 2$.

The next lemma determines the limiting distribution for our i.i.d. random vector $I_{(l,s)}$ through the use of the moment generating function.

Lemma 2.4. *Under the conditions specified in Lemma 2.1 for $r = o(2^k)$ we have*

$$P[I_{(l,s)} = (i,j)] = \begin{cases} \frac{r}{2^{k+1}} \gamma_{(i,j)}(1 + o(1)) & \text{if } 0 \le i, j \le 1, i + j = 1 \ , \\ 1 - \frac{r}{2^{k+1}}(\gamma_{(1,0)} + \gamma_{(0,1)})(1 + o(1)) & \text{if } i = j = 0 \ , \end{cases}$$

where $\gamma_{(i,j)} = (G(b_{2-i}) - G(a_{2-i}))(b_{2-i}^{'-\beta} - a_{2-i}^{'-\beta})(1 - \phi^{(m\beta)})/(1 - \phi^\beta)$ for $i + j = 1$. Additionally,

$$\sum_{l=l_0}^{k} \sum_{s=1}^{\lfloor 2^l/r \rfloor} I_{(l,s)} = (I^1_{(l,s)}, I^2_{(l,s)}) \Rightarrow Pois(\lambda_1) \times Pois(\lambda_2), \qquad \text{as } k \to \infty,$$

where $\lambda_i = (G(b_i) - G(a_i))(b_i^{'-\beta} - a_i^{'-\beta})(1 - \phi^{(m\beta)})/(1 - \phi^\beta)$ for $i = 1, 2$.

Applying the previous lemma's yields our first main result.

Theorem 2.1. *Consider the stationary BAR(1) process $\{X_t\}$ from (2.1) where F satisfies (2.2). Then for any $m \ge 1$ and disjoint sets, $R_i := [a_i, b_i] \times [a_i^{'}, b_i^{'}]$ for $i = 1, 2$ we have*

$$(\mathcal{I}_n^{(m)}(R_1), \mathcal{I}_n^{(m)}(R_2)) \Rightarrow Pois(\lambda_1) \times Pois(\lambda_2), \qquad \text{as } n \to \infty,$$

where $\lambda_i \equiv \lambda_i^{(m)} = \lim_{n\to\infty} E[\mathcal{I}_n^{(m)}(R_i)] = (G(b_i) - G(a_i))(b_i^{'-\beta} - a_i^{'-\beta})(1 - \phi^{(m\beta)})/(1 - \phi^\beta)$, for $i = 1, 2$.

Recall that if n is the total number of observations in k generations, we have $n = 2^{k+1} - 1$ or $k = \log_2(n+1) - 1$. The following corollary produces the limiting distribution for $\hat{\phi}_n$ as the number of generations tends to infinity.

Corollary 2.1. *Consider the estimator of ϕ, $\hat{\phi}_n = \bigwedge_{t=2}^n \frac{X_t}{X_{\lfloor t/2 \rfloor}}$. Suppose $0 < \phi < 1, \theta > 0, \bar{F} \in RV_{-\beta}$ and $EY^{-\gamma} < \infty$ for some $\gamma > \beta$, then*

$$\lim_{n\to\infty} P[b_n(\hat{\phi}_n - \phi) > x] = e^{-x^\beta EY^{-\beta}(1-\phi^\beta)^{-1}}, \qquad \text{for all } x > 0,$$

where Y has the stationary distribution G for the process (2.1).

We now shift our attention to the positive unknown location parameter θ. The motivation for an estimator of θ arrives from the observation that, $X_t - \hat{\phi}_n X_{\lfloor t/2 \rfloor}$ can be expressed as $-(\hat{\phi}_n - \phi) X_{\lfloor t/2 \rfloor} + \epsilon_t$. Now since $\bigwedge_{t=2}^n \epsilon_t \xrightarrow{a.s.} \theta$ and $\hat{\phi}_n \xrightarrow{p} \phi$ as $n \to \infty$ allows us to define our estimator for θ:

$$\hat{\theta}_n = \bigwedge_{t \in I_n} (X_t - \hat{\phi}_n X_{\lfloor t/2 \rfloor}), \tag{2.13}$$

where we define the index set $I_n = \{t : 2 \le t \le n \text{ and } X_{\lfloor t/2 \rfloor} \le (a_n b_n)^\rho\}$ where $0 < \rho < 1$ is a fixed value.
Now to determine the limiting distribution for $\hat{\theta}$ observe that,

$$\hat{\theta}_n - \theta = [(\hat{\theta}_n - \bigwedge_{t \in I_n} Y_t) + (\bigwedge_{t \in I_n} Y_t - \theta)]. \tag{2.14}$$

The following lemma show that the first term in (2.14) goes to zero in probability, thus allowing us to focus only on the second term.

Lemma 2.5. *Under the assumptions that G is regularly varying with index α at its positive left endpoint θ and $\bar{F}$ is regularly varying with index $-\beta$ at infinity, its right endpoint, and $\alpha > \beta$, then*

$$a_n^{-1}\Big(\hat{\theta}_n - \bigwedge_{t \in I_n} Y_t\Big) \xrightarrow{p} 0, \text{ as } n \to \infty,$$

where $a_n = G^{\leftarrow}(1/n) - \theta$.

Now we use point processes to show that the second term in (2.14) converges (weakly) to a Poisson point process with mean measure $y^\alpha/2$.

Theorem 2.2. *Consider the stationary BAR(1) process $\{X_t\}$ from (2.1) where G satisfies (2.7). Let $\mathcal{V}_n$ and $\mathcal{V}$ be the point processes on the space $E_2 = [0, \infty)$ defined by*

$$\mathcal{V}_n = \sum_{t=1}^{\lfloor n/2 \rfloor} \varepsilon_{(a_n^{-1}(Y_t - \theta))} \quad \text{and} \quad \mathcal{V} = \sum_{p=1}^{\infty} \varepsilon_{j_p},$$

where $\sum_{p=1}^{\infty} \varepsilon_{j_p}$ is $PRM(\nu)$ with $\nu[0, y] = y^\alpha/2$, $y > 0$. Then in $M_p(E_2)$,

$$\mathcal{V}_n \Rightarrow \mathcal{V}.$$

The following lemma is an alternative approach to show that the point process used in Theorem 4.2.2 suffices.

Lemma 2.6. *Consider the stationary BAR(1) process $\{X_t\}$ from (2.1) where G satisfies (2.7). Let $\tilde{\mathcal{V}}_n$ and $\mathcal{V}_n$ be the point processes on the space $E_3 = [0, \infty) \times [0, \infty)$ defined by*

$$\tilde{\mathcal{V}}_n = \sum_{t=1}^{\lfloor n/2 \rfloor} \varepsilon_{(a_n^{-1}(Y_t - \theta), c_n)} \text{ and } \quad \mathcal{V}_n = \sum_{t=1}^{\lfloor n/2 \rfloor} \varepsilon_{(a_n^{-1}(Y_t - \theta), 0)},$$

where $c_n = a_n^{-1}|(\hat{\phi}_n - \phi)|X_t$. Then for $t \in I_n$

$$d(\tilde{\mathcal{V}}_n, \mathcal{V}_n) \xrightarrow{p} 0,$$

where d is the vague metric on $M_p(E_3)$.

The following corollary uses the continuous mapping theorem to obtain the limiting distribution for $\hat{\theta}_n$ as n tends to infinity.

Corollary 2.2. *Consider the estimator of θ, $\hat{\theta}_n = \bigwedge_{t \in I_n}(X_t - \hat{\phi}_n X_{\lfloor t/2 \rfloor})$. Suppose $\theta > 0$ and F is RV_α at θ. If $\alpha > \beta$ then for any $y > 0$ we have*

$$\lim_{n \to \infty} P[a_n^{-1}(\hat{\theta}_n - \theta) > y] = e^{-y^\alpha/2}.$$

3 Simulation Study

In this section, we assess the reliability and performance of our extreme value estimation method. In doing so, we will examine the finite sample properties of $(\hat{\phi}_n, \hat{\theta}_n)$, where $\hat{\theta}_n = \bigwedge_{t \in I_n}(X_t - \hat{\phi}_n X_{\lfloor t/2 \rfloor})$ and $0 < \rho < 1$ is a fixed value, $b_n = F^{\leftarrow}(1 - 1/2^n)$ and $a_n = F^{\leftarrow}(1/2^n) - \theta$. To study the performance of the estimators $\hat{\phi}_n = \bigwedge_{t=2}^{n} \frac{X_t}{X_{\lfloor t/2 \rfloor}}$ and $\hat{\theta}_n = \bigwedge_{t \in I_n}(X_t - \hat{\phi}_n X_{\lfloor t/2 \rfloor})$, we generated 5,000 independent samples of size $n = 2^{k+1} - 1$, where k is the number of generations and $\{X_t\}$ is a BAR(1) process satisfying the difference equation

$$X_t = \phi X_{\lfloor t/2 \rfloor} + \epsilon_t, \quad \text{for } 2 \le t \le n \quad \text{and } \epsilon_t \ge \theta.$$

The autoregressive parameter ϕ is taken to be in the range from 0 to 1 guaranteeing a nonnegative time series since the unknown location parameter θ is positive. In order to perform this simulation the following ad hoc approach was adopted to generate bivariate random variables from a distribution that is regularly varying at both endpoints. First, let Z_1 and Z_2 be independent random variables that are taken from

$$F(z) = \begin{cases} c(z - \theta)^{\alpha} & \text{if } \theta < z < \theta + 1, \\ 1 - d(z - \theta)^{-\beta} & \text{if } \theta + 1 < z < \infty. \end{cases}$$

For this innovation distribution let c and d be nonnegative constants such that $c + d = 1$. Then this distribution is regularly varying at both endpoints with index of regular variation $-\beta$ at infinity and index of regular variation α at θ. Now define

$$\epsilon_{2t} \stackrel{d}{=} a_1 Z_1 + b_1 Z_2 \quad \text{and} \quad \epsilon_{2t+1} \stackrel{d}{=} a_2 Z_1 + b_2 Z_2,$$

where a_i and b_i are nonnegative constants such that $a_i + b_i = 1$, for $i = 1, 2$.
For this simulation study we have considered two different cases: The first case is when the innovations are i.i.d and the second case is when the innovation pairs are dependent. In order to achieve this two distributions were considered:

(i) Define F_1 such that $c = d = .5, a_1 = 1, a_2 = 0, b_1 = 0, b_2 = 1$

(ii) Define F_2 such that $c = d = .5, a_1 = .7, a_2 = .6, b_1 = .3, b_2 = .4$.

Observe in case (i) that $(\epsilon_{2t}, \epsilon_{2t+1})$ are i.i.d. with a regular varying tail distribution at infinity with index $-\beta$ and regular varying at θ with index α, whereas in case (ii) $(\epsilon_{2t}, \epsilon_{2t+1})$ are dependent with a regular varying tail with index $-\beta$ at infinity and index α at θ.
Choosing $k = 4, 9, 13, 17$ and $n = 2^{k+1} - 1$, the sample mean and standard error (s.e.) of the estimates are given in Table 1 and 2 for $\phi = .3$, $\theta = 2$, respectively. Additionally, the average lengths for 95 percent empirical confidence intervals with exact coverage are also reported. Note that the confidence intervals were directly constructed from the empirical distributions of $n^{1/\beta}(\hat{\phi}_n - \phi)$ and $n^{1/\alpha}(\hat{\theta}_n - \theta)$ respectively, while the exponent ρ inside the index set I_n was set to .9.
We first examine the simulation results in Table 1 for different number of generations. As k increases, the standard errors and biases of $\hat{\phi}_n$ and $\hat{\theta}_n$ decrease. In particular, when $\beta = .3$ and k increases from 9 to 13, the standard error becomes

1,712 times smaller. Similarly, the 95% confidence interval average length is 5,580 times smaller.

Next we look at the behavior of the estimators as β increases. Not surprisingly, the standard errors and bias get larger as β increases. This is expected since our extreme value method of estimation depends heavily on obtaining large innovations. Thus, it can be shown when the regular varying index is small (less than 1) that the distance in quantity between the largest innovation Z_t and other innovations will be extremely large, and only in the situation when a large sample value occurs followed by innovations in the next generation at least one of which is as large, does our estimator behave badly. Whereas, if β takes on values larger than one, then the distance in quantity between the largest innovation and other innovations is not as likely to be as large, thus the chance that we get a bad estimate increases. Therefore, we can expect with small probability to observe some extremely poor estimates from our estimator.

Table 1: Performance of $(\hat{\phi}_n, \hat{\theta}_n)$ with $(\phi = .3, \theta = 2)$ and $\alpha = 1$ under F_1

		$\hat{\phi}_n$		$\hat{\theta}_n$		95% C.I. Avg. Length	
k	β	mean	s.e.	mean	s.e.	$\hat{\phi}_n$	$\hat{\theta}_n$
4	.3	.3002	(4.76×10^{-3})	2.09	(.0738)	5.26×10^{-3}	.0363
	.9	.3089	(8.39×10^{-1})	2.24	(.1689)	9.47×10^{-1}	.0748
9	.3	.30016	(7.21×10^{-8})	2.005	(.0035)	6.92×10^{-5}	.0195
	.9	.3003	(1.56×10^{-3})	2.091	(.0249)	2.74×10^{-2}	.0397
13	.3	.300074	(4.21×10^{-11})	2.0082	(.0028)	1.24×10^{-8}	.0092
	.9	.30078	(7.83×10^{-5})	2.047	(.0153)	7.34×10^{-3}	.0203
17	.3	.300001	(5.28×10^{-13})	2.0024	(.0009)	5.19×10^{-10}	.0033
	.9	.30023	(6.29×10^{-7})	2.006	(.0077)	3.56×10^{-5}	.0096

We now turn our attention to Table 2. The purpose of this table is to see whether or not the correlation between $(\epsilon_{2t}, \epsilon_{2t+1})$ affects our estimates for the autocorrelation parameter ϕ. The results are expected from a biological viewpoint, as one expects the environmental correlation between the sisters to be distinct from the environmental correlations inherited from the mother. Hence, the results seem to suggest that a cell's attributes can be explained by inheritance from its mother, suggesting that a BAR(1) model for a single line of descent is appropriate.

Table 2: Performance of $(\hat{\phi}_n, \hat{\theta}_n)$ with $(\phi = .3, \theta = 2)$ and $\alpha = 1$ under F_2

		$\hat{\phi}_n$		$\hat{\theta}_n$		95% C.I. Avg. Length	
k	β	mean	s.e.	mean	s.e.	$\hat{\phi}_n$	$\hat{\theta}_n$
4	.3	.3011	(2.44×10^{-2})	2.17	(.1148)	8.45×10^{-3}	.0543
	.9	.3109	(9.67×10^{-1})	2.44	(.2408)	9.95×10^{-1}	.1148
9	.3	.30079	(8.34×10^{-5})	2.05	(.0108)	5.72×10^{-4}	.0262
	.9	.3071	(4.52×10^{-2})	2.03	(.0967)	8.34×10^{-2}	.0432
13	.3	.30059	(6.28×10^{-7})	2.0089	(.0051)	3.19×10^{-7}	.0147
	.9	.30052	(3.74×10^{-4})	2.0078	(.0427)	5.21×10^{-3}	.0379
17	.3	.300031	(3.67×10^{-10})	2.0021	(.0018)	7.16×10^{-9}	.0079
	.9	.30058	(3.24×10^{-6})	2.0028	(.0104)	3.22×10^{-5}	.0142

4 The Elegance of An Extreme Value Approach

In this section we use demonstrate the effectiveness of an extreme value approach by not only filling in the gaps for the proof of Proposition 2 in Zhang (2011), but also verifying that our alternative approach under the correct parameterization can obtain the same limit law found in Theorem 2 of Zhou and Basawa (2005b). While this approach may seem unless under the given conditions, it does reveal the complexity and difficulty found in Zhou and Basawa (2005b) is unnecessary. In order to obtain the same limit law found in Zhou and Basawa 92005b) for our estimator $\hat{\phi}_n$ we must first apply the assumption that $Y_t = \epsilon_{2t} \wedge \epsilon_{2t+1}$ is positive and the marginal distribution is regularly varying at $\theta = 0$ with index α. In Section 2, we obtained the limit law for $\hat{\phi}_n$ under the assumption that the innovation distribution for $\{\epsilon_t\}$ was regularly varying at infinity with index $-\beta$. We now continue in this section with the usual first-order bifurcating autoregressive process defined by

$$X_t = \phi X_{\lfloor \frac{t}{2} \rfloor} + \epsilon_t, \quad \text{for } 2 \leq t \leq n, \tag{4.1}$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x. For the model in (4.1) we assume that $0 < \phi < 1$ and the innovations ϵ_t are such that $(\epsilon_{2t}, \epsilon_{2t+1})$ are i.i.d with $(\epsilon_{2t}, \epsilon_{2t+1}) \sim F_1$ and ϵ_t has the same marginal distribution F satisfying $F(\theta) = 0$. By defining $G(x) = 2F(x) - F_1(x, x)$, we assume G is regularly varying at $\theta = 0$ with index α. The following Theorem generalizes the result presented in Zhang (2011) where the author assumes Weibull type innovations.

Theorem 4.1. *Let $d_n = G^{\leftarrow}(1/n)$ and consider the estimator of ϕ, $\hat{\phi}_n = \bigwedge_{t=2}^{n} X_t / X_{\lfloor t/2 \rfloor}$. Under the assumption that $G \in RV_\alpha$ and $EX_1^\gamma < \infty$ for some $\gamma > \alpha$, we have*

$$\lim_{n\to\infty} P[d_n^{-1}(\hat{\phi}_n - \phi) > y] = e^{-y^\alpha} EX_1^\alpha / 2,$$

where X_1 has the stationary distribution H for the process in (4.1).

Remark 2. The stationary distribution H is the same as the stationary distribution of the AR(1) sequence $\tilde{X}_t = \phi \tilde{X}_{t-1} + \tilde{\epsilon}_t$, where the $\tilde{\epsilon}_t$ are i.i.d. with the same marginal distribution as ϵ_t.

The following Corollary is a verification that our extreme value method under the specified bivariate exponential innovation distribution is in agreement with the limit law presented in Theorem 2 of Zhou and Basawa (2005b). That is, suppose the joint distribution of $(\epsilon_{2t}, \epsilon_{2t+1})$ is specified by

$$\bar{F}_1(x_1, x_2) = P[\epsilon_{2t} > x_1, \epsilon_{2t+1} > x_2] = \exp\left(-\alpha(x_1 + x_2) - \beta(x_1 \vee x_2)\right), \quad x_1, x_2 > 0, \tag{4.2}$$

where α and β are the model parameters satisfying $\alpha > 0, \beta > 0$.

Observe that the marginal distribution of ϵ_{2t} and ϵ_{2t+1} are exponential with mean $(\alpha + \beta)^{-1}$ and correlation $\rho = \beta(2\alpha + \beta)^{-1}$.

Now we consider the parameterization

$$\alpha = \frac{1-\rho}{(1+\rho)\lambda} \quad \text{and} \quad \beta = \frac{2\rho}{(1+\rho)\lambda}, \tag{4.3}$$

where $\lambda > 0$ and $0 \leq \rho < 1$. With this parameterization, the marginal distributions of ϵ_{2t} and ϵ_{2t+1} are both exponential with mean λ and correlation ρ. Observe when $\rho = 0$, the innovations $\{\epsilon_t\}$ in (4.1) will be independent and identically exponential distributed random variables.

Corollary 4.1. *Suppose* $(\epsilon_{2t}, \epsilon_{2t+1}) \sim F_1$, *where* F_1 *is specified in* (4.2)*. Then with the parameterization defined in* (4.3) *we have*

$$\lim_{n\to\infty} P\left[\frac{n}{(1+\rho)(1-\phi)}(\hat{\phi}_n - \phi) > y\right] = e^{-y}.$$

5 Selected Proofs

Prior to proving any results, our first objective is to look at the dependency among $X_t^{(m)}$. Upon doing so, we realized that the dependency within each tree segment is more delicate than anticipated, hence the following argument should be considered heuristic. The thought process to obtain independence was to determine how much distance was needed between observations. With this in mind, we begin by partitioning the index set $[1, 2^{k+1} - 1]$.
Thus, for the l^{th} generation we consider intervals

$$Z_{l,s} = [2^l + (s-1)r, 2^l + sr - 1],$$

for $s = 1, \ldots, \lfloor 2^l/r \rfloor, l = l_0, \ldots, k$, where l_0 is such that $2^{l_0 - 1} < r \le 2^{l_0}$ and $r \ge 2^{m+1} + 1$. Notice that the number of indices in each interval is at most $2^{l+1} - 1 - 2^l - r(2^l/r - 1) + 1 = r$. Now we consider trimming the intervals by 2^{m+1} in hopes of achieving the necessary independence. Thus, we define $Q_{l,s} = [2^l + (s - 1)r, 2^l + sr - 1 - 2^{m+1}]$. Then

$$dist(Q_{l,s}, Q_{l,s+1}) = 2^l + sr - (2^l + sr - 1 - 2^{m+1}) = 2^{m+1} + 1 > 2^{m+1}.$$

Therefore, $X_t^{(m)}$ is independent of $X_{t'}^{(m)}$, for all $t \in Q_{l,s}$ and $t' \in Q_{l',s'}$ provided $(l, s) \neq (l', s')$. This is true by construction when $l = l'$ and $s \neq s'$. In the case $l \neq l'$, we have

$$dist(Q_{l,s}, Q_{l-1,\lfloor 2^{l-1}/r \rfloor}) = 2^l - (2^{l-1} + \lfloor 2^{l-1}/r \rfloor - 1 - 2^{m+1} \ge 2^{m+1} + 1 > 2^{m+1}.$$

Remark 1. As stated above, it is not the case that $X_t^{(m)}$ will be independent of $X_{t'}^{(m)}$ for all t and t'. That is, most of the time the distance between indices t and t' will be large enough so that observations $X_t^{(m)}$ and $X_{t'}^{(m)}$ will be independent, but there are a few scenarios where this is not true. While this is a concern, we verified through simulation that asymptotically this dependency does not affect our estimators.
Proof of Lemma 2.3.

Proof. With out loss of generality suppose $j = 1$. Then applying Lemma 4.24 in Resnick (1987) for $\epsilon > 0$ and k large we have

$$\begin{aligned} P[I_{(l,s)} = (1, 0)] &\le r2^{k+1} P[a_1 \le Y_1 \le b_1] P[b_n a_1' \le X_1^{*(m)} \le b_n b_1'] \\ &\le (1+\epsilon) r 2^{k+1} (G(b_1) - G(a_1)) P[b_n a_1' \le \sum_{j=0}^{m} \phi^j \epsilon_{2^j} \le b_n b_1'] \\ &= (1+\epsilon)\frac{r}{2^{k+1}}(G(b_1) - G(a_1))(b_1'^{-\beta} - a_1'^{-\beta})\frac{1-\phi^{(m\beta)}}{1-\phi^\beta}. \end{aligned}$$

Thus,

$$\limsup_{k\to\infty} \frac{2^{k+1}}{r} P[I_{(l,s)} = (1,0)] \leq \gamma_1.$$

Now to obtain a lower bound observe that

$$\begin{aligned} P[I_{(l,s)} = (1,0)] &\geq (r - 2^{m+1}) P[a_1 \leq Y_1 \leq b_1] P[b_n a_1^{'} \leq X_1^{*(m)} \leq b_n b_1^{'}] \\ &- r^2 P^2 [b_n a_1^{'} \leq X_1^{*(m)} \leq b_n b_1^{'}] \\ &\geq (1-\epsilon) \frac{r}{2^{k+1}} \gamma_1 - 2 \left(\frac{r}{2^{k+1}} \right)^2 (b_1^{'-2\beta} - a_1^{'-2\beta}) \sum_{j=0}^{m} \phi^{(2\beta)j}. \end{aligned}$$

Then

$$\liminf_{k\to\infty} \frac{2^{k+1}}{r} P[I_{(l,s)} = (1,0)] \geq \gamma_1,$$

which completes the proof. □

Proof of Corollary 2.1.

Proof. First observe that

$$P[b_n(\hat{\phi}_n - \phi) > x] = P \left[\bigwedge_{t=1}^{\lfloor n/2 \rfloor} \left(\frac{Y_t}{b_n^{-1} X_t} \right) > x \right]. \tag{5.1}$$

Now define a subset of $\mathbb{R}_+^2$ by $A_x = \{y_1, y_2 : y_1/y_2 \leq x, y_1, y_2 > 0\}$. Then it suffices to show that there are no points t that satisfies the condition in A_x. Thus, if we let $y_1 = Y_t$ and $y_2 = b_n^{-1} X_t$, then notice that (5.1) is equivalent to $(\mathcal{I}_n(A_x) = 0)$. Furthermore, observe that A_x is a bounded set in $E = [\theta, \infty) \times (0, \infty]$ provided $\theta > 0$. Therefore, assuming $\phi > 0$ and applying Lemma 2.5 in Bartlett and McCormick (2012) we have that the point process $\mathcal{I}_n^{(m)} = \sum_{t=1}^{\lfloor n/2 \rfloor} \varepsilon_{(Y_t, b_n^{-1} X_t^{(m)})}$ is equivalent to $\mathcal{I}_n = \sum_{t=1}^{\lfloor n/2 \rfloor} \varepsilon_{(Y_t, b_n^{-1} X_t)}$. Hence using Theorem 2.1 and (5.1), we have that

$$\begin{aligned} \lim_{n\to\infty} P[b_n(\hat{\phi}_n - \phi) > x] &= \lim_{n\to\infty} P[\mathcal{I}_n(A_x) = 0] \\ &= \exp\left(-(1-\phi^\beta)^{-1} \int_\theta^\infty \left(\frac{y_1}{x} \right)^{-\beta} \bar{G}(y_1/x) dy_1 \right) \\ &= \exp\left(-(1-\phi^\beta)^{-1} x^\beta \int_\theta^\infty y_1^{-\beta} \bar{G}(y_1/x) dy_1 \right) \\ &= e^{-x^\beta E Y^{-\beta} (1-\phi^\beta)^{-1}}. \end{aligned}$$

□

Proof of Theorem 2.2.

Proof. First observe from (2.7) we have

$$nP[a_n^{-1}(Y_1 - \theta) \in \cdot] \xrightarrow{v} \nu \quad \text{in} \quad E_2$$

where

$$\nu[0,y] = \lim_{n\to\infty} P[a_n^{-1}(Y_t-\theta)\le y] = \lim_{n\to\infty} n/2P[Y_1 \le (\theta + a_n y)]$$
$$= \lim_{n\to\infty} 1/2 \frac{G(\theta+a_n y)}{G(\theta+a_n)}$$
$$= y^\alpha/2.$$

The result now follows from the fact that $\{Y_t, t = 1,\ldots,\lfloor n/2\rfloor\}$ are i.i.d. random elements of $(E_2,\mathcal{B})$ where E_2 is locally compact, $\mathcal{B}$ is the Borel σ-algebra, and ν is a Radon measure on $(E_2,\mathcal{B})$. Therefore, by (proposition 3.21 in Resnick (1987)) we have

$$\mathcal{V}_n \Rightarrow \mathcal{V}.$$

□

Proof of Corollary 2.2.

Proof. By Lemma 2.5 and (2.14) we have

$$\lim_{n\to\infty} P[a_n^{-1}(\hat\theta-\theta) > y] = \lim_{n\to\infty} P[a_n^{-1}(\bigwedge_{t\in I_n} Y_t - \theta) > y] + o(1).$$

Now observe that

$$0 \le P\Big[a_n^{-1}(\bigwedge_{t\in I_n} Y_t-\theta) > y\Big] - P\Big[a_n^{-1}(\bigwedge_{t=1}^{\lfloor n/2\rfloor} Y_t-\theta) > y\Big]$$
$$= P\Big[a_n^{-1}(\bigwedge_{t=1}^{\lfloor n/2\rfloor} Y_t-\theta) \le y < a_n^{-1}(\bigwedge_{t\in I_n} Y_t-\theta)\Big]$$
$$\le P\Big[\bigcup_{1\le t\le\lfloor n/2\rfloor} \big(X_t > (a_n b_n)^\rho \text{ and } a_n^{-1}(Y_t-\theta)\le y\big)\Big]$$
$$\le nP[X_1^* > (a_n b_n)^\rho] P[Y_1 \le \ a_n y + \theta] = o(1). \tag{4.5.3}$$

It then follows from (4.5.3) that

$$\lim_{n\to\infty} P\left[\bigwedge_{t\in I_n}^{\lfloor n/2\rfloor}\left(\frac{Y_t-\theta}{a_n}\right) > y\right] = \lim_{n\to\infty} P\left[\bigwedge_{t=1}^{\lfloor n/2\rfloor}\left(\frac{Y_t-\theta}{a_n}\right) > y\right] + o(1).$$

Now observe from Lemma 2.6 that the point process $\mathcal{V}_n$ suffices. Therefore, if we define the subset $B_y = \{z : z \le y, z > 0\}$, then $\bigwedge_{t=1}^{\lfloor n/2\rfloor}\left(\frac{Y_t-\theta}{a_n}\right) > y$ is equivalent to $(\mathcal{V}_n(B_y) = 0)$. The result now follows from Theorem 2.2 since

$$\lim_{n\to\infty} P[a_n^{-1}(\hat\theta_n-\theta) > y] = \lim_{n\to\infty} P\left[\bigwedge_{t=1}^{\lfloor n/2\rfloor} a_n^{-1}(Y_t-\theta) > y\right] + o(1)$$
$$= \lim_{n\to\infty} P[\mathcal{V}_n(B_y) = 0]$$
$$= P[\mathcal{V}(B_y) = 0] = e^{-y^\alpha/2}.$$

□

6 References

[1] Bartlett, A., McCormick W.P. (2012). Estimation for Non-negative First order Autoregressive Process with an Unknown Location Paramter. *Applied Mathematics*, **3 No. 12A**, 2133 - 2147.

[2] Bercu, B., Saporta DE, Gegout-Petit B. (2009). Asymptotic analysis for bifurcating autoregressive processes via a martingale approach. *Electron. J. Probab.,* **14**, 2492 - 2526.

[3] Cowan, R. Staudte, R. G. (1986). The bifurcating autoregressive model in cell lineage studies. *Bioetrics*, **42**, 769 - 783.

[4] Davis, R.A., McCormick, W.P. (1989). Estimation for first-order autoregressive processes with positive or bounded innovations. *Stochastic Process. Appl.* **31**, 237–250.

[5] Huggins, R.M., Basawa, I.V. (1999). Extensions of the bifurcating autoregressive model for cell lineage studies. *J. Appl. Probab.* **36**, 1225 - 1233.

[6] Huggins, R.M., Basawa, I.V. (2000). Inference for the extended bifurcating autoregressive model for cell lineage studies. *Aust. N. Z. J. Stat.* **42**, 423 - 432.

[7] Hwang, S.Y., Basawa, I.V. (2009). Local asymptotic normaility for bifurcating autoregressive processes and related asymptotic inference. *Statistical Methodology.* **6**, 61 - 69.

[8] Hwang, S.Y., Basawa, I.V. (2009). Branching Markov processes and related asymptotics. *Journal of Multivariate Analysis* **100**, 1155 - 1161.

[9] Kallenberg, O. (1976). *Random Measures*, Akademie-Verlag, Berlin.

[10] Resnick, S.I., (1985). Point Processes, regular variation and weak convergence. *Adv. Appl. Probab.* **18**, 66 – 138.

[11] Resnick, S.I. (1987). *Extreme Values, Regular Variation, and Point Processes*, Springer, New York.

[12] Resnick, S.I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*, Springer, New York.

[13] Zhang, C. (2011). Parameter Estimation for First-order Bifurcating Autoregressive Processes with Weibull Type Innovations. *Statistics and Probability Letters*, **81**, 1961 - 1969.

[14] Zhou, J. Basawa, I.V. (2004). Non-Gaussian bifurcating models and quasi-likelihood estimation. Stochastic methods and their applications. *J.Appl. Probab.*, **41A**, 55 - 64.

[15] Zhou, J. Basawa, I.V. (2005a). Least-squares estimation for bifurcating autoregressive processes. *Statist. Probab. Lett.* **74**, 77 - 88.

[16] Zhou, J. Basawa, I.V. (2005b). Maximum Likelihood Estimation for a First-Order Bifurcating Autoregressive Process with Exponential Errors. *J. Time Series Anal.*, **26**, 825 - 842.