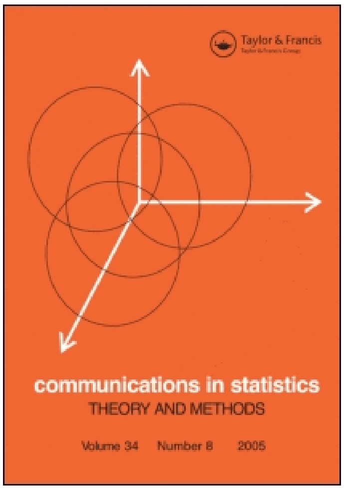

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597238

Estimation for the semipareto processes

N. Balakrishna[a]
[a] Department of Statistics, Cochin University of Science & Technology, Cochin, India

To cite this Article Balakrishna, N.(1998) 'Estimation for the semipareto processes', Communications in Statistics - Theory and Methods, 27: 9, 2307 — 2323
To link to this Article: DOI: 10.1080/03610929808832229
URL: http://dx.doi.org/10.1080/03610929808832229

COMMUN. STATIST.—THEORY METH., 27(9), 2307-2323 (1998)

ESTIMATION FOR THE SEMIPARETO PROCESSES

N.BALAKRISHNA

Department of Statistics
Cochin University of Science & Technology,
Cochin-682 022
INDIA

Key Words: *α-mixing; consistent and asymptotically normal estimators; ergodicity; Pareto process.*

ABSTRACT

This paper proposes different estimators for the parameters of SemiPareto and Pareto autoregressive minification processes. The asymptotic properties of the estimators are established by showing that the SemiPareto process is α-mixing. Asymptotic variances of different moment and maximum likelihood estimators are compared.

1. INTRODUCTION

It is well-known that the autoregressive models of appropriate orders are extensively used for modelling time series data. Even though the classical analysis of time series rests heavily on the Gaussian assumption, in recent years there are many models introduced for explaining the time series data using non-

Gaussian distributions. In some cases these models are proved to be more suitable than their Gaussian counterparts (See eg. Lawrance and Lewis (1985)).

In the study of non-Gaussian time series modelling, we assume the form of a stochastic model and suppose under certain conditions that it generates a stationary sequence of random variables (r.v.s) of specified marginal distributions. The main tools used for studying these models are generating functions such as Laplace transforms or characteristic functions. If these generating functions do not have closed form expressions then it is difficult to handle the model mathematically. In such cases alternative models having minification (instead of addition) structures are introduced. These minification models possess most of the properties of autoregressive models. The existence of such models and their properties can be easily studied using the survival function of the underlying r.v. For a general discussion on minification processes see Lewis and Mckenzie (1991).

It is needless to say that if one wants to verify the suitability of the model to explain real life situations, one has to have reasonably good estimators for the unknown parameters. However, not much work is available on inference for non-Gaussian time series models, when the marginals are not exponentials. The estimation problems related to autoregressive processes generating exponential r.v.s are studied by Adke and Balakrishna (1992), Billard and Mohamed (1991), Karlsen and Tjφstheim (1988), Smith (1986) and Raftery (1980,1981). Jin-Guan and Yuan (1991) discuss estimation in some discrete time series models. Adke and Balakrishna (1992) have considered the parameter estimation of the exponential minification process of Tavares (1980). Except this, no detailed study on estimating the minification processes is available in the literature.

This paper discusses the problem of estimating the parameters of the SemiPareto process of Pillai (1991) and Pareto process of Yeh et al. (1988). Let us describe these models.

A r.v. X is said to have a SemiPareto distribution if its survival function is of the form

$$\bar{F}_x(x) = \Pr(X>x) = \frac{1}{1+\psi(x)}, \qquad x \geq 0 \tag{1.1}$$

where $$\psi(x) = \tfrac{1}{p}\psi\left(p^{1/\alpha}x\right), \qquad 0<p<1,\ \alpha > 0. \tag{1.2}$$

The solution of the equation (1.2) can be shown to be $\psi(x) = x^{\alpha}h(x)$, where the function $h(x)$ is periodic in $\ln x$ with period $\dfrac{2\pi\alpha}{-\ln p}$ [cf. Pillai (1991)].

For example, if $\psi(x) = (x/\sigma)^{\alpha}$, $\sigma>0$, then (1.1) becomes the survival function of a Pareto type III distribution. That is,

$$\bar{F}_x(x) = \Pr(X>x) = \frac{1}{1+(x/\sigma)^{\alpha}}, \qquad x \geq 0,\ \sigma>0, \alpha>0. \tag{1.3}$$

From (1.2) it follows that

$$\lim_{x\to 0}\psi(x) = 0 \text{ and } \lim_{x\to\infty}\psi(x) = \infty. \tag{1.4}$$

A large variety of socio-economic variables have distributions which are heavy tailed and well fitted by Pareto type distributions. The SemiPareto distribution is a generalization of the Pareto type III distribution which is more flexible in the sense that a variety of survival functions can be defined by properly choosing the parameters α, p and the function $\psi(.)$ satisfying (1.2). So whenever we have a data, with a tendency to follow heavy tailed distribution, it is more advantageous to use SemiPareto distribution where in we can change the function $\psi(.)$ until it fit in to the data properly. However, in many situations the observations at different time points are not independent. In such cases a Markov dependent sequence will be a better choice than a sequence of i.i.d.r.v.s. Even though linear autoregressive models are used to generate Markov dependent sequences in classical time series, due to the reasons explained earlier, it becomes difficult in the case of semiPareto distributions as its Laplace transform has no simple expression. Hence, the following minification model is used for the purpose.

A first order SemiPareto autoregressive minification (SPAM(1)) process $\{X_n\}$ is defined by the model

$$X_n = p^{-1/\alpha} \min(X_{n-1}, Y_n), \; n = 1,2,\ldots, \tag{1.5}$$

where $\{Y_n\}$ is a sequence of independent and identically distributed (iid) extended real-valued r.v.s called innovations, having common survival function

$$\bar{F}_y(y) = \{p + p\psi(y)\} / \{p + \psi(y)\} \tag{1.6}$$

and X_0 has the survival function (1.1) which is independent of Y_j, $j \geq 1$. Then $\{X_n\}$ defines a stationary Markov sequence with each X_n having a survival function (1.1). Using (1.4) we have

$$\lim_{y \to \infty} \bar{F}_y(y) = p > 0.$$

Yeh et al. (1988) have studied the first order Pareto type III minification process (PAM(1)), where each X_n defined by (1.5) has the survival function (1.3). For using the above model in emperical studies, one has to have some estimators for the unknown parameters. The work of this paper is an attempt towards that end. We may note that if X has the survival function (1.3), $E(X^\delta)$ will be finite only if $-\alpha<\delta<\alpha$. Therefore, any estimators based on sample moments require assumptions on existence of moments of appropriate orders. However, by the properties of the model (1.5) discussed in the next section, we are able to propose some consistent and asymptotically normal (CAN) estimators of the parameters without any additional assumptions.

The Section 2 discusses some special properties of $\{X_n\}$, such as ergodicity and mixing. In Section 3, we propose different estimators for α, p and ψ and prove their CAN properties. The properties of estimators in the case of PAM(1) process are studied in Section 4.

2. SOME USEFUL PROPERTIES OF SPAM (1) PROCESS

Let $\{X_n\}$ be a SPAM(1) process defined by (1.5). For some $t > 0$, define

$$Z_n(t) = \begin{cases} 1, & X_n > t \\ 0, & X_n \le t \end{cases} , \quad n = 0,1,\ldots \tag{2.1}$$

Then it is proved that (see Pillai (1991)) $\{Z_n(t), n \ge 0\}$ is a Markov chain on $\{0,1\}$ with transition probabilities:

$$\begin{aligned} p_{01} &= (1-p)/\{1+\psi(t)\}, \; p_{00} = 1-p_{01} \\ p_{10} &= (1-p)\, \psi(t)/\{1+\psi(t)\}, \; p_{11} = 1-p_{10}. \end{aligned} \tag{2.2}$$

Moreover, the autocorrelation function of $\{Z_n(t)\}$ is given by

$$\rho_z(h) = \mathrm{Corr}(Z_n, Z_{n+h}) = p^h, \; h = 1,2,\ldots$$

Note that, if $\{X_n\}$ is defined by (1.5) then the minimal sigma field induced by $\{X_0, X_1, \ldots X_n\}$ remains same as the one induced by the set of independent r.v.s $\{X_0, Y_1, Y_2, \ldots, Y_n\}$. As a consequence it follows that the sequence $\{X_n\}$ and hence $\{Z_n(t)\}$ are stationary and ergodic. The Markov property of $\{Z_n(t)\}$ distinguishes the semi-Pareto (& Pareto) process from other minification processes (see Arnold and Hallett (1988)). This helps us in estimating the parameters of these models. The following results are useful for establishing the properties of our estimators.

A sequence $\{X_n\}$ of r.v.s is said to be α-mixing in the sense of Billingsley (1986) if

$$|P(A \cap B) - P(A)P(B)| \le \alpha_h \tag{2.3}$$

for all $A \in \sigma\{X_0, X_1, \ldots X_n\}$ and $B \in \sigma\{X_{n+h}, X_{n+h+1}, \ldots\}$, where $\sigma\{X_0, X_1, \ldots X_n\}$ is the minimal sigma field induced by $\{X_0, X_1, \ldots X_n\}$. The sequence $\{\alpha_h\}$ is called mixing parameters and $\alpha_h \to 0$ as $h \to \infty$. Now we state the following Lemma from Billingsley (1986), p.376 for our reference.

Lemma 2.1: Suppose that $\{X_n, n>0\}$ is stationary and α-mixing sequence with

$$E(X_n) = \mu, \; \alpha_n = O(n^{-5}) \text{ and } E\left(X_n^{12}\right) < \infty. \tag{2.4}$$

Let $S_n = X_1+X_2+...+X_n$, then

$$n^{-1}\,\mathrm{Var}(S_n) \to \sigma^2 = \mathrm{Var}(X_1)+2\sum_{j=1}^{\infty}\mathrm{cov}(X_1, X_{1+j}), \qquad (2.5)$$

where the series converges absolutely. If $\sigma^2>0$ then $\sqrt{n}(\overline{X}_n-\mu)/\sigma \xrightarrow{L} N(0,1)$ as $n\to\infty$, where $\xrightarrow{L}$ denotes the convergence in distribution, N(0,1) is a standard normal variate and $\overline{X}_n = S_n/n$. That is, $\overline{X}_n$ is asymptotically normal with mean μ and asymptotic variance σ^2/n. We denote it by $\overline{X}_n \sim \mathrm{AN}(\mu,\sigma^2/n)$.

The joint survival function $\overline{F}_h(x,y)$ of X_n and X_{n+h} is given by

$$\overline{F}_h(x,y) = \Pr[X_n>x,\ X_{n+h}>y]$$

$$= \begin{cases} \dfrac{1+p^h\psi(y)}{\{1+\psi(x)\}\{1+\psi(y)\}}, & 0\le y< p^{\frac{-h}{\alpha}}x \\ \\ \dfrac{1}{\{1+\psi(y)\}}, & y\ge p^{\frac{-h}{\alpha}}x. \end{cases} \qquad (2.6)$$

Over the range $y< p^{\frac{-h}{\alpha}}x$, the density function corresponding to $\overline{F}_h(x,y)$ is given by

$$h(x,y) = \frac{(1-p^h)d\psi(x).d\psi(y)}{\{1+\psi(x)\}^2\{1+\psi(y)\}^2} \qquad (2.7)$$

Lemma 2.2: The SPAM(1) process is α-mixing with the mixing parameter

$$\alpha_h = \frac{2}{1-p^h}p^h\log p^{-h}, \qquad h = 1,2,... \qquad (2.8)$$

Proof:

Let A and B be two events such that $A\in\sigma\{X_0,X_1,...X_n\}$ and $B\in\sigma\{X_{n+h},X_{n+h+1},...\}$. Consider

$$\Pr\{A\cap B\} = \int_0^\infty\int_0^\infty P\{A\cap B|X_n = x, X_{n+h} = y\}dF_h(x,y)$$

$$= \int_0^\infty \int_{yp^{h/\alpha}}^\infty P\{A\cap B|X_n = x, X_{n+h} = y\}h(x,y)dxdy$$

$$+ \Pr\left(X_{n+h} = p^{-h/\alpha}X_n\right) P\left(A\cap B\middle|X_{n+h} = p^{-h/\alpha}X_n\right),$$

where F (.,.) and h(.,.) are as defined by (2.6) and (2.7). By definition of the model, it is true that

$$\Pr\left(X_{n+h} = p^{-h/\alpha}X_n\right) = \Pr\ (Y_1 > X_0, Y_2 > p^{-1/\alpha}X_0, \ldots, Y_h > p^{-(h-1)/\alpha}X_0).$$

Now by conditioning on X_0 and using the equations (1.6) and (1.2) we get

$$\Pr\left(X_{n+h} = p^{-h/\alpha}X_n\right) = \int_0^\infty \frac{p + p\psi(x)}{[p + p^{-(h-1)}\psi(x)]} \frac{d\psi(x)}{[1+\psi(x)]^2} dx$$

$$= \{-p^h \log p^h\}/(1-p^h). \tag{2.9}$$

Using the Markov property of $\{X_n\}$, we can write

$$|P(A\cap B) - P(A).P(B)| \le \Pr\left(X_{n+h} = p^{-h/\alpha}X_n\right) + \int_0^\infty \int_0^{yp^{h/\alpha}} f(x)f(y)dxdy$$

$$+ \int_0^\infty \int_{yp^{h/\alpha}}^\infty |h(x,y) - f(x).f(y)|dxdy,$$

where $f(x)$ and $f(y)$ are the density functions of X_n and X_{n+h}. On simplifying the integrals and using (2.9) it is readily proved that the right side of the above inequality reduces to α_h defined by (2.8). Clearly $\alpha_h \to 0$ as $h \to \infty$. Hence the lemma is proved.

3. ESTIMATION OF SPAM(1) MODEL

The structure of a minification model makes the likelihood based inference more complicated when we use observations from $\{X_n\}$. The method of moments requires an explicit form for the function $\psi(.)$ as well as the existence of moments. In this section, we propose estimators for α, p and $\psi(.)$ using the properties discussed in Section 2. For the model (1.5), the statistic,

$$\hat{k}_n = \max_{1\le i\le n}\left(\frac{X_i}{X_{i-1}}\right) \tag{3.1}$$

may be taken as an estimator of $k = p^{-1/\alpha}$. It can be proved that $\hat{k}_n$ is strongly consistent estimator of $p^{-1/\alpha}$, but $\sqrt{n}(\hat{k}_n - k)$ always converges in distribution to a degenerate r.v and hence this estimator is not CAN. Throughout the discussion we take $\hat{k}_n$ as the estimator of $k = p^{-1/\alpha}$ and then estimate α by

$$\hat{\alpha}_n = \{-\log \hat{p}_n\} / \log \hat{k}_n, \tag{3.2}$$

where $\hat{p}_n$ is an estimator of p.

This relation suggests that, we need to estimate p and ψ and deduce an estimator of α from (3.2). In the rest of this section we concentrate in estimating $\psi(t)$ and p by the methods of moments and maximum likelihood using appropriate functions of the sequence $\{X_n\}$ defined by (1.5). A moment estimator of $\psi(t)$ based on $\{Z_n(t)\}$ defined by (2.1) becomes

$$\hat{\psi}_n(t) = \{1-\overline{Z}_n(t)\} / \overline{Z}_n(t), \tag{3.3}$$

where $\overline{Z}_n(t) = (Z_1(t)+Z_2(t)+\ldots+Z_n(t))/n$.

As we have already noted, p is the first order autocorrelation coefficient of $\{Z_n(t)\}$. Hence p can be estimated by the sample auto-correlation. However, for studying the properties of the estimator, we need an explicit form of $\psi(.)$. For estimating p in the general setup, we define another sequence $\{U_n\}$ of r.v.s, where

$$U_n = \begin{cases} 1 & \text{if } X_{n+1} > X_n \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

Clearly, $\{U_n\}$ is also a stationary and ergodic sequence, with Pr $[U_n = 1] = (1+p)/2$. Moreover $\{U_n\}$ is α-mixing with mixing coefficients α_{h-1}, where α_h is given by (2.8). Now the moment estimator of p based on U_n is given by

$$\hat{p}_n = 2\overline{U}_n - 1. \tag{3.5}$$

In order to study the asymptotic properties of $(\hat{\psi}_n, \hat{p}_n)$ we need to prove the following theorem.

Theorem 3.1: As $n \to \infty$, the random vector $(\overline{Z}_n(t), \overline{U}_n) \sim AN(\mu_1, \frac{1}{n}\Sigma_1)$,

where $\mu_1 = \left(\frac{1}{1+\psi(t)}, \frac{1+p}{2}\right)$, $\Sigma_1 = \begin{bmatrix} \sigma_1^2 & (\sigma_{12}+\sigma_{21})/2 \\ (\sigma_{12}+\sigma_{21})/2 & \sigma_2^2 \end{bmatrix}$

and

$$\sigma_1^2 = [\{\psi(t)\}/\{1+\psi(t)\}^2][(1+p)/(1-p)], \tag{3.6}$$

$$\sigma_2^2 = \frac{1-p^2}{4} + \sum_{h=1}^{\infty} \frac{(1-p)^2}{2\left(1-p^{h-1}\right)^3} p^{h-1} \left\{p^{2h-2} - p^{h-1}\log p^{h-1} - 1\right\} \tag{3.7}$$

$$\sigma_{12} = -\frac{(1-p)\psi(t)}{2[1+\psi(t)]} \left[\frac{1}{1+\psi(t)} + \sum_{h=1}^{\infty} \left\{\frac{p^h}{p^h + \psi(t)}\right\}\right] \tag{3.8}$$

$$\sigma_{21} = -\frac{(1-p)\psi(t)}{2[1+\psi(t)]} \left[\frac{1}{1+\psi(t)} - \sum_{h=1}^{\infty} \left\{\frac{p^{h-1}}{1+p^{h-1}\psi(t)}\right\}\right]. \tag{3.9}$$

Proof:

If we define

$$A_n = s_1 Z_n(t) + s_2 U_n, \quad n = 0, 1, \ldots \tag{3.10}$$

then the stationarity and ergodicity of $\{X_n\}$ implies that $\{A_n\}$ is also stationary and ergodic for any arbitrary constants s_1 and s_2. Further it can be proved that the sequence $\{A_n\}$ is α-mixing with mixing coefficients $\alpha_h^* = \alpha_{h-1}$, where α_h is given by (2.8) and it satisfies the condition (2.4). Hence by Lemma 2.1,

$$\sqrt{n}\left(\overline{A}_n - \left\{s_1\left(\frac{1}{1+\psi(t)}\right) + s_2\left(\frac{1+p}{2}\right)\right\}\right) \xrightarrow{L} N\left(0,\gamma^2\right), \tag{3.11}$$

where $\overline{A}_n = (A_1+A_2+\ldots+A_n)/n$ and

$$\gamma^2 = s_1^2\ \mathrm{Var}(Z_0) + 2\ s_1 s_2\ \mathrm{Cov}(Z_0, U_0) + s_2^2\ \mathrm{Var}(U_0) + 2\sum_{h=1}^{\infty}\Big\{s_1^2\ \mathrm{Cov}(Z_0, Z_h)$$

$$+ s_1 s_2\ \mathrm{Cov}(Z_0, U_h) + s_1 s_2\ \mathrm{Cov}(Z_h, U_0) + s_2^2 Cov(U_0, U_h)\Big\}.$$

The expressions of these covariances can be obtained using the definitions of Z_n and U_n. On simplifying the algebra and rearranging the terms we get

$$\gamma^2 = s_1^2\sigma_1^2 + s_1 s_2(\sigma_{12} + \sigma_{21}) + s_2^2\sigma_2^2,$$

where σ_i and σ_{ij} are as in (3.6)-(3.9). An application of ratio test shows that all the summations above are finite. This shows that $\sqrt{n}\left(\overline{A}_n - E(A_1)\right)$ converges in distribution to a r.v. with mean 0 and variance γ^2 for any choice of s_1 and s_2. Now an argument using the continuity theorem on characteristic function and the Cramer-Wold device for reducing multivariate convergence to that of univariate, it follows that the result of the theorem is true. Hence the proof is complete.

Corollary: The estimator $\left(\hat{\psi}_n, \hat{p}_n\right)$ is strongly consistent and asymptotically normal for (ψ, p). That is, as $n \to \infty$

$$\left(\hat{\psi}_n, \hat{p}_n\right) \sim AN\left(\mu_2, \tfrac{1}{n}\Sigma_2\right), \tag{3.12}$$

where $\mu_2 = (\psi, p)$, $\Sigma_2 = \begin{bmatrix} \sigma_1^2\{1+\psi(t)\}^4 & -(\sigma_{12}+\sigma_{21})\{1+\psi(t)\}^2 \\ -(\sigma_{12}+\sigma_{21})\{1+\psi(t)\}^2 & 4\sigma_2^2 \end{bmatrix}$

and σ_i, σ_{ij} are as before.

Proof:

Define the functions

$g_1(x_1,x_2) = (1-x_1)/x_2$ and $g_2(x_1,x_2) = 2x_1-1$, $0<x_i<1$, $i=1,2$.

Then $\left(g_1\left(\bar{Z}_n,\bar{U}_n\right), g_2\left(\bar{Z}_n,\bar{U}_n\right)\right) = \left(\hat{\psi}_n,\hat{p}_n\right)$ and the partial derivatives of $g_i(x_1,x_2)$ do not vanish at $(E(Z_1), E(U_1))$. Hence the corollary follows from Theorem 3.1 using the well-known results on functions of asymptotic normal random vectors (see Serfling (1980), p.122).

In a similar way, it can be shown that as $n \to\infty$

$$\left(\hat{\psi}_n,\hat{\alpha}_n\right) \sim AN\left((\psi,\alpha),\tfrac{1}{n}\Sigma_3\right), \tag{3.13}$$

where

$$\Sigma_3 = \begin{bmatrix} \sigma_1^2\{1+\psi(t)\}^4 & \dfrac{-\alpha(\sigma_{12}+\sigma_{21})\{1+\psi(t)\}^2}{p\log p} \\ \dfrac{-\alpha(\sigma_{12}+\sigma_{21})\{1+\psi(t)\}^2}{p\log p} & \dfrac{4\sigma_2^2\alpha^2}{\{p\log p\}^2} \end{bmatrix}.$$

We will use these results to study the properties of estimators of PAM(1) process in the next section.

In the following discussion, we obtain the maximum likelihood estimator (mle) of ψ and p using the Markov chain $\{Z_n(t)\}$. The transition probabilities $\{p_{ij}, i,j = 0,1\}$ of $\{Z_n(t)\}$ are given by (2.2). In what follows we write Z_n in the place of $Z_n(t)$. It is well-known that the mle's $\tilde{p}_{ij}$ of p_{ij} based on the realization $(Z_1,Z_2,\ldots,Z_n)$ are given by

$$\tilde{p}_{ij} = \frac{n_{ij}}{n_i}, \qquad i,j=0,1, \tag{3.14}$$

where n_{ij} is the frequency of one-step transitions from state i to stage j and $n_i=\sum_j n_{ij}$, $n_j=\sum_i n_{ij}$. As we have a two-state Markov chain on $\{0,1\}$, $n_1=\sum_{j=1}^{n} Z_j$ and $n_0 = n-n_1$. It is also readily verified that as $n \to\infty$, $\tilde{p}_{ij} \xrightarrow{a.s} p_{ij}$, $i, j =0,1$ and

$$\sqrt{n}\left(\tilde{p}_{01}-p_{01},\ \tilde{p}_{10}-p_{10}\right) \xrightarrow{L} N_2(0,\Lambda), \tag{3.15}$$

where $N_2(0,\Lambda)$ denotes a bivariate normal vector with mean **0** and dispersion matrix

$$\Lambda = \begin{pmatrix} \mu & 0 \\ 0 & \theta \end{pmatrix}, \tag{3.16}$$

$$\mu = \frac{(1-p)\{p+\psi(t)\}}{\psi(t)\{1+\psi(t)\}} \quad \text{and} \quad \theta = \frac{(1-p)\psi(t)\{1+p\psi(t)\}}{\{1+\psi(t)\}}.$$

Note that, since $\tilde{p}_{00} = 1-\tilde{p}_{01}$, $\tilde{p}_{11} = 1-\tilde{p}_{10}$, the asymptotic distribution of any function of $\tilde{p}_{ij}$, $i,j=0,1$, can be obtained using (3.15).

The mle $(\tilde{\psi}_n, \tilde{p}_n)$ of (ψ, p) can be expressed as a function of $\tilde{p}_{ij}$, i,j=0,1, using (2.2) and then we have the following theorem.

Theorem 3.2: The estimator $(\tilde{\psi}_n, \tilde{p}_n)$ is strongly consistent and asymptotically normal for (ψ, p).

Proof:

By (2.2), we note that the mle's of ψ and p may be expressed as functions of $\tilde{p}_{01}$ and $\tilde{p}_{10}$ as

$$\tilde{\psi}_n(t) = \tilde{p}_{10}/\tilde{p}_{01} \quad \text{and} \quad \tilde{p}_n = 1 - \tilde{p}_{01} - \tilde{p}_{10}. \tag{3.17}$$

Obviously $(\tilde{\psi}_n, \tilde{p}_n)$ is strongly consistent for (ψ, p). Now using the results on functions of CAN estimators, it is easy to show that

$$(\tilde{\psi}_n, \tilde{p}_n) \sim \text{AN}\left((\psi,p), \tfrac{1}{n}\Lambda_1\right), \tag{3.18}$$

where

$$\Lambda_1 = \begin{bmatrix} \psi(t)\{1+\psi(t)\}^2\left(\dfrac{1+p}{1-p}\right) & p\{1-\psi^2(t)\} \\ p\{1-\psi^2(t)\} & \dfrac{p(1-p)\{1+\psi^3(t)\}}{\psi(t)\{1+\psi(t)\}} + (1-p) \end{bmatrix}.$$

This completes the proof.

It can be noted that the asymptotic variances of $\hat{\psi}_n$ and $\tilde{\psi}_n$ remain same.

Having estimated p by $\hat{p}_n$ or $\tilde{p}_n$ an estimate of α can be obtained from (3.2), Further it easily follows that as $n\to\infty$,

$$\hat{\alpha}_n \sim \text{AN}\left(\alpha, \frac{4\sigma_2^2\alpha^2}{n(p\log p)^2}\right)$$

and

$$\tilde{\alpha}_n \sim \text{AN}\left(\alpha, \frac{\lambda_{22}\alpha^2}{n(p\log p)^2}\right), \tag{3.19}$$

where λ_{22} is the $(2,2)^{\text{th}}$ element of Λ_1. It is also straight forward to show that $(\tilde{\psi}_n, \tilde{\alpha}_n)$ is asymptotically normal and the asymptotic dispersion matrix can be easily computed as in the case of $(\hat{\psi}_n, \hat{\alpha}_n)$. In the next section we concentrate on estimating PAM(1) process.

4. PARAMETRIC ESTIMATION OF PAM(1) PROCESS

We have already noted in Section 1 that the Pareto process defined by Yeh et al.(1988) is a special case of the semi-Pareto process (1.5), where

$$\psi(t) = (t/\sigma)^{\alpha},\ \sigma>0, \alpha>0. \tag{4.1}$$

It may be recalled that Yeh et al.(1988) have suggested an estimation method, where $k = p^{-1/\alpha}$ can be determined exactly in a long enough realization of $\{X_n\}$ and then σ can be estimated by identifying the innovations. They also proposed moment estimators of the parameters, but their properties are not established.

In this section we study the properties of estimators of σ using the results from previous sections. We continue to estimate k by $\hat{k}_n$ as it does not depend on a particular distribution of X_n and hence the estimator of α also remains same as before (see (3.2)). However, estimator of α depends on that of p. An estimator $\hat{\sigma}_n$ of σ based on the moment estimator $\hat{\psi}_n(t)$ defined by (3.3) is

$$\hat{\sigma}_n = t\{\hat{\psi}_n(t)\}^{1/\hat{\alpha}_n}. \tag{4.2}$$

Note that $\hat{\sigma}_n$ is a function of $(\hat{\psi}_n(t), \hat{\alpha}_n)$ and hence using (3.13) it is readily proved that

$$\hat{\sigma}_n \sim \mathrm{AN}\left(\sigma, \delta_1^2 / n\right), \tag{4.3}$$

where

$$\delta_1^2 = \sigma_1^2 \left\{1+\left(\frac{t}{\sigma}\right)^{\alpha}\right\}^4 \left(\frac{\sigma}{\alpha}\right)^2 \left(\frac{t}{\sigma}\right)^{2\alpha} + 4\alpha^2\sigma_2^2(\sigma/\alpha)^2 \frac{\{\log(t/\sigma)\}^2}{\{p\log p\}^2}$$

$$+\ 2\alpha(\sigma_{12}+\sigma_{21}) \left\{1+\left(\frac{t}{\sigma}\right)^{\alpha}\right\}^2 \left(\frac{\sigma}{\alpha}\right)^2 \left(\frac{\sigma}{t}\right)^{\alpha} \left\{\log\left(\frac{t}{\sigma}\right)\right\} / \{p\log p\}.$$

Similarly, we can get the mle $\tilde{\sigma}_n$ of σ from $\tilde{\psi}_n(t)$ and its asymptotic variance can be obtained from Λ_1 given by (3.18).

As $\psi(.)$ has the form (4.1), we can write the likelihood function based on $(Z_0, Z_1, \ldots, Z_n)$ and obtain the mles. By omitting the term corresponding Z_0, the log-likelihood function may be expressed as

$$\log \mathrm{L} = \sum_i \sum_j n_{ij} \log p_{ij},$$

where p_{ij}'s are defined by (2.2) with $\psi(t) = (t/\sigma)^{\alpha}$.

The likelihood equations with respect to σ, α and p show that the parameters σ and α are not identifiable. Hence we consider the mle of (σ, p) and deduce an estimator of α from (3.2). Thus the mle of (σ, p) can be obtained by solving the following equations:

$$\frac{n_{00}(t/\sigma)^{\alpha}}{p+(t/\sigma)^{\alpha}} + n_{10} + \frac{n_{11}p(t/\sigma)^{\alpha}}{1+p(t/\sigma)^{\alpha}} - \frac{n(t/\sigma)^{\alpha}}{1+(t/\sigma)^{\alpha}} = 0 \tag{4.4}$$

$$\frac{n_{00}}{p+(t/\sigma)^{\alpha}} + \frac{n_{11}(t/\sigma)^{\alpha}}{1+p(t/\sigma)^{\alpha}} - \frac{n_{01}+n_{10}}{1-p} = 0. \tag{4.5}$$

The solution of these equations do not have closed form expressions and one has to go for numerical methods. Let (α_n^{*}, p_n^{*}) be the resulting mle. On applying the well-known results on asymptotic properties of mle's under certain regularity conditions (cf. Basawa and Prakasa Rao (1980), p.56-57), it is proved that

$$(\alpha_n^{*}, p_n^{*}) \sim \mathrm{AN}\left((\alpha, p), \tfrac{1}{n}F^{-1}\right),$$

where

$$F^{-1} = \begin{bmatrix} \left\{1+\left(\frac{t}{\sigma}\right)^{\alpha}\right\}^{2}\left(\frac{\sigma}{t}\right)^{\alpha}\left(\frac{\sigma}{\alpha}\right)^{2}\left\{\frac{1+p}{1-p}\right\} & p\left\{1-\left(\frac{t}{\sigma}\right)^{2\alpha}\right\}\left(\frac{\sigma}{t}\right)^{\alpha}\left(\frac{\sigma}{\alpha}\right) \\ p\left\{1-\left(\frac{t}{\sigma}\right)^{2\alpha}\right\}\left(\frac{\sigma}{t}\right)^{\alpha}\left(\frac{\sigma}{\alpha}\right) & (1-p)\left\{1+\frac{p\left[1+\left(\frac{t}{\sigma}\right)^{3\alpha}\right]}{\left(\frac{t}{\alpha}\right)^{\alpha}\left[1+\left(\frac{t}{\sigma}\right)^{\alpha}\right]}\right\} \end{bmatrix}$$

Remark: It is also readily verified that, if we deduce an estimator $\tilde{\sigma}_n$ from $\tilde{\psi}_n(t) = (t/\tilde{\sigma}_n)^{\tilde{\alpha}_n}$, then as n $\to\infty$, $(\tilde{\sigma}_n, \tilde{p}_n) \sim \mathrm{AN}\left((\sigma, p), \tfrac{1}{n}F^{-1}\right)$.

That is, (σ_n^{*}, p_n^{*}) and $(\tilde{\sigma}_n, \tilde{p}_n)$ have the same asymptotic dispersion matrices. Thus for all practical purposes we may use $(\tilde{\sigma}_n, \tilde{p}_n)$ in the place of (σ_n^{*}, p_n^{*}) when the sample size is large. These asymptotic results can be used to obtain asymptotic tests and confidence intervals for the parameters.

ACKNOWLEDGEMENTS

I thank the referee for his suggestions and comments on the earlier version of this manuscript

BIBLIOGRAPHY

Adke, S.R. and Balakrishna, N (1992). "Estimation of the mean of some stationary Markov sequences," *Commun. Statist. Theory Meth.*, 21(1), 137-159.

Arnold, B.C. and Hallett, J.T.(1989). "A characterization of the Pareto process among stationary stochastic processes of the form $X_n = c \min(X_{n-1}, Y_n)$," *Statistics and Probability Letters*, 8,

Basawa, I.V and Prakasa Rao, B.L.S (1980). *Statistical Inference for Stochastic Processes*, London Academic Press.

Billard,L. and Mohamed,F.Y.(1991). "Estimation of the parameters of an EAR(p) process," *J. of Time Ser. Anal.*, 2(3), 179-192.

Billingsley, P.(1986). *Probability and Measure* (2nd edition) New York: John Wiley & Sons.

Du Jin-Guan and Li Yuan (1991). "The integer-valued autoregressive (INAR(p)) model," *J. of Time Ser. Anal.*, 12(2), 129-142.

Karlsen, H. and Tjϕstheim, D. (1988). "Consistent estimates for the NEAR(2) and NLAR(2) time series models", *J.R. statist. soc.B*, 50(2), 313-320.

Lawrance, A.J. and Lewis, P.A.W. (1985). "Modelling and residual analysis of non-linear autoregressive time series in exponential variables," *J.R. statist soc. B.*, 47(2), 165-202.

Lewis, P.A.W and Mckenzie, E.D. (1991). "Minification processes and their transformations," *J.Appl.Prob.*, 28, 45-57.

Pillai, R.N(1991). "Semi-Pareto processes," *J. Appl. Prob.*, 28, 461-465.

Raffery, A.E. (1981). "Unprocesses autoregressifa coi marginal exponential; properties asymptotique and estimation de maximum of de vrasemblance," *Annales Scientifiques de L'University de Olermont*. 69,149-160.

Raffery,A.E.(1980). "Estimation efficace pour un processes autoregressif exponential a densite discontinue," *Put.Inst. stat. Unive Paris xxv fase* 1-2, 69-91.

Serfling,R.J.(1980). *Approximation theorems of Mathematical Statistics*, New York: John Wiley & sons.

Smith, R.L.(1986). "Maximum likelihood estimation for the NEAR(2) model," *J.R. Statist. soc, B*. 48(2), 251-257.

Yeh,H.C., Arnold,B.C. and Robertson,C.A.(1988). "Pareto processes," *J.Appl. Prob.*, 25, 291-301.

Tavares, L.V.(1980). "An exponential Markovian stationary processes," *J.Appl. Prob*. 17, 1117-1120.

Received March, 1997; Revised March, 1998.