DEPARTMENT OF STATISTICS
UNIVERSITY OF CALICUT

Dr. K. JAYAKUMAR.
Lecturer (Sr. Scale)

CALICUT UNIVERSITY (P. O.)
KERALA, INDIA PIN: 673 635

December 10, 2007

Certificate

This is to certify that the work reported in this Thesis entitled **'STUDY ON MINIFICATION PROCESSES'** that is being submitted by Smt. Krishnarani, S.D. for the award of the Degree of **Doctor of Philosophy**, to the University of Calicut, is based on the bonafide research work carried out by her under my supervision and guidance in the Department of Statistics, University of Calicut. The results embodied in this Thesis have not been included in any other Thesis submitted previously for the award of any degree or diploma of any other University or Institution.

11/12/07

Dr. K. Jayakumar
(Supervising Teacher)

DECLARATION

I hereby declare that this Thesis entitled '**Study on Minification Processes**' submitted to the University of Calicut, for the award of the Degree of Doctor of Philosophy under the Faculty of Science, is an independent work done by me under the supervision and guidance of Dr. K. Jayakumar in the Department of Statistics, University of Calicut.

I also declare that this Thesis contains no material which has been accepted for the award of any other degree or diploma of any University or Institution and to the best of my knowledge and belief, it contains no material previously published by any other person, except where due references are made in the text of the Thesis.

Krishnarani.S.D

Calicut University Campus

December 10, 2007

KRISHNARANI, S.D.

ACKNOWLEDGEMENT

First of all, I thank God, the Almighty for the boundless blessings, which led me to the completion of this work overcoming all difficulties.

I am greatly indebted to Dr. K. Jayakumar, Lecturer (Sr. scale), *Department of Statistics, University of Calicut for his constant encouragement and invaluable guidance during this research work. I find no words to express my gratitude for his help and deep involvement in every stage of this Thesis.*

I owe my deep gratitude to the Head and all the Faculty Members of the Department of Statistics for their extensive help during the entire span of my research work. The courteous co-operation by the Research Scholars, Office Staff, Librarian and Students of the Department of Statistics, University of Calicut is here remembered and acknowledged.

I am thankful to all my colleagues and friends at Bhavan's Vidyashram, Chelembra, SAFI Institute of Advanced Studies, Vazhayoor and Govt. Rajah's Higher Secondary School, Kottakkal for the numerous way in which they helped me during the days of this research work.

I extend my personal thanks to all my family members for their moral support and good wishes throughout my research work.

Last but not least, I am most grateful to my beloved husband –Dr. M. Manoharan, my son-Gautham and my daughter – Gayathri for their support and patience throughout the period of my research.

KRISHNARANI, S.D.

CONTENTS

CHAPTER I

INTRODUCTION

CHAPTER I

INTRODUCTION

1.1 Introduction

Observations are frequently made sequentially over time in order to understand the world around us. Values in the future depend, usually in a random manner, on the observations available at present. Such a dependence, called the Markov dependence, makes it worthwhile to predict the future from its past. Our task is to depict the underlying dynamics from which the observed data are generated. This will help us to forecast and possibly control future events. Motivated by this we study in this Thesis some probability models for time series processes.

A *time series* is a set of observations generated sequentially in time. The primary objective of time series analysis is to develop mathematical models that provide plausible descriptions for sample data. A model is selected in such a way that the selected model is a simple one and reasonably reflects the physical law that governs the data. Examples of time series occur in a variety of fields ranging from economics to engineering. Share prices on successive days, average incomes in successive months, company profit in successive years, rainfall on successive days, monthly unemployment figures and air temperature measured in successive hours, days or months etc. are common examples. In medicine, blood pressure measurement traced over time could be useful for evaluating drugs used in treating hypertension. An epidemiologist might be interested in the number of influenza cases observed over some time period.

The special feature of time series analysis is the fact that successive observations are usually not independent. When successive observations are dependent, future values may be predicted from past observations. If a time series can be predicted exactly, it is said to be a deterministic time series. But in most cases, the future is only partly determined by the past observations. Such time series are called stochastic time series. In that case, exact predictions are impossible and therefore the future values have a probability distribution, which is conditioned by knowledge of past values. Time series analysis is concerned with techniques for the analysis of this dependence. This requires the development of stochastic and dynamic models for time series data. Finding a suitable model for a given time series depends on a number of factors, including the properties of the series as assessed by a visual examination of the data, the number of observations available and the way the model is to be used. It is important to understand the three main stages in model building, which can be described as model formulation, estimation and diagnostic checking.

The objectives of time series analysis are diverse, depending on the background applications. Statisticians usually view a time series as a realization from a stochastic process. A fundamental task is to unveil the probability law that governs the observed time series. With such a probability law, we can understand the underlying dynamics, forecast future events, and control future events via intervention. Those are the three main objectives of time series analysis.

When the observations are partly determined by its past values, the model can be written as,

$$X_t = f_t + \varepsilon_t,$$

where $X_t,\ t = 1,2,\cdots$ are observations on time series made at t equally distant time points. The observed series is made up of a completely determined sequence $\{f_t\}$, which is called the *systematic part*, and of a random sequence $\{\varepsilon_t\}$, which obeys a probability law and $\{\varepsilon_t\}$ is called an *innovation sequence* or a *white noise process*.

In the present work we construct several autoregressive minification models yielding useful stationary marginal distributions. The marginal distributions of the usual autoregressive models can be obtained only if the generating functions have closed form. But we can use minification models as an alternative to the usual additive model, which can generate given distribution as marginal. Also Weibull or extreme value random variables are commonly used for modeling run-off series. But processes with these marginal distributions cannot be generated with linear models. So minification processes are important as a source of time series for such processes.

In Chapter II, a new class of autoregressive processes with minification structure is constructed, which can generate autoregressive minification models with a given distribution as marginal. A necessary and sufficient condition for the process to be stationary is given and properties of the class are studied. Several examples are given. The new process so defined is extended to higher orders. The corresponding moving average and autoregressive moving average models are also given. We develop another class of autoregressive process, called autoregressive Lehmann process, which can generate any distribution from the Lehmann family. Properties of the same are discussed. Estimation of the parameters is also discussed.

In Chapter III, we introduce generalized half semi-logistic distribution generated through the Marshall-Olkin form. The autoregressive minification process with generalized half semi-logistic distribution as marginal is constructed and properties are given. It is also shown that Pareto, Weibull, exponential and half logistic distributions belong to this class of distributions.

Chapter IV deals with a generalized autoregressive minification process. This newly introduced process is a generalization of the processes we discussed in the previous Chapters. A necessary and sufficient condition for this new process to be stationary is given. Properties are studied and examples are given. Higher order processes are constructed. Generalized half semi-logistic process and generalized semi-logistic process are discussed as particular cases. Estimation of the parameters of the process is also done.

Random coefficient first order autoregressive minification model is constructed in Chapter V. Again a generalized random coefficient first order process is introduced and properties are studied. Also random coefficient autoregressive Lehman process is constructed and studied.

In Chapter VI, bivariate semi-logistic and half semi-logistic distributions are introduced and studied. The first order minification processes with these distributions as marginals are also discussed. Further a general process useful in generating bivariate autoregressive minification processes with a given distribution as marginals is introduced and studied.

Chapter VII gives an overall summary of the Thesis and concluding remarks.

1.2 Some Distributions Used in the Thesis

In this Section, we briefly present some distributions used in this Thesis.

Definition 1.2.1: A random variable X on $(0, \infty)$ is said to have semi-*Pareto distribution* and write $X \underline{\underline{d}}\ SP(\beta, p)$, if its survival function is

$$\overline{F}(x) = P(X > x) = \frac{1}{1+\psi(x)}, x \geq 0, \tag{1.2.1}$$

where $$\psi(x) = \frac{1}{p}\psi\left(p^{\frac{1}{\beta}}x\right), \ \beta > 0, \quad 0 < p < 1. \tag{1.2.2}$$

□

The solution of the functional equation (1.2.2) is $\psi(x) = x^{\beta} h(x)$,

where h(x) is periodic in $\ln x$ with period $\frac{2\pi\beta}{(-\ln p)}$.

If $\psi(x) = \left(\frac{x}{\sigma}\right)^{\beta}$, $\sigma > 0$ then we get a *Pareto type III distribution* denoted by $P(\sigma, \beta)$ with survival function

$$\overline{F}(x) = \frac{1}{1+\left(\frac{x}{\sigma}\right)^{\beta}} \quad , \quad \sigma, \beta > 0 \,;\, x > 0. \tag{1.2.3}$$

Definition 1.2.2: We say that random variable X on $(-\infty, \infty)$ has *semi-logistic distribution* and write $X \underset{=}{d} SL(\beta, p)$, if it has the survival function

$$\overline{F}(x) = \frac{1}{1+\psi(x)} \quad , \tag{1.2.4}$$

where $\psi(x)$ satisfies the functional equation

$$\psi(x) = \frac{1}{p}\psi\left(\frac{1}{\beta}\ln p + x\right). \tag{1.2.5}$$

□

The solution of the functional equation is

$$\psi(x) = e^{\beta x} h(x), \tag{1.2.6}$$

where $h(x) = h\left(\frac{1}{\beta}\ln p + x\right)$. (1.2.7)

That is, h(x) is periodic in x with period $\frac{1}{\beta}\ln p$. This distribution can be used to model real data that exhibit periodic movements.

Definition 1.2.3: A random variable with support $(-\infty, \infty)$ is said to have *logistic distribution* and write $X \underline{\underline{d}}\, L(\mu, \sigma^2)$ if its density function is given by,

$$f(x,\mu,\sigma)=\frac{\pi}{\sigma\sqrt{3}}\frac{e^{-\pi(x-\mu)/\sigma\sqrt{3}}}{\left\{1+e^{-\pi(x-\mu)/\sigma\sqrt{3}}\right\}^{2}}\ ,-\infty<x,\mu<\infty,\sigma>0. \quad (1.2.8)$$

□

The corresponding survival function is

$$\overline{F}(x)=\frac{1}{1+e^{\pi(x-\mu)/\sigma\sqrt{3}}},\ -\infty<x,\mu<\infty,\ \sigma>0$$

$$=\frac{1}{2}\left\{1-\tanh\left(\frac{\pi(x-\mu)}{2\sigma\sqrt{3}}\right)\right\}.$$

Another form of the probability density function is

$$f(x,\mu,\sigma)=\frac{\pi}{4\sigma\sqrt{3}}\sec h^{2}\left(\frac{\pi(x-\mu)}{2\sigma\sqrt{3}}\right)\ ,\ -\infty<x,\mu<\infty,\ \sigma>0.$$

Because of this form logistic distribution is called sech-*squared distribution.*

When $Y=\frac{X-\mu}{\sigma}$, the probability density function of Y is

$$f(y)=\frac{\pi}{\sqrt{3}}\frac{e^{-\pi y/\sqrt{3}}}{(1+e^{-\pi y/\sqrt{3}})^{2}}\ ,\ -\infty<y<\infty$$

and the corresponding survival function is

$$\overline{F}(y)=\frac{1}{1+e^{\pi y/\sqrt{3}}},-\infty<y<\infty.$$

This is the standardized form of the *logistic distribution* denoted by L (0,1). Note that the random variable Y has mean 0 and variance 1.

But for the sake of convenience, we may take $Z=\frac{\pi(x-\mu)}{\sigma\sqrt{3}}$. Then the probability density function of Z is $f(z)=\frac{e^{-z}}{\left(1+e^{-z}\right)^2}$, $-\infty<z<\infty$ and the survival function is

$$\overline{F}(z)=\frac{1}{1+e^{z}},-\infty<z<\infty. \tag{1.2.9}$$

Then $Z\sim L\left(0,\frac{\pi^2}{3}\right)$ is used to indicate that the random variable Z distributed as *logistic* with mean 0 and variance $\frac{\pi^2}{3}$.

The logistic density function is symmetric about zero and is more peaked in the center than the normal density function. Since Z is symmetric about zero all the odd moments of Z are zero whereas the even order moments of Z are

$$E(Z^{2r})=2\Gamma(2r+1)\left(1-\frac{1}{2^{2r-1}}\right)h(2r) \quad \textit{for } r=1,2,\ldots.$$

where $h(s)=\sum_{j=1}^{\infty}j^{-s}$ is the Riemann zeta function. Also the coefficient of skewness of Z is zero and coefficient of kurtosis is 4.2. This means that the logistic distribution has longer tails than the normal distribution. By comparing the cumulative distribution functions of the standard normal variable and the standard logistic variable Johnson et al. (2004), showed that the two are very close and the maximum value of the difference is about 0.0228.

Remark 1.2.1: When $\psi(x) = e^x$ in (1.2.4) we get *logistic distribution* with survival function of the form (1.2.9).

Now we consider some generalizations of the logistic distribution namely, *Type I and Type II generalized logistic distributions*.

Definition 1.2.4: We say that a random variable X on $(-\infty, \infty)$ has the *Type I generalized logistic distribution* if it has the survival function,

$$\overline{F}_1(x) = 1 - \frac{1}{(1+e^{-x})^{\gamma}}, \quad \gamma > 0. \tag{1.2.10}$$

□

Note that the Type I distribution coincides with the logistic distribution when $\gamma = 1$.

Definition 1.2.5: A random variable X on $(-\infty, \infty)$ has the *Type II generalized logistic distribution* if it has the survival function,

$$\overline{F}_2(x) = \left(\frac{1}{1+e^{x}}\right)^{\gamma}, \quad \gamma > 0. \tag{1.2.11}$$

□

If Y is a random variable distributed as Type I, then –Y has a Type II distribution.

Definition 1.2.6: A random variable X is said to follow *folded logistic distribution* on $(0, \infty)$ if it has the survival function,

$$\bar{F}(x) = \frac{1}{1+e^{\left(\frac{x-\mu}{\sigma}\right)}} + \frac{1}{1+e^{\left(\frac{x+\mu}{\sigma}\right)}} \tag{1.2.12}$$

$$= 1 - \frac{1}{2}\left[\tanh\left(\frac{x-\mu}{2\sigma}\right) + \tanh\left(\frac{x+\mu}{2\sigma}\right)\right], x \geq 0, -\infty < \mu < \infty, \sigma > 0. \quad \square$$

If Y is a logistic random variable with density

$g(y) = \dfrac{e^{\frac{y-\mu}{\sigma}}}{\sigma(1+e^{\frac{y-\mu}{\sigma}})^2}$, $-\infty < y < \infty, -\infty < \mu < \infty, \sigma > 0$, then (1.2.12) is the survival function of the random variable $X = |Y|$.

Cooray et al. (2006) considered the folded logistic probability density and estimated the parameters of the same. They discussed a number of situations where this distribution arises.

Definition 1.2.7: A random variable X on $(0, \infty)$ is said to follow the *half logistic distribution* if its survival function is,

$$\bar{F}(x) = \frac{2}{1+e^{x}}, x \geq 0. \tag{1.2.13}$$

$\square$

When $\mu = 0$ and $\sigma = 1$ in (1.2.12), we get (1.2.13).

Definition 1.2.8: A random variable X on $(0, \infty)$ is said to have *half semi-logistic distribution* and write $X \underset{=}{d} HSL(\beta, p)$ if it has the survival function

$$\overline{F}(x) = \frac{2}{1+\psi(x)}, \tag{1.2.14}$$

where $\psi(x)$ satisfies the functional equation (1.2.5). □

Definition 1.2.9: A random variable X on $(0,\infty)$ is said to have Type 1 generalized half semi-logistic distribution if it has the distribution function

$$F(x) = \left(\frac{\psi(x)-1}{\psi(x)+1}\right)^{\gamma}, \gamma > 0, \tag{1.2.15}$$

where $\psi(x)$ satisfies the same functional equation (1.2.5). □

Definition 1.2.10: A random variable X on $(0,\infty)$ is said to have Type II generalized half semi- logistic distribution if it has the survival function

$$\overline{F}(x) = \left(\frac{2}{1+\psi(x)}\right)^{\gamma}, \quad \gamma > 0, \tag{1.2.16}$$

where $\psi(x)$ satisfies the functional equation (1.2.5). □

1.3 Models of Time Series

1.3.1 Autoregressive Processes

Many time series are well approximated by the representation,

$$X_t = \lambda_1 X_{t-1} + \lambda_2 X_{t-2} + \cdots + \lambda_p X_{t-p} + \varepsilon_t, \tag{1.3.1}$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with mean 0 and variance σ^2 and $\lambda_1, \lambda_2, \ldots, \lambda_p$ are constants. Then the process $\{X_t\}$ is said to be an autoregressive process of order p and we write $\{X_t\} \sim AR(p)$. Here X_t is regressed on its past p values and hence it is called an autoregressive process of order p. The model is easy to implement and therefore is the most popular time series model in practice.

An autoregressive process of order 1 is of the form, $X_t = \lambda X_{t-1} + \varepsilon_t$. Then the autocovariance function is

$$Cov(X(t), X(t+h)) = \frac{\sigma^2 \lambda^h}{1-\lambda^2},$$

and the autocorrelation function of order h is

$$\rho(h) = Corr(X(t), X(t+h)) = \lambda^h, \quad h > 0.$$

1.3.2 Moving Average Processes

Suppose that $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with mean zero and variance σ^2 and $\rho_1, \rho_2, \ldots, \rho_q$ are constants. Then a process $\{X_t\}$ is said to be moving average process of order q denoted by MA(q) if,

$$X_t = \varepsilon_t + \rho_1 \varepsilon_{t-1} + \cdots + \rho_q \varepsilon_{t-q}, \tag{1.3.2}$$

where $\{\rho_i\}$ are constants.

MA model expresses a time series as a moving average of a white noise process or innovation process. The correlation between X_t and X_{t-h} is due to the fact that they may depend on the same ε_{t-j}'s. Obviously X_t and X_{t-h} are uncorrelated when $h > q$. Advantage of MA models lies in their theoretical tractability and it is very easy to generate.

The autocorrelation function is,

$$\begin{aligned} \rho(h) = Corr(X(t), X(t+h)) &= \rho_q \rho_{q-h} + \ldots + \rho_{h+1}\rho_1 + \rho_h \quad && if \quad h \le q \\ &= \quad 0 && if \quad h > q. \end{aligned}$$

1.3.3 Autoregressive Moving Average Process

A useful class of models for time series is formed by combining MA(q) and AR(p) processes. It is said to be an autoregressive moving average process given by,

$$X_t = \lambda_1 X_{t-1} + \cdots + \lambda_p X_{t-p} + \varepsilon_1 + \rho_1 \varepsilon_{t-1} + \cdots + \rho_q \varepsilon_{t-q}, \tag{1.3.3}$$

where $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with mean zero and variance σ^2. Here $p, q \ge 0$ are integers and (p, q) is called the order of the model. It is denoted by ARMA(p, q). Using the back shift operator, the model can be written as

$$b(B)X_t = a(B)\varepsilon_t,$$

where B denotes the back shift operator, which is defined as

$$B^k X_t = X_{t-k}, k = 1, 2, \ldots$$

and a(.) and b(.) are polynomials defined as

$$b(z) = 1 - b_1 z - \cdots - b_p z^p, \; a(z) = 1 + a_1 z + \cdots + a_q z^q.$$

ARMA models are one of the most frequently used families of parametric models in time series analysis. This is due to their flexibility in approximating many stationary processes.

1.3.4 Autoregressive Integrated Moving Average Process

Real time series data often exhibit some time trend (such as slowly increasing or cyclic features) that is beyond the capacity of ARMA models. The common practice is to preprocess the data to remove those unstable components. Taking difference is a convenient and effective way to detrend. After removing time trends, we can model the new and remaining series by a stationary ARMA model. Since the original series is the integration of the differenced series, it is called an autoregressive integrated moving average process (ARIMA).

A time series $\{Y_t\}$ is called an autoregressive integrated moving average process with order p, d and q, denoted as $\{Y_t\} \sim ARIMA(p,d,q)$, if its d-order difference, $X_t = (1-B)^d Y_t$ is a stationary ARMA(p, q) process, where $d \geq 1$ is an integer namely, $b(B)(1-B)^d Y_t = a(B)\varepsilon_t$.

1.4 Stationarity and Markov Property

Statistical inference is about learning something that is unknown from the known. In order to achieve this, it is necessary to assume that at least some features of the underlying probability law are sustained over a period of interest. This leads to the assumption of different types of stationarity, depending on the nature of the problem at hand. We give two types of stationarity namely, weak stationarity and strict stationarity.

Definition 1.4.1: A time series $\{X_t, t = 0, \pm 1, \pm 2, ...\}$ is weakly stationary if $E\left(X_t{}^2\right) < \infty$ for each t,

i) $E(X_t)$ is a constant, independent of t and

ii) $\mathrm{Cov}(X_t, X_{t+k})$ is independent of t for each k. □

Definition 1.4.2: A time series $\{X_t, t = 0, \pm 1, \pm 2, ...\}$ is strictly stationary if $(X_1, X_2, ..., X_n)$ and $(X_{1+k}, X_{2+k}, ..., X_{n+k})$ have the same joint distribution for any integer $n \geq 1$ and any integer k. In other words shifting the time origin by an amount k has no effect on the joint distributions. □

Note: A strict stationary time series with finite second moment is widely stationary. But a wide sense stationary time series will not necessarily be strictly stationary.

It is relatively easy to check stationarity in linear time series models. But it is by no means easy to check whether a time series defined by a nonlinear model is strictly stationary. The common practice is to represent a time series as a Markov chain and to establish that the Markov chain is ergodic. Stationarity follows from the fact that an ergodic Markov chain is stationary. So we give a brief introduction of Markov chains.

Definition 1.4.3: A sequence of scalar variables $\{X(t), 0 < t < \infty\}$, is said to be a Markov process on a state space S, if for every $n \geq 2$,

$$0 \leq t_1 < t_2 < \cdots < t_n < t < \infty, \quad i_1, i_2, \cdots, i_n, j \in S,$$

$$P(X(t) = j / X(t_1) = i_1, \cdots, X(t_n) = i_n) = P(X(t) = j / X(t_n) = i_n). \qquad \square$$

The Markovian property requires that, given the present and the past, the future depend on the present only. The conditional probability of X_{t+1} given X_t is called the transition probability function at time t. If the transition probability function is independent of time t, the Markov chain is called homogeneous.

1.5 Gaussian and Non-Gaussian Time Series Models

A sequence $\{X_n, n \in Z\}$ is a Gaussian time series if for n ≥ 1, the joint distribution of $X_{n1}, X_{n2}, ..., X_{nk}$ is a k-variate normal. This multivariate normal distribution is completely characterized by its first and second moments and so it follows that second order stationarity implies strict stationarity for normal processes and vice versa.

The models used in classical analysis of time series are linear in nature. Moreover the time series $\{X_n\}$ is assumed to be Gaussian sequence. One of the linear stochastic models used in time series analysis is the AR(p) model defined by (1.3.1). The modeling and analysis of time series in the classical set up heavily depends on the assumption that the series is a realization from a Gaussian sequence. However most of the series we come across in practical situations are far from Gaussian and hence a study of such non-Gaussian models is of interest.

In modeling of non-Gaussian models, the standard technique is to make a suitable transformation to remove the skewness of the data and then fit a Gaussian model. But the stringent condition that the transformed sequence must be Gaussian is very unlikely to be true in practice (see Sim (1990)). As a result, the time series that does not fit to the Gaussian set up needs a separate treatment.

A number of different models have been constructed for the generation of non-Gaussian time series. For example, Gaver and Lewis (1980), Jacobs and Lewis (1977), Lawrance and Lewis (1977, 1980, 1981) and McKenzie (1986). The need for such models arises from the fact that many naturally occurring time series are clearly non-Gaussian. Some non-Gaussian time series models are given below.

Lawrance and Lewis (1977) considered a stationary sequence of random variables $\{X_n\}$, which are formed from an i.i.d. exponential sequence $\{\varepsilon_n\}$ according to the linear model,

$$X_n = \begin{cases} \rho\,\varepsilon_n & w.p. \quad \rho \\ \rho\,\varepsilon_n + \varepsilon_{n-1} & w.p. \quad (1-\rho) \end{cases} \tag{1.5.1}$$

(w.p. stands for with probability). Then $\{X_n\}$ form a sequence of exponential random variables. This is the first order exponential moving average process. Also Lawrance and Lewis (1980) extended this model to the qth order, called exponential moving average (EMA(q)) model.

Gaver and Lewis (1980) considered the AR(1) model

$$X_n = \lambda\, X_{n-1} + \varepsilon_n,\ \ n = 0,1,2,\ \dots. \tag{1.5.2}$$

They have shown that if X_n's are exponentially distributed then ε_n's can be written as $\varepsilon_n = I_n E_n$, where I_n's are Bernoulli λ and E_n's are exponentially distributed. Then the model (1.5.2) is called exponential autoregressive of order 1 (EAR(1)). They have also shown that if X_n' s are gamma (ν, k) variables then the Laplace Stieltjes transform of ε_n is given by,

$$\phi(s) = \left(\lambda + (1-\lambda)\frac{\nu}{\nu+s}\right)^k. \tag{1.5.3}$$

This is the gamma autoregressive process. The simplicity of the EAR(1) allows one to model in an intuitive way dependencies in stochastic systems.

Jayakumar and Pillai (1993) introduced the first order autoregressive semi-Mittag-Leffler process (SMLAR(1)) and they have shown that EAR(1) process of Gaver and Lewis (1980) is a special case of SMLAR(1) process.

Sim (1986) has given the model

$$X_n = V_n X_{n-1} + \varepsilon_n \tag{1.5.4}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. exponential random variables with parameter $\theta > 0$, $\{V_n, \quad n = 1,2,...\}$ is a sequence of i.i.d. random coefficients with standard power function distribution $F_V(v) = v^{\beta}$, $\beta > 0$, $0 < v < 1$ and ε_n's and V_n's are mutually independent. He discussed the model in the study of hydrological modeling.

Also, Sim (1993) proposed an AR(1) model that can be used to generate logistic processes. The model he constructed has the structure (1.5.2) where $|\lambda| < 1$. He has shown that if $\{X_n\}$ is stationary with L(0,1) marginal distribution then $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables with probability density function

$$f_{\varepsilon}(z) = \frac{\sin(\lambda\pi)}{2\lambda\pi[\cosh(z) + \cos(\lambda\pi)]}$$

which is symmetrical about $z = 0$ and closely resembles the logistic distribution L(0,1) as $\lambda \to 0$.

Grunwald et al. (2000) gave a general formulation of a non-Gaussian conditional linear AR(1) model subsuming most of the non-Gaussian AR(1) models that have appeared in the literature. They have organized models into broad classes to clarify similarities and differences and facilitate application in particular situations.

A mixed process, called exponential autoregressive moving average process of order (p, q) (EARMA(p, q)), incorporating aspects of both EAR(p) and EMA(q) correlation structures is developed by Lawrance and Lewis (1980). Sim (1987) developed a mixed gamma ARMA(1,1) model for generating river flow time series.

1.6 Minification Processes

Models with minification structures have been introduced in the literature as an alternative to the non-Gaussian time series models where the generating functions, usually used as a main tool for their analysis, do not have closed form expressions. In their study we assume the form of a stochastic model and the existence of a stationary sequence of random variables of specified marginal distributions under certain conditions. These minification models possess most of the properties of additive autoregressive models. The existence of such models and their properties can be easily studied using the survival function of the underlying random variable.

The study on minification process began with the work of Tavares (1980). He introduced an autoregressive minification process

$$X_n = \begin{cases} X_0 & n=0 \\ k\min(X_{n-1}, \varepsilon_n) & n \geq 1 \end{cases} \tag{1.6.1}$$

where $k>1$ is a constant and $\{\varepsilon_n\}$ is an innovation process of i.i.d. random variables such that $\{X_n\}$ is a stationary Markov process. He considered a particular case, where $\{\varepsilon_n,\ n=1,2,\ ...\}$ is a sequence of i.i.d. exponential random variables with mean $\theta(k-1)$ and X_0 is exponential with mean θ. This model generates first order autoregressive exponential process with mean θ and it is useful in hydrological applications. Because of the structure of (1.6.1) it is called minification process.

Sim (1986) developed a first order autoregressive Weibull process and studied its properties. He has shown that $\{X_n\}$ in (1.6.1) are stationary Weibull random variables with survival function $e^{\left(\frac{-\theta x^c}{k^c-1}\right)}$ if and only if $\{\varepsilon_n,\ n=1,2,\ldots\}$ is a sequence of i.i.d. Weibull random variables with survival function $e^{-\theta x^c}$.

Yeh et al. (1988) developed an autoregressive minification process having Pareto marginals analogous to the EAR(1) process of Gaver and Lewis (1980). The structure is,

$$X_n = \begin{cases} \lambda^{-1/\beta} X_{n-1} & \text{w.p.} \quad \lambda \\ \min\left(\lambda^{-1/\beta} X_{n-1},\ \varepsilon_n\right) & \text{w.p.} \quad (1-\lambda). \end{cases} \qquad (1.6.2)$$

If X_0 has the distribution $P(\sigma,\beta)$ and $\{\varepsilon_n\}$ be i.i.d. $P(\sigma,\beta)$ random variables then the process (1.6.2) is strictly stationary.

The time-reversed form of this process was introduced first in Arnold (1989) as a process based on geometric minimization. The structure of the model is,

$$X_n = \begin{cases} \min\left(X_{n-1},\ (1-\lambda)^{-1/\beta}\varepsilon_n\right) & \text{w.p.} \quad \lambda \\ (1-\lambda)^{-1/\beta}\varepsilon_n & \text{w.p.} \quad (1-\lambda). \end{cases} \tag{1.6.3}$$

If $\{\varepsilon_n\}$ is a sequence of i.i.d. $P(\sigma,\beta)$ random variables and $X_0 \sim P(\sigma,\beta)$ then the process (1.6.3) is strictly stationary $P(\sigma,\beta)$ random variables. Later in 1993, he has extended the same and developed a logistic process using Markovian (0,1) sequence.

Arnold and Robertson (1989) introduced autoregressive minification model with logistic marginals. If X_0 has logistic distribution with survival function $\frac{1}{1+e^{\beta x}}$ and $\{\varepsilon_n\}$ is a sequence of i.i.d. logistic random variables with the same survival function then $\{X_n\}$ in the model,

$$X_n = \begin{cases} X_{n-1} - \frac{1}{\beta}\ln p & \text{w.p.} \quad \text{p} \\ \min\left(X_{n-1} - \frac{1}{\beta}\ln p,\ \varepsilon_n\right) & \text{w.p.} \quad (1-\text{p}) \end{cases} \tag{1.6.4}$$

is stationary logistic.

Arnold and Hallett (1989) considered processes of the form $X_n = k\min(X_{n-1}, Y_n),\ k>0$ and verified that the associated level crossing processes, $Z_n(t) = \begin{cases} 1 & if \quad X_n > t \\ 0 & if \quad X_n \le t \end{cases}$ are Markovian for every t if and only if $\{X_n\}$ is a Pareto process.

Pillai (1991) extended the Pareto process of Yeh et al. (1988) to obtain a first order autoregressive semi-Pareto process. Such minification process has the structure same as (1.6.2).

Lewis and McKenzie (1991), discussed about minification processes and their transformations. They considered processes of the form (1.6.1) and gave a necessary and sufficient condition on the hazard rate, $h(x)=\dfrac{f(x)}{\bar{F}(x)}$ for $\{X_n\}$ to be a stationary process. Several examples are given. And also they have shown that many of the important features of the process are invariant under monotone transformation. For instance, if g is a monotone increasing transformation,

$$Y_n = g(X_n) \quad and \quad W_n = g(Z_n),$$

then (1.6.1) becomes $Y_n = g\,[Kg^{-1}\min(Y_{n-1}, W_n)]$.

Pillai et al. (1995) introduced a new class of distributions called distributions of universal geometric minima and brought out its role in defining autoregressive minification processes. Their minification model is,

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p \\ \min(X_{n-1}, \varepsilon_n) & w.p. \quad (1-p) \end{cases} \tag{1.6.5}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables such that $\{X_n\}$ is a stationary Markov process with a given marginal distribution.

Dziubdziela (1997) has done characterizations of some minification processes. Jayakumar and Mathew (2004) extended the process of Arnold and Robertson (1989) and showed that (1.6.4) has semi-logistic marginals if and only if $\{\varepsilon_n\}$ is a sequence of semi-logistic random variables.

Jayakumar and Pillai (2002) introduced a general Markov process with innovation, using which we can develop any autoregressive process of first order. The model they constructed is,

$$X_n = \begin{cases} \phi^{-1}\left(\lambda^{-1}\phi(X_{n-1})\right) & w.p. \quad \lambda \\ \min\left(\lambda^{-1}\phi(X_{n-1}), \varepsilon_n\right) & w.p. \quad 1-\lambda \end{cases}, \tag{1.6.6}$$

where $\phi(x) = \dfrac{1}{\overline{F}(x)} - 1$, and $\phi(x)$ is strictly monotone with $\phi(0) = 0$ and $\phi(\infty) = \infty$, $0 < \lambda < 1$. If $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables with common distribution function F and X_0 has distribution function F , then (1.6.6) defines a stationary sequence.

Ristic (2007) generalized the two-parameter semi-Pareto minification process of Pillai (1991) to a semi-Pareto minification process with three parameters. The model he constructed is

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad \lambda_1 \\ \lambda_2^{-1/\beta} X_{n-1} & w.p. \quad \lambda_2(1-\lambda_1) \\ \min\left(\lambda_2^{-1/\beta} X_{n-1}, \varepsilon_n\right) & w.p. \quad (1-\lambda_1)(1-\lambda_2) \end{cases} , \tag{1.6.7}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables which follows $SP(\beta, \lambda_2)$, X_{n-1} and ε_n are independent , $0 < \lambda_1, \lambda_2 < 1$, $\beta > 0$. If X_0 is distributed as ε_1 then $\{X_n\}$ in (1.6.7) is stationary with $SP(\beta, \lambda_2)$ distribution.

Many researchers worked on the higher order versions of these processes. For a detailed study on the higher order versions of the Pareto processes one may refer to Arnold (2001).

Davis and Resnick (1989) provide material relevant to the study of minification ARMA process. Several researchers discussed multivariate extensions of most of these processes. Note that in this case the innovation variables become innovation vectors.

Now we discuss integer valued time series models. Discrete variate time series occur in many contexts, often as counts of events, objects or individuals in consecutive intervals or at consecutive points in time. Some simple examples are the numbers of accidents in a manufacturing plant each month, the numbers of patients treated by a hospital's accident and emergency unit each hour, the numbers of fish caught in a particular area of sea each week, and the numbers of lifts in a tall office building which are fully operational at the start of business each day. Such data may also arise from the discretization of continuous variate time series. An example of this is the reduction of daily rainfall volumes to a binary series of ones and zeros, that is wet and dry days (see McKenzie (2003)).

One attempt to find alternative structural forms for the AR(1) process is represented by the work of Littlejohn (1992) on discrete minification processes. These are discrete analogues of the positive continuous variate minification processes of Lewis and McKenzie (1991). Kalamkar (1995) investigated the stationarity of minification processes when the marginal is a discrete distribution and obtained a necessary and sufficient condition for a discrete distribution to be the marginal of a stationary minification process.

1.7 Random Coefficient Autoregressive Models

Extensions of the model (1.3.1) have been proposed by replacing λ_i by random variables to get random coefficient autoregressive models. During the last 30 years, there has been an increasing interest in nonlinear time series models. One of the first examples is the random coefficient model introduced and studied by Nicholls and Quinn (1982). These types of time series have been used in the context of random perturbations of dynamical systems and they have found a variety of applications, for example, in finance and biology (see, Nicholls and Quinn (1982), Tong (1990)).

The sequence $\{X_n\}$ is said to follow the p[th] order random coefficient autoregressive model if,

$$X_n = \sum_{i=1}^{p}(b_i + V_{i,n})X_{n-i} + \varepsilon_n,\ n = 1,2,\dots. \tag{1.7.1}$$

The following assumptions are made on this model.

- $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables with mean 0 and variance σ^2.
- $\{Z_n = (V_{1,n}, V_{2,n}, \dots, V_{p,n})\}$ is a sequence of i.i.d. random vectors with mean 0 and dispersion matrix Γ.
- $\{\varepsilon_n\}$ and $\{Z_n\}$ are statistically independent.
- $b = (b_1, b_2, \dots, b_p)$ is a vector of real constants.

The random coefficient autoregressive model of order 1, is given by the equations

$$X_n = (V_{1,n} + b_1)X_{n-1} + \varepsilon_n,\ \ n = 1,2,\dots \tag{1.7.2}$$

where $V_{1,n}$ and ε_n are independent and identically distributed pairs of random variables.

Discussion on conditions for existence of the model can be found in Dewald and Lewis (1985), Gaver and Lewis (1980) and Sim (1986,1990). Lawrance and Lewis (1985) generated exponential random variables using random coefficient model. Also we can see uses of these types of models in Liu (1990) and Basu and Das (1993).

In the minification models given by (1.6.1) and (1.6.4), the numbers k and p associated with the minification processes are considered as the coefficients of the processes. When we allow these coefficients of the process to be random, we call such a process as a random coefficient minification process. Such models were discussed in Jayakumar and Mathew (2002).

1.8 Some Bivariate Minification Processes

Gumbel (1961) proposed two bivariate logistic distributions; one with joint distribution function $F(x,y)=\dfrac{1}{1+e^{-x}+e^{-y}}$ that appears to be a natural generalization of the univariate logistic distribution but asymmetrical, called Type I bivariate logistic density and the other one is symmetrical. A generalization of Gumbel's Type I bivariate distribution is given in Satterthwaite and Hutchinson (1978). Ali et al. (1978) studied a family of bivariate distributions with joint distribution function $F(x,y)=\dfrac{1}{1+H(x,y)}$ where $H(x,y)$ is the bivariate odds function.

Marshall and Olkin (1997) introduced a scheme for adding a new parameter to distribution F. The corresponding survival function so generated is given by

$$\bar{G}(x) = \frac{\alpha \bar{F}(x)}{1-(1-\alpha)\bar{F}(x)}, \alpha > 0. \tag{1.8.1}$$

In the bivariate case, it is $\bar{G}(x,y) = \dfrac{\alpha \bar{F}(x,y)}{1-(1-\alpha)\bar{F}(x,y)}$, (1.8.2)

where $\bar{F}(x,y)$ is a bivariate survival function.

Many researchers studied bivariate minification models, see for example Balakrishna and Jayakumar (1996,1997), Alice and Jose (2004), Ristic (2006) and Kuttikrishnan and Jayakumar (2007).

Balakrishna and Jayakumar (1996) introduced a bivariate autoregressive minification process with the following structure.

$$X_n = \begin{cases} \phi_1^{-1}\left(\lambda_1^{-1}\phi_1(X_{n-1})\right) & w.p. \quad \lambda_1 \\ \min\left(\phi_1^{-1}\left(\lambda_1^{-1}\phi_1(X_{n-1})\right), \varepsilon_n\right) & w.p. \quad (1-\lambda_1) \end{cases}$$

and (1.8.3)

$$Y_n = \begin{cases} \phi_2^{-1}\left(\lambda_2^{-1}\phi_2(X_{n-1})\right) & w.p. \quad \lambda_2 \\ \min\left(\phi_2^{-1}\left(\lambda_2^{-1}\phi_2(X_{n-1})\right), \eta_n\right) & w.p. \quad (1-\lambda_2) \end{cases}$$

where $\{\varepsilon_n, \eta_n\}$ is a bivariate random vector independent of $\{X_i, Y_i\}, i < n$. They have shown that this is useful in constructing stationary bivariate AR(1) processes with required stationary bivariate distribution.

Balakrishna and Jayakumar (1997) introduced a bivariate semi-Pareto minification process with the structure,

$$\begin{aligned} X_n &= \min\left(\lambda^{-1/\beta_1} X_{n-1}, \varepsilon_n\right) \\ Y_n &= \min\left(\lambda^{-1/\beta_2} Y_{n-1}, \eta_n\right) \end{aligned} \qquad (1.8.4)$$

where $0 \le \lambda \le 1$, $\beta_1, \beta_2 \ge 0$ and $(\varepsilon_n, \eta_n) = \begin{cases} (\infty, \infty) & w.p. \quad \lambda \\ (\xi_n, \kappa_n) & w.p. \quad (1-\lambda) \end{cases}$

where ξ_n, κ_n are real valued random variables. Then (1.8.4) defines a stationary bivariate semi-Pareto process if $(X_0, Y_0) \underline{\underline{d}} (\xi_1, \kappa_1)$ and $\{\xi_n, \kappa_n\}$ is bivariate semi-Pareto.

Ristic (2006) introduced a stationary bivariate minification process having the form,

$$\begin{aligned} X_n &= K \min(X_{n-1}, Y_{n-1}, \varepsilon_n) \\ Y_n &= L \min(X_{n-1}, Y_{n-1}, \kappa_n), \end{aligned} \qquad (1.8.5)$$

where $\{\varepsilon_n, \kappa_n\}$ is a sequence of i.i.d. non-negative random vectors with common survival function $\overline{G}(x, y)$ and (X_0, Y_0) and $(\varepsilon_1, \kappa_1)$ are independent, $K > 1$ and $L > 1$. He obtained a necessary and sufficient condition for $\{X_n, Y_n\}$ to be stationary.

1.9 Estimation Problems

To verify the suitability of the model to explain real life situations, one has to have reasonably good estimators for the unknown parameters. Gaver and Lewis (1980) discussed some problems of estimation of their models. Lawrance and Lewis (1981) briefly considered estimation aspects of the models. Ristic and Popovic (2000a) considered the problem of estimation of the parameter k of the uniform autoregressive processes of the first-order with positive and negative lag one-autocorrelation functions. Ristic and Popovic (2000b) estimated the unknown parameters of the new uniform autoregressive process of order 1. Smith (1986) has done maximum likelihood estimation for the second order New Exponential Autoregressive (NEAR(2)) model. Karlsen and Tjøstheim (1988) introduced consistent estimates for the NEAR(2) and New Laplace Autoregressive (NLAR(2)) time series models.

Balakrishna (1998) discussed the problem of estimating the parameters of the semi-Pareto process of Pillai (1991) and Pareto process of Yeh et al. (1988). Balakrishna and Jacob (2003) estimated the parameters of the process (1.6.1).

CHAPTER II

A CLASS OF AUTOREGRESSIVE MINIFICATION PROCESSES

CHAPTER II

A CLASS OF AUTOREGRESSIVE MINIFICATION PROCESSES[♠]

2.1 Introduction

There are distributions where the Laplace transform is not in closed form but the survival function exists in a closed form. For such distributions it is difficult to find the distribution of the innovation sequence $\{\varepsilon_n\}$ such that the model $X_n = \lambda X_{n-1} + \varepsilon_n$ can be defined. As an alternative to this, Tavares (1977) introduced a minification process by replacing addition by minimum. The exponential minification process studied in Tavares (1980) has the form

$$X_0 = \varepsilon_0 \ , \ \ X_n = K \min(X_{n-1}, \varepsilon_n), \text{ for } n = 1,2,\cdots ,$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. exponential random variables chosen so that the stationary distribution of X_i is exponential. Analogous to this several minification processes appeared in the literature. Lewis and McKenzie (1991) have obtained the conditions under which a stationary minification process exists. They also have studied the structure of the minification process under a transformation. Krishnarani and Jayakumar (2007a) introduced an autoregressive minification process using monotone transformation.

Our aim is to construct a class of minification models that can be used to generate autoregressive minification models with a given marginal. For this we consider a

[♠] Some part of this Chapter is based on Krishnarani and Jayakumar (2007a).

monotone transformation of the distribution function, F(x) and we use the same to construct a minification model.

In Section 2, we introduce an autoregressive process and give a necessary and sufficient condition for the process to be stationary. In Section 3, we define a general stationary Markov process with innovation and give some examples. Some properties of this process are also studied. Higher order extension of the process is studied in Section 4. In Section 5, moving average processes are dealt with and in Section 6 ARMA process is constructed. In Section 7, Lehmann family of distributions and processes are introduced with several examples. Last Section deals with estimation problems.

2.2 A General Class of Autoregressive Minification Processes

Let F(.) be a non-degenerate distribution function with $F(-\infty)=0$ and $F(\infty)=1$.

Consider the monotone transformation,

$$\phi(x) = \ln \frac{F(x)}{\bar{F}(x)} , \tag{2.2.1}$$

where $\phi(-\infty) = -\infty, \phi(\infty) = \infty$ and $\bar{F}(x) = 1 - F(x)$.

Define the general Markov process as

$$X_n = \begin{cases} \phi^{-1}[\phi(X_{n-1}) - \ln p] & \text{w.p.} \quad p \\ \min[\phi^{-1}(\phi(X_{n-1}) - \ln p), \varepsilon_n] & \text{w.p.} \quad (1-p) \end{cases} \tag{2.2.2}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables, ε_n is independent of X_i's $(i < n)$, $0 < p < 1$.

Next we seek a necessary and sufficient condition for the $\{X_n\}$ to be stationary.

Theorem 2.2.1

Let X_0 has distribution function F. The process $\{X_n\}$ in (2.2.2) is a strictly stationary Markov process if and only if ε_n's are i.i.d. with distribution function F.

Proof:

Let ε_n's be i.i.d. random variables with distribution function F and let X_0 has distribution function F.

Expressing (2.2.1) in terms of survival functions, we have

$$\bar{F}_{X_n}(x) = p\bar{F}_{X_{n-1}}[\phi^{-1}(\phi(x)+\ln p)]+(1-p)\bar{F}_{X_{n-1}}[\phi^{-1}(\phi(x)+\ln p)]\bar{F}_{\varepsilon_n}(x).$$

When n=1, this becomes,

$$\bar{F}_{X_1}(x) = p\bar{F}_{X_0}[\phi^{-1}(\phi(x)+\ln p)]+(1-p)\bar{F}_{X_0}[\phi^{-1}(\phi(x)+\ln p)]\bar{F}_{\varepsilon_1}(x). \tag{2.2.3}$$

Since $\phi(x) = \ln \dfrac{F(x)}{\bar{\mathrm{F}}(\mathrm{x})}$, $\bar{F}(x) = \dfrac{1}{1+e^{\phi(x)}}$.

Using this in (2.2.3), we get $\bar{F}_{X_1}(x) = \dfrac{1}{1+e^{\phi(x)}} = \bar{F}_{X_0}(x)$ and hence $X_1 \overset{d}{=} X_0$.

Suppose X_{n-1} has distribution function F. Then we get $X_n \overset{d}{=} X_{n-1} \overset{d}{=} X_0$.

Hence the process (2.2.2) is stationary.

Conversely, suppose that $\{X_n\}$ is stationary.

Then $\bar{F}_X(x) = \bar{F}_X[\phi^{-1}(\phi(x) + \ln p)]\,[p + (1-p)\;P(\varepsilon_n > x)]$.

Since X_0 has distribution function F, we have

$$p + (1-p)\;\bar{F}_{\epsilon_n}(x) = \frac{\dfrac{1}{1+e^{\phi(x)}}}{\dfrac{1}{1+pe^{\phi(x)}}} = \frac{1 + pe^{\phi(x)}}{1 + e^{\phi(x)}}.$$

That is, $\bar{F}_{\varepsilon_n}(x) = \dfrac{1}{1 + e^{\phi(x)}}$ and hence the proof. □

Theorem 2.2.2

If ε_n's are i.i.d. with distribution function F and X_0 is arbitrary, then $\{X_n\}$ in (2.2.2) converges in distribution to Z where Z has the distribution function F.

Proof:

Suppose ε_n's are i.i.d. with distribution function F.

Then from (2.2.3), we have

$$\bar{F}_{X_1}(x) = \bar{F}_{X_0}\left[\phi^{-1}(\phi(x) + \ln p)\right]\left[p + (1-p)\,\frac{1}{1 + e^{\phi(x)}}\right]$$

$$= \bar{F}_{X_0}\left[\phi^{-1}(\phi(x) + \ln p\right]\left[\frac{1 + pe^{\phi(x)}}{1 + e^{\phi(x)}}\right]. \qquad (2.2.4)$$

Similarly, when n=2,

$$\bar{F}_{X_2}(x) = \bar{F}_{X_1}[\phi^{-1}(\phi(x) + \ln p)][p + (1-p)\bar{F}_\varepsilon(x)].$$

$$= \bar{F}_{X_1}\left[\phi^{-1}(\phi(x) + \ln p)\right]\left[\frac{1 + pe^{\phi(x)}}{1 + e^{\phi(x)}}\right]. \tag{2.2.5}$$

Substituting (2.2.4) in (2.2.5), we get

$$\bar{F}_{X_2}(x) = \bar{F}_{X_0}\left[\phi^{-1}(\phi(x) + \ln p^2)\right]\left[\frac{1 + p^2 e^{\phi(x)}}{1 + e^{\phi(x)}}\right].$$

Suppose, $\bar{F}_{X_n}(x) = \bar{F}_{X_0}\left[\phi^{-1}(\phi(x) + \ln p^n)\right]\left[\frac{1 + p^n e^{\phi(x)}}{1 + e^{\phi(x)}}\right]$

Then, it can be proved that,

$$\bar{F}_{X_{n+1}}(x) = \bar{F}_{X_0}\left[\phi^{-1}(\phi(x) + \ln p^{n+1})\right]\left[\frac{1 + p^{n+1} e^{\phi(x)}}{1 + e^{\phi(x)}}\right]$$

Thus, $\bar{F}_{X_n}(x) = \bar{F}_{X_0}\left[\phi^{-1}(\phi(x) + \ln p^n)\right]\left[\frac{1 + p^n e^{\phi(x)}}{1 + e^{\phi(x)}}\right]$

As n $\to \infty$, $\quad \bar{F}_{X_n}(x) \to \frac{1}{1 + e^{\phi(x)}}$.

Hence $X_n \xrightarrow{d} Z$ where Z has distribution function F. □

Theorem 2.2.3

Suppose ε_n's have distribution function H. A necessary and sufficient condition for the first order autoregressive process $\{X_n\}$ in (2.2.2) to be stationary is

$$(1-p)\, H(x) = P[X_0 \le x / X_0 > \phi^{-1}(\phi(x) + \ln p)] \tag{2.2.6}$$

Proof:

Suppose ε_n's have distribution function H and $\{X_n\}$ in (2.2.2) is stationary. Let G_i be the distribution function of X_i, i = 0, 1, 2, ... We have $X_n \underline{\underline{d}} X_0$ for all n.

Expressing (2.2.2) in terms of survival function and substituting n=1,

$$\overline{G}_1(x) = \overline{G}_0 \; [\phi^{-1}(\phi(x) + \ln p)]\,[p + (1-p)\;\overline{H}(x)].$$

If $X_1 \underline{\underline{d}} X_0$, then the above implies that

$$(1-p)\overline{H}(x) = \frac{\overline{G}_0(x) - p.\overline{G}_0 \; [\phi^{-1}(\phi(x) + \ln p)]}{\overline{G}_0 \; [\phi^{-1}(\phi(x) + \ln p)]}.$$

That is, $(1-p)H(x) = \dfrac{\overline{G}_0 \, [\phi^{-1}(\phi(x) + \ln p)] - \overline{G}_0(x)}{\overline{G}_0 \; [\phi^{-1}(\phi(x) + \ln p)]}$.

Hence $(1-p)\; H(x) = P[X_0 \le x / \; X_0 > \phi^{-1}\,[\phi(x) + \ln p)]$.

Conversely, suppose $(1-p)\; H(x) = P[X_0 \le x / \; X_0 > \phi^{-1}\,[\phi(x) + \ln p)]$.

This means that $(1-p)\; H(x) = \dfrac{\overline{G}_0(y) - \overline{G}_0(x)}{\overline{G}_0(y)}$ where $y = \phi^{-1}[\phi(x) + \ln p]$

and hence $(1-p)\; \overline{H}(x) = \dfrac{\overline{G}_0(x) - p\,\overline{G}_0(y)}{\overline{G}_0(y)}$.

From (2.2.2), $\overline{G}_n(x) = \overline{G}_{n-1}[\phi^{-1}(\phi(x) + \ln p)]\,[p + (1-p)\;\overline{H}(x)]$.

For $n = 1$, $\overline{G}_1(x) = \overline{G}_0(y)\;[p + (1-p)\overline{H}(x)]$.

If we substitute $(1-p)\,\overline{H}(x) = \dfrac{\overline{G}_0(x) - p\,\overline{G}_0(y)}{\overline{G}_0(y)}$, we get $\overline{G}_1(x) = \overline{G}_0(x)$

That is, $X_1 \overset{d}{=} X_0$.

Suppose $X_n \overset{d}{=} X_0$, then we get $\overline{G}_{n+1}(x) = \overline{G}_0(x)$.

Thus $X_n \overset{d}{=} X_0 \Rightarrow X_{n+1} \overset{d}{=} X_0$.

Hence $X_n \overset{d}{=} X_0$ for all n.

Therefore, the process is stationary. □

Based on the above theorems we define the new autoregressive process with innovation in the next Section.

2.3 Stationary Markov Process with Innovation

The stationary Markov process with innovation is defined below.

Definition 2. 3.1

Let $\phi(x) = \ln \dfrac{F(x)}{\overline{F}(x)}$. We say that $\{X_n\}$ is a general stationary Markov process with innovation if

$$X_n = \begin{cases} \phi^{-1}[\phi(X_{n-1}) - \ln p] & \text{w.p.} \quad p \\ \min\,[\phi^{-1}\,(\phi(X_{n-1}) - \ln p),\, \varepsilon_n] & \text{w.p.} \quad (1-p) \end{cases} \tag{2.3.1}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables with common distribution function F, ε_n is independent of X_i, $i = 0, 1, 2, \ldots, n-1$ with X_0 having distribution function F, $0 < p < 1$. □

Now we study some properties of the process given in (2.3.1). The joint survival function of (X_n, X_{n+1}) is,

$$P(X_n > x_n, X_{n+1} > x_{n+1})$$

$$= \begin{cases} \dfrac{1}{1+e^{\phi(x_{n+1})}} & 0 \le x_n \le \phi^{-1}[\phi(x_{n+1}) + \ln p] \\ \dfrac{1+p\,e^{\phi(x_{n+1})}}{(1+e^{\phi(x_n)})(1+e^{\phi(x_{n+1})})} & 0 < \phi^{-1}[\phi(x_{n+1}) + \ln p] < x_n < \infty \end{cases} \qquad (2.3.2)$$

It can be easily seen that the process $\{X_n\}$ is not time reversible.

Also it can be seen that $P\ (X_{n+1} > X_n) = \dfrac{p+1}{2}$. (2.3.3)

Thus we have, as p increases more up runs in the process $\{X_n\}$ can be observed.

The distribution of extremes of the process $\{X_n\}$ in (2.3.1) is given below.

Let N be a geometric random variable with probability mass function $p(N = n) = p(1-p)^{n-1}$ $n = 1, 2, \ldots$. Assuming that N is independent of X_i's , $i = 1, 2, \ldots$, we define geometric minimum (maximum) as,

$$T_N = \min(X_1, \ldots, X_N) \left[M_N = \max(X_1, \ldots, X_N)\right].$$

The distribution of the geometric minimum and maximum are evaluated below.

$$\bar{F}_{T_N}(x) = P\left[\min(X_1, ..., X_N) > x\right]$$

$$= \frac{1}{1 + \frac{1}{p} e^{\phi(x)}}. \tag{2.3.4}$$

$$F_{M_N}(x) = P[\max(X_1, ..., X_N) < x]$$

$$= \frac{e^{\phi(x)}}{\frac{1}{p} + e^{\phi(x)}}. \tag{2.3.5}$$

Some examples of $\{X_n\}$ in (2.3.1) are given below.

Example 2.3.1

Let $F(x) = 1 - \dfrac{1}{1 + e^{\beta x}}, \quad -\infty < x < \infty$. Then

$$X_n = \begin{cases} X_{n-1} - \frac{1}{\beta} \ln p & \text{w.p.} \quad \text{p} \\ \min(X_{n-1} - \frac{1}{\beta} \ln p, \varepsilon_n) & \text{w.p.} \quad (1-\text{p}) \end{cases} \tag{2.3.6}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. logistic random variables with distribution function F. Then $\{X_n\}$ defines a first order autoregressive logistic process.

Note that if we assume ε_n's as i.i.d. semi-logistic random variables and $X_0 \stackrel{d}{=} \varepsilon_1$ then (2.3.6) defines a stationary AR(1) process with semi-logistic distribution as marginal.

(2.3.6) is the logistic process studied in Arnold and Robertson (1989) and also the semi-logistic process in Jayakumar and Mathew (2004).

Example 2.3.2

Let $\bar{F}(x) = \dfrac{1}{1 + x^{\beta}}$, $x > 0$, $\beta > 0$. Then (2.3.1) becomes,

$$X_n = \begin{cases} p^{-\frac{1}{\beta}} X_{n-1} & \text{w.p.} \quad \text{p} \\ \min\left(p^{-\frac{1}{\beta}} X_{n-1}, \varepsilon_n \right) & \text{w.p.} \quad (1-\text{p}) \end{cases} \tag{2.3.7}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. Pareto III random variables with distribution function F(x). Note that (2.3.7) is the Pareto process studied in Yeh et al. (1988).

Example 2.3.3

If $F(x) = 1 - e^{-x}$, $x>0$, $\alpha > 0$, (2.3.8)

then the process (2.3.1) takes the form,

$$X_n = \begin{cases} \ln\left(1 + \frac{1}{p}\left(e^{X_{n-1}} - 1\right)\right) & w.p. \quad p \\ \min\left(\ln(1+\frac{1}{p}\,(e^{X_{n-1}} - 1), \varepsilon_n \right) & w.p. \quad (1-p) \end{cases} \tag{2.3.9}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. exponential random variables with distribution function (2.3.8). That is the model (2.3.1) defines a stationary exponential process when $X_0 \overset{d}{=} \varepsilon_1$ where ε_1 has the exponential distribution function.

Example 2.3.4

When X has the Laplace distribution with probability density function,

$f(x) = \frac{1}{2}e^{-|x|}, \quad -\infty < x < \infty,$ the process (2.3.1) becomes

$$X_n = \begin{cases} \ln\left(\frac{p-1}{2p} + \frac{e^{|X_{n-1}|}}{p}\right) & \text{w.p. } p \\ \min\left(\ln\left(\frac{p-1}{2p} + \frac{e^{|X_{n-1}|}}{p}, \varepsilon_n\right)\right) & \text{w.p. } (1-p) \end{cases} \tag{2.3.10}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. Laplace random variables. (2.3.10) is a stationary sequence of Laplace random variables if $X_0 \underset{=}{d} \varepsilon_1$.

Example 2.3.5

When X_n's have uniform distribution, $f(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \textit{otherwise}, \end{cases}$

then (2.3.1) becomes

$$X_n = \begin{cases} \frac{X_{n-1}}{p+(1-p)X_{n-1}} & \text{w.p. } p \\ \min\left(\frac{X_{n-1}}{p+(1-p)X_{n-1}}, \varepsilon_n\right) & \text{w.p. } (1-p) \end{cases} \tag{2.3.11}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. uniform (0, 1) random variables. This process generates a stationary sequence of uniform random variables if $X_0 \underset{=}{d} \varepsilon_1$.

The sample path behaviour of the uniform autoregressive process (2.3.11) is presented in **Figures 2.3.1** below.

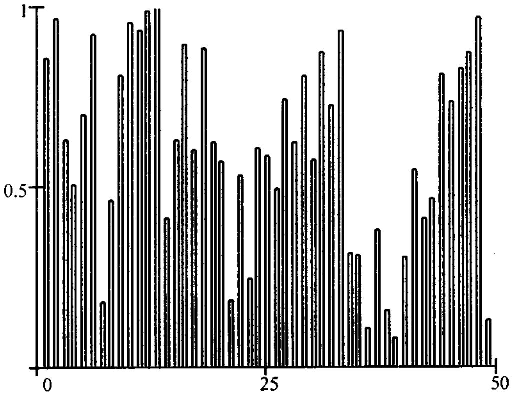


Figure 2.3.1a (p=0.2)

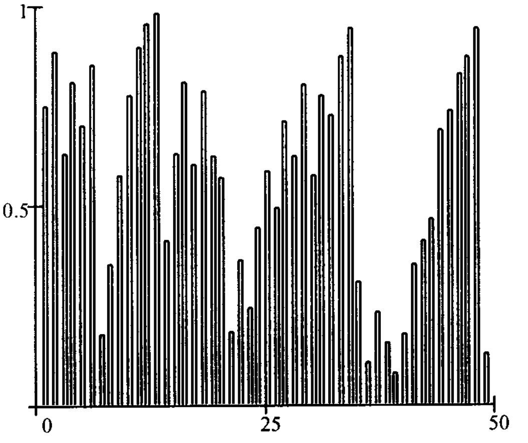


Figure 2.3.1b (p=0.4)

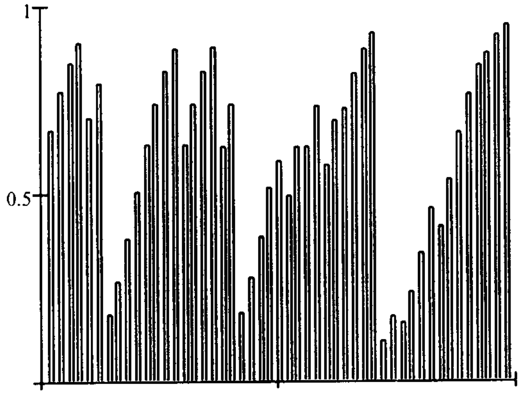


Figure 2.3.1c (p=0.6)

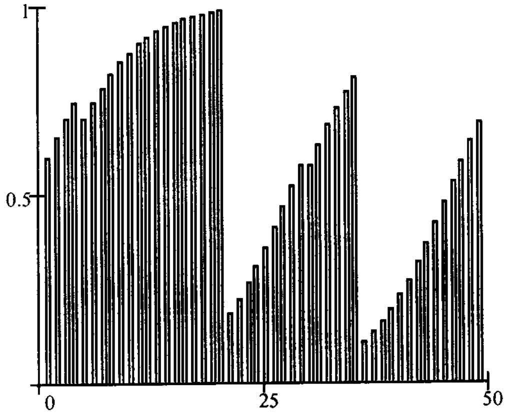


Figure 2.3.1d (p=0.8)

Figure 2.3.1 Sample path behaviour of uniform autoregressive process.

2.4 Higher Order Extensions

2.4.1 Autoregressive Process of Order 2 (AR(2))

Note that one may be able to define different extensions of the AR(1) process $\{X_n\}$ in (2.3.1). Now we develop an AR(2) process as a generalization of the model.

Let $\phi(x)$ be as in (2.2.1). Define

$$X_n = \begin{cases} \phi^{-1}(\phi(X_{n-1}) - \ln p_1) & \text{w.p.} \quad p_1 \\ \min\left((\phi^{-1}(\phi(X_{n-1}) - \ln p_1), \varepsilon_n\right) & \text{w.p.} \quad p_2 \\ \min\left(\phi^{-1}(\phi(X_{n-2}) - \ln p_1), \varepsilon_n\right) & \text{w.p.} \quad p_3 \end{cases} \tag{2.4.1}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables, $0 < p_1, p_2, p_3 < 1$, $p_1 + p_2 + p_3 = 1$.

Theorem 2.4.1

Let X_0 has distribution function F. The process $\{X_n\}$ in (2.4.1) is stationary if and only if $\varepsilon_n \underset{=}{d} X_0$.

Proof:

(2.4.1) in terms of survival function is

$$\overline{F}_{X_n}(x) = \overline{F}_{X_{n-1}}[\phi^{-1}(\phi(x) + \ln p_1)]\,[p_1 + p_2\,\overline{F}_{\varepsilon_n}(x) + p_3\,\overline{F}_{\varepsilon_n}(x)]. \tag{2.4.2}$$

If X_n's are stationary and X_0 has distribution function F where $\overline{F}(x) = \dfrac{1}{1 + e^{\phi(x)}}$,

then

$$p_1 + (1-p_1)\,\bar{F}_{\varepsilon_n}(x) = \frac{\dfrac{1}{1+e^{\phi(x)}}}{\dfrac{1}{1+p_1 e^{\phi(x)}}} = \frac{1+p_1 e^{\phi(x)}}{1+e^{\phi(x)}}$$

and hence $\bar{F}_{\varepsilon_n}(x) = \dfrac{1}{1+e^{\phi(x)}}$.

That is, $\varepsilon_n \underline{\underline{d}}\, X_0$.

Conversely, let $X_0 \underline{\underline{d}}\, \varepsilon_n$ and ε_n's are i.i.d. with distribution function F.

When n = 1, (2.4.2) becomes

$$\bar{F}_{X_1}(x) = \bar{F}_{X_0}[\phi^{-1}(\phi(x) + \ln p_1)]\,[p_1 + p_2\,\bar{F}_{\varepsilon}(x) + p_3\,\bar{F}_{\varepsilon}(x)]$$

$$= \frac{1}{1+p_1 e^{\phi(x)}}\left[p_1 + (1-p_1)\frac{1}{1+e^{\phi(x)}}\right]$$

$$= \frac{1}{1+e^{\phi(x)}}.$$

That is, $X_1 \underline{\underline{d}}\, X_0$.

If we assume that $X_{n-1} \underline{\underline{d}}\, X_0$, then we can prove that $X_n \underline{\underline{d}}\, X_0$.

Therefore the process $\{X_n\}$ is stationary.

Hence the Theorem. □

Now as an illustration of this AR(2) process, we define a second order logistic process.

In this case, the process (2.4.1) takes the form

$$X_n = \begin{cases} X_{n-1} - \frac{1}{\beta} \ln p_1 & \text{w.p.} \quad p_1 \\ \min\left(X_{n-1} - \frac{1}{\beta} \ln p_1, \ \varepsilon_n\right) & \text{w.p.} \quad p_2 \\ \min\left(X_{n-2} - \frac{1}{\beta} \ln p_1, \ \varepsilon_n\right) & \text{w.p.} \quad p_3 \end{cases} \tag{2.4.3}$$

where $0 < p_i < 1 (i = 1, 2, 3)$, $p_1 + p_2 + p_3 = 1$ and $\{\varepsilon_n\}$ is a sequence of i.i.d. logistic random variables with survival function

$$\overline{F}_\varepsilon(x) = \frac{1}{1 + e^{\beta x}}, \quad \beta > 0, -\infty < x < \infty \text{ and } X_0 \stackrel{d}{=} \varepsilon_1.$$

We propose another second order model as,

$$X_n = \begin{cases} \min\left(\phi^{-1}(\phi(X_{n-1}) - \ln p_1), \varepsilon_n\right) & w.p. \quad (1 - p_2) \\ \min\left(\phi^{-1}(\phi(X_{n-2}) - \ln p_2), \ \varepsilon_n\right) & w.p. \quad p_2 \end{cases} \tag{2.4.4}$$

where p_1 and p_2 are constants, $0 < p_1, p_2 < 1$.

Here X_n is always a function of either X_{n-1} or X_{n-2}.

Theorem 2.4.2

If the process $\{X_n\}$ in (2.4.4) is a strictly stationary Markov process with survival function ,

$$\overline{F}(x) = \frac{1}{1 + e^{\phi(x)}} \tag{2.4.5}$$

then ε_n 's can be written as a mixture of three random variables in the form,

$$\varepsilon_n = \begin{cases} 0 & w.p. \quad \dfrac{p_1}{1-p_2+p_1} \\ X_n & w.p. \quad \dfrac{(1-p_2)(1-p_1)}{1-p_2(1-p_2+p_1)} \\ \phi^{-1}(\phi(X_n)-\ln p_2(1-p_2+p_1)) & w.p. \quad \dfrac{(1-p_2)(p_1-p_2)^2}{(1-p_2+p_1)(1-p_2(1-p_2+p_1))} \end{cases} \tag{2.4.6}$$

Proof:

Suppose X_n 's are stationary with survival function (2.4.5).

From (2.4.4), we get

$$\overline{F}_X(x) = (1-p_2)\overline{F}_X\left(\phi^{-1}(\phi(x)+\ln p_1)\right)\overline{F}_{\varepsilon_n}(x) + p_2\overline{F}_X\left(\phi^{-1}(\phi(x)+\ln p_2)\right)\overline{F}_{\varepsilon_n}(x) .$$

Using (2.4.5) we have

$$\overline{F}_{\varepsilon_n}(x) = \frac{1+p_2e^{\phi(x)}+p_1e^{\phi(x)}+p_1p_2e^{2\phi(x)}}{\left(1+p_2e^{\phi(x)}(1-p_2+p_1)\right)\left(1+e^{\phi(x)}\right)}$$

$$= \frac{p_1}{1-p_2+p_1} +$$

$$\frac{(1-p_2)(1-p_1)}{(1-p_2(1-p_2+p_1))\left(1+e^{\phi(x)}\right)} +$$

$$\frac{(1-p_2)(p_1-p_2)^2}{(1-p_2+p_1)(1-p_2(1-p_2+p_1))\left(1+e^{\phi(x)}p_2(1-p_2+p_1)\right)}$$

This means that ε_n can be written as a mixture of three random variables such that,

$$\varepsilon_n = \begin{cases} 0 & w.p. \quad \dfrac{p_1}{1-p_2+p_1} \\ X_n & w.p. \quad \dfrac{(1-p_2)(1-p_1)}{1-p_2(1-p_2+p_1)} \\ \phi^{-1}(\phi(X_n)-\ln p_2(1-p_2+p_1)) & w.p \quad \dfrac{(1-p_2)(p_1-p_2)^2}{(1-p_2+p_1)(1-p_2(1-p_2+p_1))} \end{cases}$$

where first is degenerate at 0 with probability $\dfrac{p_1}{1-p_2+p_1}$, the second has survival function $\overline{F}$ with probability $\dfrac{(1-p_2)(1-p_1)}{1-p_2(1-p_2+p_1)}$ and the third has survival function $\dfrac{1}{1+e^{\phi(x)}p_2(1-p_2+p_1)}$ with probability $\dfrac{(1-p_2)(p_1-p_2)^2}{(1+p_1-p_2)(1-p_2(1-p_2+p_1))}$.

Hence the theorem. □

2.4.2 Autoregressive Process of Order q

The q[th] order general autoregressive minification process may be constructed similarly and hence it may be written as,

$$X_n = \begin{cases} \min\left(\phi^{-1}(\phi(X_{n-2})-\ln p_1), \varepsilon_n\right) & w.p. \quad a_1 \\ \min\left(\phi^{-1}(\phi(X_{n-2})-\ln p_2), \varepsilon_n\right) & w.p. \quad a_2 \\ \vdots & \vdots \quad \vdots \\ \min\left(\phi^{-1}(\phi(X_{n-2})-\ln p_q), \varepsilon_n\right) & w.p. \quad a_q \end{cases} \tag{2.4.7}$$

where $a_1 = 1-p_2, a_q = p_2 p_3 \cdots p_q \quad and \quad a_k = \prod_{j=2}^{k} p_j(1-p_{k+1})$.

But it is difficult to derive the distribution of ε_n's in this case.

2.5 Moving Average Processes

Another objective is to obtain a MA model corresponding to (2.2.2). We define the MA model as,

$$X_n = \begin{cases} \phi^{-1}(\phi(\varepsilon_{n-1}) - \ln p) & \text{w.p.} \quad p \\ \min\left(\phi^{-1}(\phi(\varepsilon_{n-1}) - \ln p),\ \epsilon_n\right) & \text{w.p.} \quad (1-p) \end{cases} \tag{2.5.1}$$

where $\phi(x) = \ln \dfrac{F(x)}{\overline{F}(x)}$ and $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables with survival function, $\overline{F}(x) = \dfrac{1}{1 + e^{\phi(x)}}$.

MA(2) model is,

$$X_n = \begin{cases} \phi^{-1}(\phi(\varepsilon_n) - \ln p_2) & w.p. \quad p_2 \\ \min\left(\phi^{-1}(\phi(\varepsilon_n) - \ln p_2), \phi^{-1}(\phi(\varepsilon_{n-1}) - \ln p_1)\right) & w.p. \quad (1-p_2)p_1 \\ \min\left(\phi^{-1}(\phi(\varepsilon_n) - \ln p_2), \phi^{-1}(\phi(\varepsilon_{n-1}) - \ln p_1), \varepsilon_{n-2}\right) & w.p. \quad (1-p_2)(1-p_1) \end{cases} \tag{2.5.2}$$

where $0 \le p_1, p_2 \le 1$, $p_1 + p_2 = 1$.

Similarly the MA(q) model can be written as,

$$X_n = \begin{cases} \phi^{-1}\left(\phi(\varepsilon_n) - \ln p_q\right) & w.p. \quad b_{q+1} \\ \min\left(\phi^{-1}\left(\phi(\varepsilon_n) - \ln p_q\right), \phi^{-1}\left(\phi(\varepsilon_{n-1}) - \ln p_{q-1}\right)\right) & w.p. \quad b_q \\ \vdots & \vdots \quad \vdots \\ \min\left(\phi^{-1}\left(\phi(\varepsilon_n) - \ln p_q\right), \ldots, \phi^{-1}\left(\phi(\varepsilon_{n-(q-1)}) - \ln p_1\right), \varepsilon_{n-q}\right) & w.p. \quad b_1 \end{cases} \tag{2.5.3}$$

where,

$$b_i = \begin{cases} p_q & i = q+1 \\ (1-p_q)(1-p_{q-1})\dots(1-p_i)p_i & 2 \le i \le q \\ (1-p_q)\dots(1-p_1) & i = 1 \end{cases} \tag{2.5.4}$$

It is easy to check that $\sum_{i=1}^{q+1} b_i = 1$.

2.6 Autoregressive Moving Average Process

In this Section we combine the MA(q) model and AR(p) model discussed earlier. Thus we have ARMA (p, q) process as defined below.

$$X_n = \begin{cases} \phi^{-1}\left(\phi(\varepsilon_n) - \ln p_q\right) & w.p. \quad b_{q+1} \\ \min\left(\phi^{-1}\left(\phi(\varepsilon_n) - \ln p_q\right), \phi^{-1}\left(\phi(\varepsilon_{n-1}) - \ln p_{q-1}\right)\right) & w.p. \quad b_q \\ \vdots & \vdots \quad \vdots \\ \min\left(\phi^{-1}\left(\phi(\varepsilon_n) - \ln p_q\right), \dots, \phi^{-1}\left(\phi(\varepsilon_{n-(q-1)}) - \ln p_1\right), A^{(p)}{}_{n-q}\right) & w.p. \quad b_1 \end{cases}$$

where $A^{(p)}{}_{n-q}$ is an AR(p) process and it is independent of $\varepsilon_n, \varepsilon_{n-1}, \dots, \varepsilon_{n-(q-1)}$.

2.7 Lehmann Family of Distributions and Autoregressive Processes

Lehmann family of distributions is generated from a given survival function (see Lehmann (1953)). Let $\overline{F}(x)$ be an arbitrary known survival function. If γ is positive then,

$$\overline{S}(x) = \left(\overline{F}(x)\right)^{\gamma}, \tag{2.7.1}$$

is also a survival function. If in particular γ is a positive integer then, it represents the survival function of $\min(X_1, X_2, \dots, X_n)$ where X_i's are i.i.d. random variables with F(x) as the common distribution function.

The density function corresponding to $\bar{S}(x)$ is

$$f_\gamma(x) = \gamma\left(\bar{F}(x)\right)^{\gamma-1} f(x) \tag{2.7.2}$$

and the failure rate is

$$h_\gamma(x) = \frac{f_\gamma(x)}{\bar{S}(x)} = \frac{\gamma\left(\bar{F}(x)\right)^{\gamma-1} f(x}{\left(\bar{F}(x)\right)^\gamma} = \gamma\, h(x)\,. \tag{2.7.3}$$

where $h(x)$ is the failure rate of $F(x)$. Thus hazards are proportional. Hence Lehmann family is also known as proportional hazards family.

We define a model having the structure,

$$X_n = \min\left(\phi^{-1}(\phi(X_{n-1}) - \ln p), \eta_n\right) \tag{2.7.4}$$

where $\eta_n = \min\limits_{1 \le i \le \gamma} \left(\eta_{in}\right)$ and $\eta_{in} = \begin{cases} \infty & w.p. \quad p \\ \varepsilon_n & w.p. \quad (1-p) \end{cases}$ (2.7.5)

where $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables with survival function $\dfrac{1}{1 + e^{\phi(x)}}$ and ε_n is independent of X_i's $(i < n)$, $0 < p < 1$.

Note that when $\phi\,(x) = \ln \dfrac{F(x)}{\bar{F}(x)}$,

$$\bar{F}(x) = \frac{1}{1 + e^{\phi(x)}}\,. \tag{2.7.6}$$

Next we seek a necessary and sufficient condition for the $\{X_n\}$ to be stationary.

Theorem 2.7.1

Let X_0 has survival function $\bar{S}(x) = \bar{F}^{\gamma}(x)$. The process $\{X_n\}$ in (2.7.4) is a strictly stationary Markov process if and only if ε_n's in (2.7.5) are i.i.d. with survival function $\bar{F}(x)$.

Proof:

Suppose X_0 has survival function $\bar{S}(x)$, where $\bar{S}(x) = \left(\bar{F}(x)\right)^{\gamma}$, γ is a positive integer.

Let ε_n's are i.i.d. with survival function $\bar{F}(x)$.

Denote the survival functions of X_n as $\bar{G}_{X_n}(x)$ for n=1, 2, 3,....

Then from (2.7.4) and (2.7.5),

$$\bar{G}_{X_n}(x) = \bar{G}_{X_{n-1}}\left(\phi^{-1}(\phi(x) + \ln p)\right)\bar{F}_{\eta_n(x)}$$

$$= \bar{G}_{X_{n-1}}\left(\phi^{-1}(\phi(x) + \ln p)\right)\left(P(\eta_{in} > x)\right)^{\gamma}$$

$$= \bar{G}_{X_{n-1}}\left(\phi^{-1}(\phi(x) + \ln p)\right)\left(p + (1-p)\bar{F}_{\varepsilon}(x)\right)^{\gamma}$$

When n=1, the above equation becomes

$$\bar{G}_{X_1}(x) = \bar{G}_{X_0}\left(\phi^{-1}(\phi(x) + \ln p)\right)\left(p + (1-p)\bar{F}_{\varepsilon}(x)\right)^{\gamma}$$

$$= \left(\frac{1}{1 + pe^{\phi(x)}}\right)^{\gamma}\left(p + (1-p)\frac{1}{1 + e^{\phi(x)}}\right)^{\gamma}$$

$$= \left(\frac{1}{1 + e^{\phi(x)}}\right)^{\gamma}.$$

That is, X_1 has survival function $\bar{S}(x)$.

Assume X_{n-1} has survival function $\bar{S}(x)$. Proceeding as above, we can prove that X_n also has survival function $\bar{S}(x)$.

Thus $\{X_n\}$ is stationary with survival function $\bar{S}(x)$.

Conversely assume $\{X_n\}$ is stationary and X_0 has survival function $\bar{S}(x)$.

Then from (2.7.4) and (2.7.5),

$$\bar{S}(x) = \bar{S}\left(\phi^{-1}(\phi(x) + \ln p)\right)\left(p + (1-p)\bar{F}_{\varepsilon_n}(x)\right)^{\gamma}.$$

That is, $\left(\dfrac{1}{1+e^{\phi(x)}}\right)^{\gamma} = \left(\dfrac{1}{1+pe^{\phi(x)}}\right)^{\gamma}\left(p + (1-p)\bar{F}_{\varepsilon_n}(x)\right).$

Hence $\bar{F}_{\varepsilon_n}(x) = \dfrac{1}{1+e^{\phi(x)}}$.

This shows that ε_n's are i.i.d. with survival function $\bar{F}(x)$.

Hence the theorem. □

For the stationary AR(1) Lehmann process,

$$P(X_n > x, X_{n+1} > y) = \begin{cases} \left(\dfrac{1}{1+e^{\phi(y)}}\right)^{\gamma} & if \quad x > \phi^{-1}(\phi(y) + \ln p) \\ \left(\dfrac{1+pe^{\phi(y)}}{\left(1+e^{\phi(x)}\right)\left(1+e^{\phi(y)}\right)}\right)^{\gamma} & if \quad x < \phi^{-1}(\phi(y) + \ln p) \end{cases} \tag{2.7.7}$$

and

$$P(X_{n+1} > X_n) = \left(\frac{\gamma + p}{\gamma + 1}\right)^{\gamma}. \tag{2.7.8}$$

2.8 Estimation of the Parameters

Here we estimate the parameters p and ϕ of the model (2.2.2) where $\phi(x)$ is as defined in (2.2.1), $\phi(-\infty) = -\infty, \phi(\infty) = \infty$ and $\overline{F}(x) = 1 - F(x)$. Suppose we know a realization $(X_0, X_1, \ldots, X_N)$ from the above-defined process.

Consider the process $\{U_n\}$ defined by,

$$U_n = \begin{cases} 1 & if \quad X_{n+1} > X_n \\ 0 & if \quad X_{n+1} < X_n \end{cases}, n = 0, 1, 2, \cdots, N-1. \tag{2.8.1}$$

From (2.3.3) $P\ (X_{n+1} > X_n) = \dfrac{p+1}{2}$.

$$\text{Therefore } E(U_n) = P\ (X_{n+1} > X_n) = \frac{p+1}{2}. \tag{2.8.2}$$

$$E\left(U_n^{\,2}\right) = P\ (X_{n+1} > X_n) = \frac{p+1}{2} \tag{2.8.3}$$

$$V(U_n) = \frac{1-p^2}{4}. \tag{2.8.4}$$

Hence an estimator of p is obtained by solving,

$$\overline{U}_N = \frac{1}{N}\sum_{i=0}^{N-1} U_i = \frac{\hat{p}+1}{2}, \text{ which yields}$$

$$\hat{p} = 2\overline{U}_N - 1. \tag{2.8.5}$$

Now let us check some of the properties of $\hat{p}$.

$E(\hat{p}) = E\left(2\overline{U}_N - 1\right) = p$ and

$$V(\hat{p}) = 4\,V\left(\overline{U}_N\right) = \frac{1-p^2}{n} \to 0 \quad as \quad n \to \infty.$$

Therefore $\hat{p}$ is an unbiased and consistent estimator.

To estimate $\phi(t)$, we define the level crossing process $\{Z_n(t)\}$ associated with $\{X_n\}$ by,

$$Z_n(t) = \begin{cases} 1 & if \quad X_n > t \\ 0 & if \quad X_n \le t \end{cases} ,n = 0,1,2,\cdots,N \tag{2.8.6}$$

Therefore $E(Z_n(t)) = P(X_n > t)$ (2.8.7)

$$= \frac{1}{1 + e^{\phi(t)}} \text{ (using (2.2.1)).} \tag{2.8.8}$$

Now $\phi(t)$ can be estimated by solving the equation,

$$\overline{Z}_{N+1}(t) = \frac{1}{N+1}\sum_{i=0}^{N} Z_i(t) = \frac{1}{1+e^{\hat{\phi}(t)}}.$$

Hence, the desired estimator is $\hat{\phi}(t) = \ln\left(\frac{1-\overline{Z}_{N+1}(t)}{\overline{Z}_{N+1}(t)}\right)$.

CHAPTER III

HALF SEMI-LOGISTIC PROCESSES

CHAPTER III

HALF SEMI-LOGISTIC PROCESSES♠

3.1 Introduction

The logistic growth function is used to describe both population and organism growth. It was first proposed as a tool in demographic studies. Some researchers used the logistic function for estimating the growth of human population and also applied to the agricultural production data. Logistic distributions are mainly applied in the study of medical diagnosis and public health. The logistic regression model was applied to a large data set of patients where there were many diagnostic categories. For a detailed survey on the applications of logistic distributions one may refer to Balakrishnan (1992).

The Marshall-Olkin parameterization scheme is already mentioned in Section 1.8. The univariate Marshall-Olkin survival function is of the form (1.8.1). Using Marshall-Olkin scheme, Krishnarani and Jayakumar (2007b) introduced a generalized half logistic distribution and studied its properties. In this Chapter we propose generalized half semi-logistic distribution and study the properties of these classes of distributions. The generalized half semi-logistic family of distributions is useful for modeling datasets having periodic fluctuations. The exponential, Pareto, Weibull etc. belong to the family of generalized half semi-logistic distributions. The

♠ This Chapter is based on Krishnarani and Jayakumar (2007b).

family of generalized half semi-logistic distributions is found to be useful in reliability studies, where the lifetimes of the components have periodic failure rate.

In this Chapter, we investigate the class of distributions generated by semi-logistic distributions through the Marshall-Olkin form given by (1.8.1). In Section 2, we introduce a generalized half semi-logistic distribution obtained through the Marshall-Olkin scheme and study its properties. This class includes a number of well-known distributions such as Pareto, Weibull, exponential, half logistic etc. In Section 3, the generalized half semi-logistic autoregressive processes are developed and their properties are studied. Some special classes of distributions that are generalizations of some well-known life distributions are studied in Section 4.

3.2 A Generalized Half Semi-Logistic Distribution

When $\bar{F}$ assumes the form

$$\bar{F}(x) = \frac{2}{1+\psi(x)}, \qquad (3.2.1)$$

where $\psi(x)$ satisfies (1.2.5) and (1.2.6) with $\psi(0)=1, \psi(\infty)=\infty$, the associated Marshall-Olkin form is given by

$$\bar{G}(x) = \frac{2\alpha}{\psi(x)+2\alpha-1}. \qquad (3.2.2)$$

Note that when $\alpha = 1/2$, $\bar{G}(x) = (\psi(x))^{-1}$ and that provides a generalization of exponential distribution. Necessarily we have $\psi(0) = 1$ and $\psi(\infty) = \infty$. When $\alpha = 1$, $\bar{G}(x)$ corresponds to half semi-logistic distribution. We call the distribution given by (3.2.2) as generalized half semi-logistic (GHSL) distribution and is denoted by GHSL(α, β, p).

We present below some basic characteristics of the generalized class of distributions given by (3.2.2).

(i) The probability density function is

$$g(x) = \frac{2\alpha\psi'(x)}{(\psi(x) + 2\alpha - 1)^2}, \, x > 0, \alpha > 0.$$

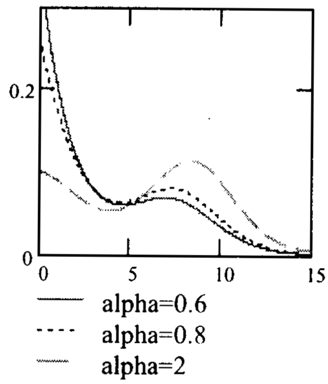
(ii) Median, $M = \psi^{-1}(2\alpha + 1)$.

(iii) Hazard rate, $r(x) = \dfrac{\psi'(x)}{\psi(x) + 2\alpha - 1}$.

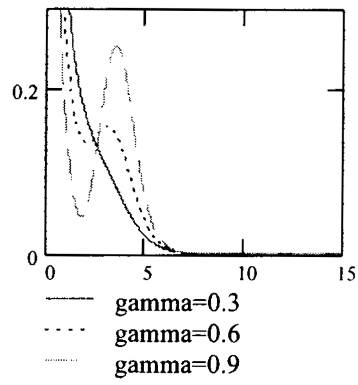
(iv) Mean Residual Life function,

$$\mu(x) = (\psi(x) + 2\alpha - 1) \int_x^\infty \frac{1}{\psi(t) + 2\alpha - 1} dt.$$

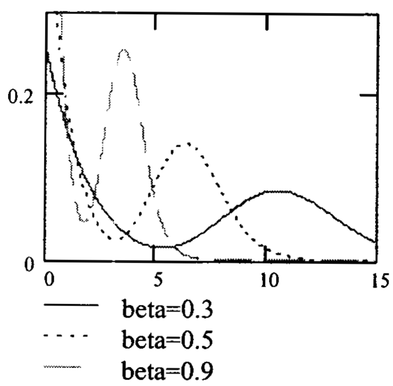
The density plot, the hazard rate and the mean residual life function of GHSL distribution are presented in **Figures 3.2.1, 3.2.2** and **3.2.3** respectively. As an illustration, we take h (x) = $e^{\gamma(\cos \beta x - 1)}$ $(0 < \beta, \gamma < 1)$. Note that h(x) is the solution of the functional equation $\psi(x) = \dfrac{1}{p}\psi\left(\dfrac{1}{\beta}\ln p + x\right)$ and it is periodic with period $\dfrac{1}{\beta}\ln p$, where $p = e^{-2\pi}$.



$\beta=0.4, \gamma=0.6$



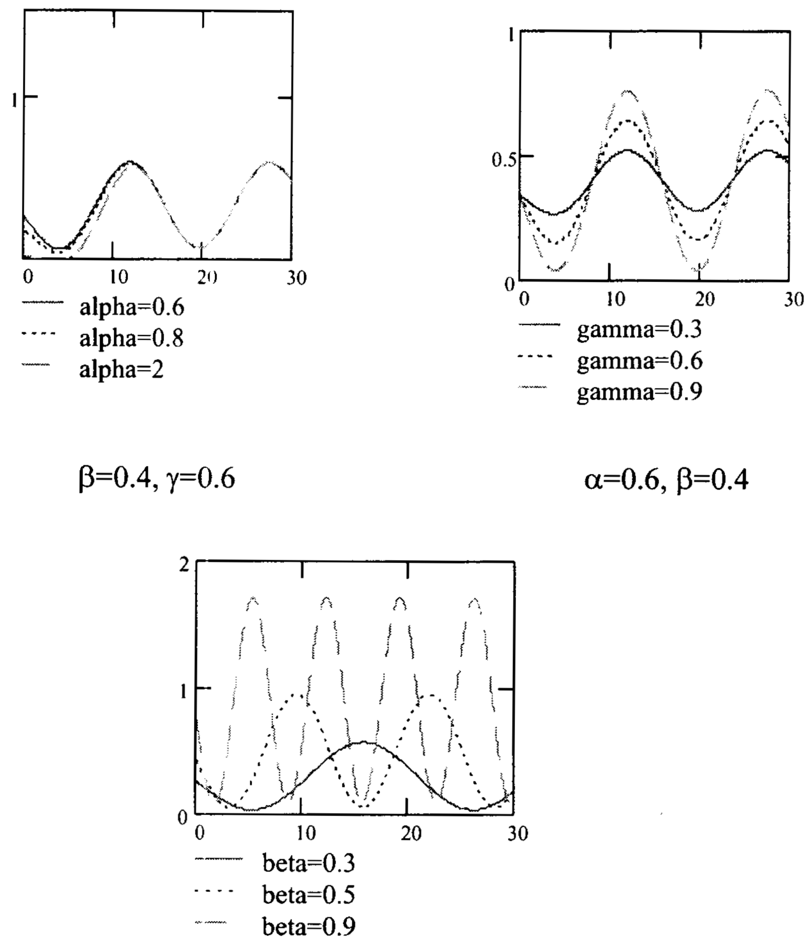
$\beta= 0.9, \gamma=0.6$



$\alpha=0.6, \gamma=0.9$

Figure 3.2.1 Density plots of GHSL.

The density plots exhibits more peakedness when the value of any parameter is increased, while the other two held fixed. The parameter γ enhances the symmetric nature of the plot.

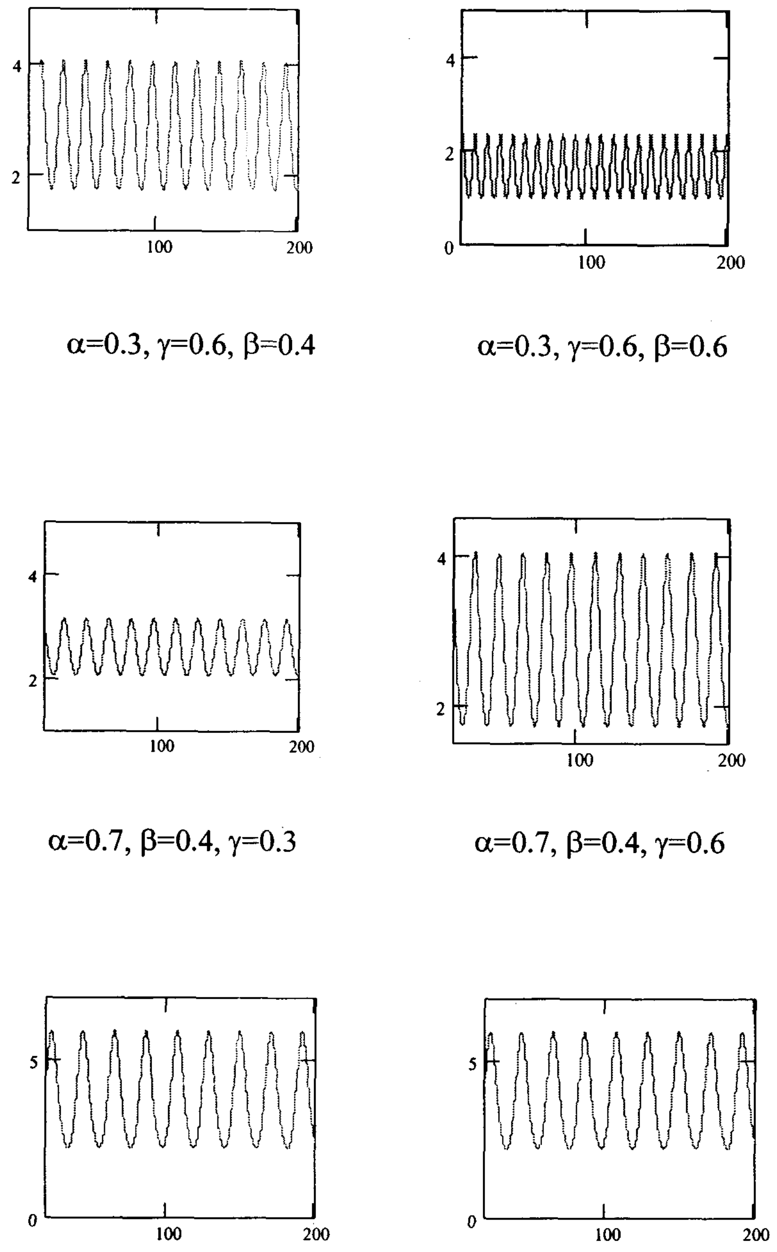


β=0.4, γ=0.6 α=0.6, β=0.4

α=0.6, γ =0.9

Figure 3.2.2 Hazard rate of GHSL distribution.

The hazard rate of GHSL distribution is periodic in nature. This property of the distribution makes it more useful in many situations, for instance, modeling of units under maintenance or replacement wherein the rate is usually periodic.



α=0.3, γ=0.6, β=0.4

α=0.3, γ=0.6, β=0.6

α=0.7, β=0.4, γ=0.3

α=0.7, β=0.4, γ=0.6

β=0.3, γ=0.7, α=0.6

β=0.3, γ=0.7, α=2

Figure 3.2.3 Mean Residual Life Function of GHSL distribution.

The MRL function of the GHSL distribution is seen to have a clear periodic nature. The amplitude appears to be smaller as the value of β increases. As the value of γ increases the amplitude becomes longer, but β makes no change.

From the foregoing discussion, it is clear that the GHSL distribution is a suitable model when the data exhibit non-monotone failure rate. The use of odds ratio and proportional odds has been common in reliability and survival analysis when the data exhibit non-proportional hazards. However, there are situations where this modeling is not suitable. Wang et al. (2003) proposed the log odds rate to characterize the distribution of failure time. The log odds rate may be viewed as a new way of modeling failure processes.

For the GHSL distribution, the log odds function is

$$\ln \frac{F(x)}{\overline{F}(x)} = \ln\left[\frac{\psi(x)-1}{2\alpha}\right].$$

The log odds rate is LOR (t) = $\dfrac{f(t)}{\overline{F}(t)F(t)}$

$$= \frac{2\alpha\beta\, e^{\beta x+\gamma(\cos\beta x-1)}\left[1-\gamma\sin\beta x\right]}{e^{\beta x+\gamma(\cos\beta x-1)}-1}.$$

Now we obtain a property of GHSL under geometric maximization.

Theorem 3.2.1

Let $X_1, X_2, \ldots$ be i.i.d. random variables with common distribution function F and N is geometric (α) distribution with P(N=n) = $\alpha(1-\alpha)^{n-1}$. Then the distribution of max $(X_1, X_2, \ldots, X_N)$ is half semi-logistic if and only if F is GHSL.

Proof:

Suppose $\overline{F} = \dfrac{2\alpha}{\psi(x) + 2\alpha - 1}$.

Then $P(\max(X_1, X_2, ..., X_N) \le x) = \sum_{n=1}^{\infty} [F(x)]^n \alpha(1-\alpha)^{n-1} = \dfrac{\psi(x) - 1}{\psi(x) + 1}$,

which is the distribution function of half semi-logistic .

Conversely suppose $P(\max(X_1, X_2, ...X_N) \le x) = \dfrac{\psi(x) - 1}{\psi(x) + 1}$.

Then, $\sum_{n=1}^{\infty} [F(x)]^n \alpha(1-\alpha)^{n-1} = \dfrac{\psi(x) - 1}{\psi(x) + 1}$.

Therefore, $\overline{F}(x) = \dfrac{2\alpha}{\psi(x) + 2\alpha - 1}$.

Hence the Theorem. □

3.3 First Order Autoregressive Generalized Half Semi-Logistic Process

A random variable X is said to be a universal geometric minimum (u.g.m.) if for every $p \in (0,1)$ there exists a sequence of i.i.d. random variables $X_1^{(p)}, X_2^{(p)}, ...,$ such that

$$X \stackrel{d}{=} \min_{1 \le i \le N(p)} \left\{X_i^{(p)}\right\}$$

where $P(N(p) = k) = p(1-p)^{k-1}$, $k = 1,2,...$ and X, N(p), $X_i^{(p)}, i = 1,2,\cdots, N(p)$ are independent (see Pillai et al. (1995)).

Here we obtain a stationary autoregressive minification process with GHSL distribution as marginals. The following theorem gives a necessary and sufficient condition for an autoregressive minification process to have GHSL distribution as stationary marginal.

Theorem 3.3.1

Let X_0 has the distribution GHSL(α, β, p). For n = 1,2, ... define

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad \alpha \\ \min(X_{n-1}, \varepsilon_n) & w.p. \quad (1-\alpha) \end{cases} \tag{3.3.1}$$

Then $\{X_n\}$ defines a stationary GHSL(α, β, p) first order autoregressive process if and only if ε_n's are i.i.d. HSL(α, β, p) random variables.

Proof:

Assume that ε_n's are i.i.d. HSL random variables and X_0 has GHSL (α, β, p) distribution.

Equation (3.3.1) in terms of survival function is

$$\bar{F}_{X_n}(x) = \bar{F}_{\varepsilon_n}(x)\left(\alpha + (1-\alpha)\bar{F}_{X_{n-1}}(x)\right).$$

When n = 1, the above equation becomes,

$$\bar{F}_{X_1}(x) = \bar{F}_{\varepsilon_1}(x)\left[\alpha + (1-\alpha)\frac{2\alpha}{\psi(x)+2\alpha-1}\right] = \frac{2\alpha}{\psi(x)+2\alpha-1}.$$

That is, X_1 has GHSL(α, β, p) distribution.

Similarly on assuming X_{n-1} has GHSL(α, β, p), we get the distribution of X_n also as GHSL$(\alpha, \beta . p)$.

Thus $\{X_n\}$ in (3.3.1) is stationary with GHSL distributions as marginals.

Conversely assume $\{X_n\}$ is stationary GHSL.

From (3.3.1), we get $\bar{F}_X(x) = \bar{F}_{\varepsilon_n}(x)\left(\alpha + (1-\alpha)\bar{F}_X(x)\right)$.

Thus $\overline{F_{\varepsilon_n}}(x) = \dfrac{2}{1+\psi(x)}$.

Hence the Theorem. □

This model (3.3.1) was introduced by Pillai et al. (1995). The above Theorem shows that GHSL(α, β, p) is a marginal distribution of (3.3.1), such that (3.3.1) defines a stationary Markov process.

The distribution of the geometric minimum, of the stationary process $\{X_n\}$ in (3.3.1) is,

$$P\left(\min(X_1, X_2, \ldots, X_N) > x\right) = \frac{2\alpha p}{\psi(x) + 2\alpha p - 1}$$

That is, $T_N \underline{\underline{d}}$ GHSL $(\alpha p, \beta, \mathrm{p})$.

Using Pillai et al. (1995), we have X belongs to the u.g.m. class.

According to them a random variable X or distribution function $F_X(x)$ to belong to the u.g.m class if and only if the survival function $\bar{F}_X(x)$ can be written in the form,

$\bar{F}_X(x) = \dfrac{1}{1+m(x)}$, where $m(x)$ has completely monotone derivative with $m(0) = 0$ and $m(\infty) = \infty$.

Hence $\bar{F}(x)$ can be written as $\dfrac{1}{1+m(x)}$, where $m(x) = \dfrac{\psi(x)-1}{2\alpha}$ and $m(0) = 0,\ m(\infty) = \infty$.

Also we can generalize (3.3.1) to the kth order process.

Now we look into some properties of the stationary GHSL process.

The joint survival function of (X_n, X_{n+1}) is

$$\overline{F}_{X_n, X_{n+1}}(x, y) = \begin{cases} \dfrac{2}{1+\psi(y)}\left[\dfrac{2\alpha^2}{\psi(x)+2\alpha-1} + \dfrac{2\alpha(1-\alpha)}{\psi(y)+2\alpha-1}\right] & if \quad x < y \\ \dfrac{4\alpha}{\left(\psi(x)+2\alpha-1\right)} \dfrac{1}{\left(1+\psi(y)\right)} & if \quad x > y \end{cases}$$

$$P(X_{n+1} > X_n) = 2\alpha \int_0^\infty \frac{\psi(y)-1}{(\psi(y)+2\alpha-1)} \frac{\psi'(y)}{(1+\psi(y))^2} dy .$$

In **Figure 3.3.1**, the sample path of the process is presented for $\psi(x) = e^{\beta x + \gamma \cos(\beta x - 1)}$ for $\alpha = 0.75,\ \beta = 0.8,\ \gamma = 1.75$ and in **Figure 3.3.2** the joint distribution of (X_n , X_{n+1}) is given.

Next we have a result giving an expression to compute the autocorrelation for the minification process.

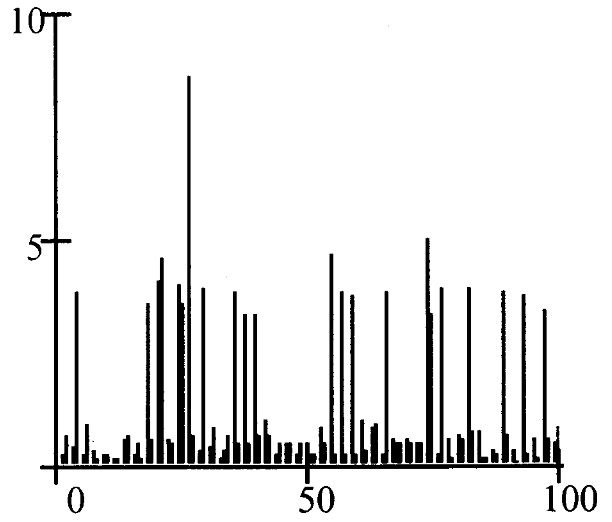


Figure 3.3.1 *Sample path of the process* $\{X_n\}$.

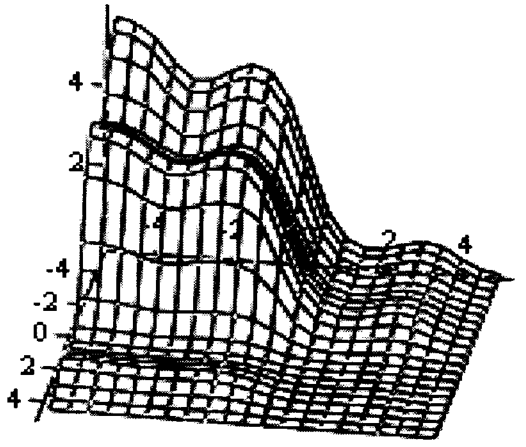


Figure 3.3.2 *Joint distribution of* (X_n, X_{n+1}).

Result 3.3.1

The autocorrelation between X_n, X_{n-1} for the minification process (3.3.1) is given by

$$\rho_X(1) = Corr(X_n, X_{n-1}) = \frac{(1-\alpha)E\left[(X_{n-1} - E(X_{n-1})) \int_0^{X_{n-1}} \bar{F}_\varepsilon(y)\,dy\right]}{V(X_n)}$$

Proof:

$$E(X_n X_{n-1}) = E(E(X_n X_{n-1} / X_{n-1}))$$

$$E(X_n X_{n-1} / X_{n-1} = x) = xE(X_n / X_{n-1} = x)$$

$$= x\alpha\, E(\varepsilon_n) + x(1-\alpha)\, E(\min(X_{n-1}, \varepsilon_n) / X_{n-1} = x)$$

$$= x\alpha\, E(\varepsilon_n) + x(1-\alpha) \int_0^x P(\varepsilon_n > y)\, dy$$

Therefore, $$E(X_n X_{n-1} / X_{n-1}) = X_{n-1}\, \alpha\, E(\varepsilon_n) + (1-\alpha) X_{n-1} \int_0^{X_{n-1}} P(\varepsilon_n > y)\, dy$$

That is, $$E(X_n X_{n-1}) = \alpha\, E(X_{n-1})\; E(\varepsilon_n) + (1-\alpha) E\left[X_{n-1} \int_0^{X_{n-1}} P(\varepsilon_n > y)\, dy\right]$$

Now $$E(X_n) = E(E(X_n / X_{n-1}))$$

$$= \alpha\, E(\varepsilon_n) + (1-\alpha) E\left(\int_o^{X_{n-1}} P(\varepsilon_n > y)\, dy\right)$$

$$Cov(X_n X_{n-1}) = \alpha\; \mathrm{cov}(\varepsilon_n X_{n-1}) + (1-\alpha) E\left[(X_{n-1} - E(X_{n-1})) \int_0^{X_{n-1}} \bar{F}_\varepsilon(y)\, dy\right]$$

$$= (1-\alpha)E\left[\left(X_{n-1} - E(X_{n-1})\right) \int_0^{X_{n-1}} \overline{F}_\varepsilon(y)\, dy\right].$$

Hence $\rho_X(1) = Corr(X_n, X_{n-1}) = \dfrac{(1-\alpha)E\left[\left(X_{n-1} - E(X_{n-1})\right) \int_0^{X_{n-1}} \overline{F}_\varepsilon(y)\, dy\right]}{V(X_n)}$. □

3.4 Some Special Classes of Distributions.

Now let us relax the condition on $\psi(x)$ taken in (3.2.1), but only assume that $\psi(x)$ is a monotone increasing function with $\psi(0) = 1$ and $\psi(\infty) = \infty$. Then

$$\overline{F}(x) = \frac{2\alpha}{\psi(x) + 2\alpha - 1} \tag{3.4.1}$$

gives the corresponding Marshall-Olkin form.

We shall investigate the properties of this new class for various choices of $\psi(x)$.

Case 1

In (3.4.1) if we put $\psi(x) = 1 + x^\beta, \beta > 0, x > 0$, we get Pareto distribution, with survival function,

$$\overline{G_1}(x) = \frac{2\alpha}{2\alpha + x^\beta} = \frac{1}{1 + \frac{1}{2\alpha} x^\beta}.$$

Case 2

In (3.4.1), if we put $\psi(x) = e^{x^\beta}$, then $\overline{G}(x)$ becomes

$$\overline{G}_2(x) = \frac{2\alpha}{e^{x^\beta} + 2\alpha - 1} \quad ,\beta > 0, \alpha > 0, x > 0. \tag{3.4.2}$$

This is a generalization of Weibull (Exponential power) distribution. We call this distribution as generalized exponential power ($GEP\ (\alpha, \beta)$) distribution.

Properties of GEP distribution

i. Probability density function is, $g_2(x) = \dfrac{2\alpha\beta\, e^{x^\beta}\ x^{\beta-1}}{\left(e^{x^\beta} + 2\alpha - 1\right)^2}$.

Figure 3.4.1 shows the behaviour of probability density function for different values of α *and* β

ii. Median $= \left[\ln\left(2\alpha + 1\right)\right]^{1/\beta}$

iii. Hazard Rate $= \dfrac{\beta e^{x^\beta}\ x^{\beta-1}}{e^{x^\beta} + 2\alpha - 1}$.

Hazard rate plot is given in **Figure 3.4.2**

iv. Log odds function $= \ln\left(\dfrac{e^{x^\beta}\ x^{\beta-1}}{2\alpha}\right)$.

v. Log odds rate $= \dfrac{2\beta\alpha\, t^{\beta-1}}{1 - e^{-t^\beta}}$.

vi. Mean Residual Life Function= $(e^{x^\beta} + 2\alpha - 1) \displaystyle\int_t^\infty \frac{1}{e^{u^\beta} + 2\alpha - 1} du$

The graph of MRL function is given in **Figure 3.4.3**

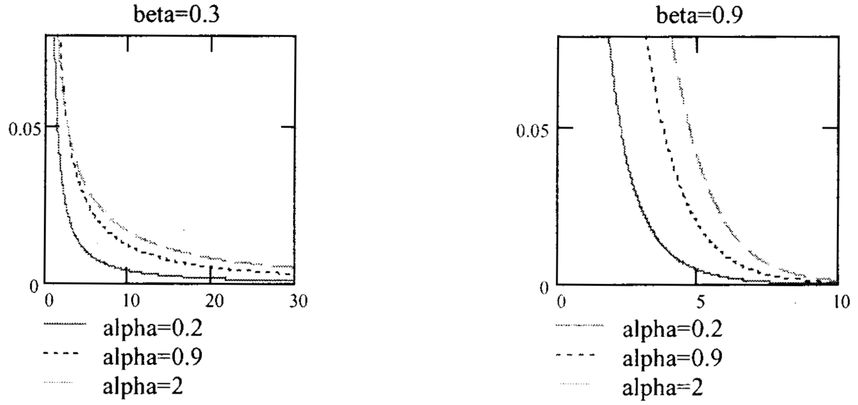


Figure 3.4.1 Density function of GEP distribution.

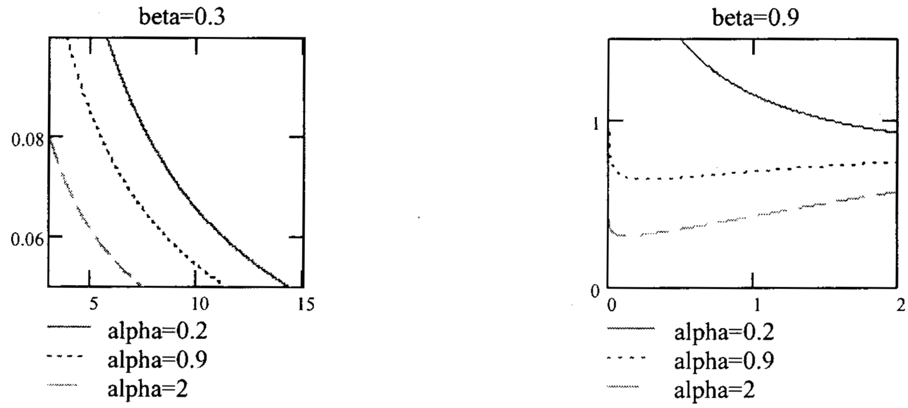


Figure 3.4.2 Hazard Rate of GEP distribution.

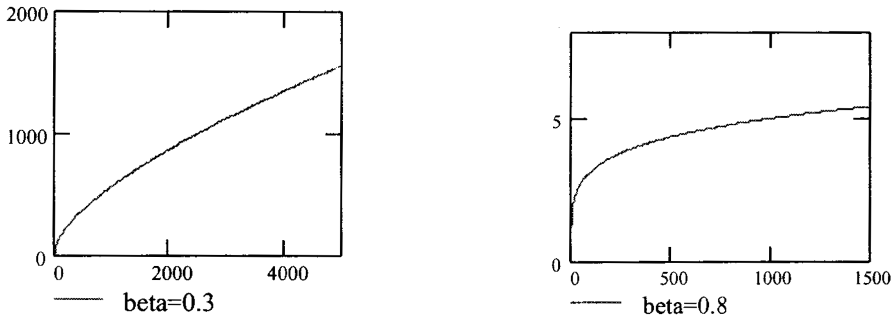


Figure 3.4.3 Mean Residual Life function of GEP distribution for α=0.6.

Remark 3.4.1: The AR(1) process with GEP distribution as marginals can be defined as follows:

In the AR(1) process (3.3.1) if $\{X_0\}$ follows GEP(α,β) and $\{\varepsilon_n\}$ has the survival function $\frac{2}{1+e^{x^\beta}}, 0<\beta<1$, then X_n' s are stationary with GEP(α,β) marginals.

Remark 3.4.2: If β=1 in (3.4.2), then $\overline{G}_2(x)$ becomes $\overline{G}_3(x)=\frac{2\alpha}{2\alpha+e^x-1}$, which is a generalization of the exponential distribution. Note that the distribution with survival function $\overline{G}_3(x)$ is called as generalized half logistic distribution (GHLD). The corresponding probability density function is

$$g_3(x)=\frac{2\alpha e^x}{\left(e^x+2\alpha-1\right)^2}. \tag{3.4.3}$$

Next we establish a characteristic property of GHLD.

Theorem 3.4.1

Let X be a random variable with $g(x)$ as probability density function. Then $g(x)$ is the probability density function of GHLD if and only if $g(x)$ satisfies the equation $g(x)=G(x)\overline{G}(x)+\frac{\overline{G}^2(x)}{2\alpha}$, where $G(x)$ and $1-G(x)$ denote respectively the distribution function and survival function.

Proof:

Suppose that $g(x)$ satisfies the equation

$$g(x)=G(x)\overline{G}(x)+\frac{\overline{G}^2(x)}{2\alpha}. \tag{3.4.4}$$

We know that any distribution function can be written in the form $G(x) = \dfrac{a(x)}{1+a(x)}$

for some function $a(x)$.

Substituting this in (3.4.4) and simplifying, we get

$$a(x) = \frac{e^x - 1}{2\alpha}.$$

Therefore $G(x) = \dfrac{e^x - 1}{e^x + 2\alpha - 1}$, which is the distribution of GHLD.

The converse part easily follows.

Hence the theorem. □

CHAPTER IV

GENERALIZED AUTOREGRESSIVE MINIFICATION PROCESSES

CHAPTER IV

GENERALIZED AUTOREGRESSIVE MINIFICATION PROCESSES

4.1 Introduction

Various autoregressive processes with minification structure have been considered in the Chapters II and III and their structural properties studied. In this Chapter, we investigate their genereralizations, viz. generalized autoregressive minification processes. Ristic (2007) has given a generalization of the semi-Pareto autoregressive minification process of first order. He introduced the minification process with the structure (1.6.7). He has shown that if X_0 is distributed as ε_1, where ε_n's are i.i.d. random variables with $SP(\beta, p)$ distribution, then X_n's are stationary $SP(\beta, p)$.

We develop a generalized autoregressive minification process from which we can construct any autoregressive minification process of required marginals. In Section 2, the generalized minification process is introduced. The properties of the process are studied in Section 3. Generalized semi-logistic process is introduced in Section 4. Section 5 deals with higher order processes. Generalized version of the half semi-logistic process is given in Section 6. Estimation of the parameters of the generalized process is given in the last Section.

4.2 Generalized Autoregressive Minification Process

Consider the process defined by,

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p_1 \\ \phi^{-1}(\phi(X_{n-1}) - \ln p_2) & w.p. \quad p_2(1-p_1) \\ \min\left(\phi^{-1}(\phi(X_{n-1}) - \ln p_2), \varepsilon_n\right) & w.p. \quad (1-p_2)(1-p_1) \end{cases} \qquad (4.2.1)$$

where $\phi(x) = \ln \frac{F(x)}{\bar{F}(x)}$ with $\phi(-\infty) = -\infty$, $\phi(\infty) = \infty, 0 < p_i < 1, i = 1,2$ and $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables, ε_n is independent of $X_i, i = 0,1,2,\cdots, n-1$.

We shall first discuss the stationarity of the process. Now following the steps similar to the proof of Theorem 2.2.1 we have the following assertion, which we state without proof.

Theorem 4.2.1

Let X_0 has distribution function F. The process $\{X_n\}$ given by (4.2.1) is strictly stationary if and only if ε_n 's are i.i.d. with distribution function F.

Proof: Similar to proof of Theorem 2.2.1 □

Theorem 4.2.2

If ε_n 's are i.i.d. with distribution function F and X_0 is arbitrary then $\{X_n\}$ in (4.2.1) converges in distribution to F.

Proof: Similar to that of Theorem 2.2.2. □

Next we seek a necessary and sufficient condition for $\{X_n\}$ in (4.2.1) to be stationary.

Let ε_n has distribution function H and $\{X_n\}$ be stationary. Let G_i be the distribution function of X_i, $i = 0,1,2,\cdots$.

Then from (4.2.1),

$$\overline{G}_1(x) = \overline{G}_0\left(\phi^{-1}(\phi(x) + \ln p_2)\right)[p_2(1-p_1) + (1-p_2)(1-p_1)\overline{H}(x)] + p_1\overline{H}(x).$$

Since $X_1 \underline{\underline{d}}\ X_0$, we may write

$$\overline{G}_0(x) = \overline{H}(x)[(1-p_2)(1-p_1)\,\overline{G}_0(\phi^{-1}(\phi(x) + \ln p_2) + p_1] \quad + \quad p_2(1-p_1)\,\overline{G}_0(\phi^{-1}(\phi(x) + \ln p_2).$$

Therefore, $\overline{H}(x) = \dfrac{\overline{G}_0(x) - \overline{G}_0\left(\phi^{-1}(\phi(x) + \ln p_2)\right)p_2(1-p_1)}{p_1 + (1-p_2)(1-p_1)\overline{G}_0\left(\phi^{-1}(\phi(x) + \ln p_2)\right)}$.

$$\text{That is } \overline{H}(x) = \frac{P(X_0 \geq x) - p_2(1-p_1)P\left(X_0 \geq \phi^{-1}(\phi(x) + \ln p_2)\right)}{p_1 + (1-p_2)(1-p_1)P\left(X_0 \geq \phi^{-1}(\phi(x) + \ln p_2)\right)}. \qquad (4.2.2)$$

Hence the above form of $\overline{H}(x)$ is necessary for stationarity of $\{X_n\}$.

Now assuming that (4.2.2) holds we have from (4.2.1),

$$\overline{G}_n(x) = \overline{G}_{n-1}\left(\phi^{-1}(\phi(x) + \ln p_2)\right)\left(p_2(1-p_1) + (1-p_2)(1-p_1)\overline{H}(x)\right) + p_1\overline{H}(x).$$

Putting n = 1, and using (4.2.2), we have

$$\overline{G}_1(x) = \overline{G}_0(x)$$

That is, $X_1 \underline{\underline{d}}\ X_0$.

Similarly assuming $X_{n-1} \underline{\underline{d}}\ X_0$, we get $X_n \underline{\underline{d}}\ X_0$

Hence the process is stationary.

Based on the above we have the following theorem.

Theorem 4.2.3

If the survival functions of ε_n 's are $\overline{H}$ then a necessary and sufficient condition for the autoregressive process $\{X_n\}$ in (4.2.1) to be stationary is

$$\overline{H}(x)=\frac{P(X_0\geq x)-p_2(1-p_1)P\left(X_0\geq \phi^{-1}(\phi(x)+\ln p_2)\right)}{p_1+(1-p_2)(1-p_1)P\left(X_0\geq \phi^{-1}(\phi(x)+\ln p_2)\right)}. \qquad \square$$

Now we study some properties of the generalized process defined by (4.2.1).

4.3 Properties of the Generalized Minification Process

The joint survival function of (X_n, X_{n+1}) is

$$\overline{F}_{X_n,X_{n+1}}(x,y)=P(X_n>x,X_{n+1}>y)$$

$$=\begin{cases}\dfrac{1+(1-p_1)e^{\phi(x)}}{(1+e^{\phi(x)})(1+e^{\phi(y)})} & \text{if} \quad x<\phi^{-1}(\phi(y)+\ln p_2)\\[2ex] \dfrac{1+p_2(1-p_1)e^{\phi(y)}}{(1+e^{\phi(x)})(1+e^{\phi(y)})} & \text{if} \quad x>\phi^{-1}(\phi(y)+\ln p_2)\end{cases}$$

In general,

$$\overline{F}_{X_n,X_{n+h}}(x,y)=\begin{cases}\dfrac{1+(1-p_1)^h e^{\phi(x)}}{(1+e^{\phi(x)})(1+e^{\phi(y)})} & \text{if} \quad x<\phi^{-1}\left(\phi(y)+\ln p_2{}^h\right)\\[2ex] \dfrac{1+p_2{}^h(1-p_1)^h e^{\phi(y)}}{(1+e^{\phi(x)})(1+e^{\phi(y)})} & \text{if} \quad x>\phi^{-1}\left(\phi(y)+\ln p_2{}^h\right)\end{cases} \qquad (4.3.1)$$

Also it can be proved that,

$$P(X_{n+1} > X_n) = \frac{1 + p_2(1-p_1)}{2}. \tag{4.3.2}$$

Next, we look into some examples leading to standard processes.

Example 4.3.1

If $\bar{F}(x) = \dfrac{1}{1+e^{\beta x}}$, $-\infty < x < \infty$ then $\phi(x) = \beta x$.

Therefore (4.2.1) becomes

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p_1 \\ X_{n-1} - \frac{1}{\beta}\ln p_2 & w.p. \quad p_2(1-p_1) \\ \min(X_{n-1} - \frac{1}{\beta}\ln p_2,\ \varepsilon_n) & w.p. \quad (1-p_2)(1-p_1) \end{cases} \tag{4.3.3}$$

If $X_0 = \varepsilon_1$ and ε_n's are i.i.d. logistic random variables then, $\{X_n\}$ is a sequence of stationary logistic random variables.

Example 4.3.2.

When $\bar{F}(x) = \dfrac{1}{1+x^{\beta}}$, $0 < x < \infty$,

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p_1 \\ p_2^{-1/\beta} X_{n-1} & w.p. \quad p_2(1-p_1) \\ \min(p_2^{-1/\beta} X_{n-1}, \varepsilon_n) & w.p. \quad (1-p_2)(1-p_1) \end{cases} \tag{4.3.4}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables with $SP(\beta, p_2)$ distribution. If $X_0 \stackrel{d}{=} \varepsilon_1$, then $\{X_n\}$ is a sequence of stationary $SP(\beta, p_2)$ random variables. This is the semi-Pareto process proposed by Ristic (2007).

Example 4.3.3

If X_n 's follow uniform distribution in (0,1), then (4.2.1) takes the form

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p_1 \\ \dfrac{X_{n-1}}{p_2 + (1-p_2)X_{n-1}} & w.p. \quad p_2(1-p_1) \\ \min\left(\dfrac{X_{n-1}}{p_2 + (1-p_2)X_{n-1}}, \varepsilon_n\right) & w.p. \quad (1-p_2)(1-p_1) \end{cases} \tag{4.3.5}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. uniform (0, 1) random variables. Thus (4.3.5) is a generalization of the uniform process defined in (2.3.11).

Now we consider the distributions of the maximum and minimum and also geometric minimum and geometric maximum of the process (4.2.1).

The distribution of $T_n = \min(X_1, X_2, ..., X_n)$ is given by its survival function as,

$$\bar{F}_{T_n}(x) = P(\min(X_1, X_2, ..., X_n) > x) .$$

$$= \frac{\left(1 + p_2(1-p_1)e^{\phi(x)}\right)^{n-1}}{\left(1 + e^{\phi(x)}\right)^n} . \tag{4.3.6}$$

The distribution of the geometric minimum when N is a random variable with p.m.f $P(N=n)=p_0(1-p_0)^{n-1}, 0<p_0<1, n=1,2,3,...$ is given by

$$\bar{F}_{T_N}(x)=P(\min(X_1,X_2,...X_N)\geq x)=\frac{1}{1+\frac{1-p_2(1-p_0)(1-p_1)}{p_0}e^{\phi(x)}}. \quad (4.3.7)$$

The distribution of the maximum is

$$F_{M_n}(x)=P(\max(X_1,X_2,...,X_n)\leq x)$$

$$=P(X_1\leq x,X_2\leq x,...X_n\leq x)$$

$$=\frac{e^{\phi(x)}\left(p_2(1-p_1)+e^{\phi(x)}\right)^{n-1}}{\left(1+e^{\phi(x)}\right)^n}. \quad (4.3.8)$$

The distribution of the geometric maximum is given by

$$F_{M_N}(x)=P(\max(X_1,X_2,...,X_N)\leq x)$$

$$=\frac{p_0\,e^{\phi(x)}}{1-p_2(1-p_0)(1-p_1)+p_0\,e^{\phi(x)}}. \quad (4.3.9)$$

Now let us check the ergodicity and mixing property of (4.2.1). For this we need the following definition.

Definition 4.3.1: A sequence $\{X_n\}$ of random variables is said to be uniformly mixing (Φ-mixing) if,

$$|P(A\cap B)-P(A)P(B)|\leq P(A)\Phi(h)$$

for all $A \in \sigma(X_0, X_1, ..., X_n), B \in \sigma(X_{n+h}, X_{n+h+1}, ...,)$, where $\sigma(X_0, X_1, ..., X_n)$ is the minimal sigma field induced by $\{X_0, X_1, \cdots, X_n\}$ and $\Phi(h) \to 0$ as $h \to \infty$. (See Billingsley (1986)). □

Also for ease of calculation define a new sequence $\{\eta_n\}$ of random variables as,

$$\eta_n = \begin{cases} \infty & w.p. \quad p_2 \\ \varepsilon_n & w.p. \quad 1 - p_2 \end{cases} \tag{4.3.10}$$

Therefore (4.2.1) becomes,

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p_1 \\ \min\left(\phi^{-1}(\phi(X_{n-1}) - \ln p_2), \eta_n\right) & w.p. \quad (1 - p_1) \end{cases}$$

Then after repeated substitution $\{X_n\}$ can be written as

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p_1 \\ \min\left(\phi^{-1}(\phi(\varepsilon_{n-1}) - \ln p_2), \eta_n\right) & w.p. \quad p_1(1-p_1) \\ \vdots & \vdots \qquad \vdots \\ \min\left(\phi^{-1}(\phi(X_0) - n \ln p_2), \phi^{-1}(\phi(\eta_1) - (n-1)\ln p_2), ..., \eta_n\right) & w.p. \quad (1-p_1)^n \end{cases} \tag{4.3.11}$$

From (4.3.10) and (4.3.11) it is clear that the minimal σ-field induced by $\{X_0, X_1, ..., X_n\}$ is same as the minimal σ-field induced by $\{X_0, \varepsilon_1, ..., \varepsilon_n\}$. Consequently $\{X_n\}$ is ergodic (see Stout (1974), Balakrishna and Jacob (2003)).

Next we have a Lemma, which gives the mixing parameters.

Lemma 4.3.1

The minification process $\{X_n\}$ generated by (4.2.1) is uniformly mixing with mixing parameters,

$$\Phi(h) = \frac{-2p_2^{\,h}\ln p_2^{\,h}}{1-p_2^{\,h}}(1-p_1)^h \qquad h = 1, 2, \ldots$$

Proof:

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

$$P(A \cap B) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(A \cap B / X_n = x, X_{n+h} = y)\, dF_h(x, y)$$

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

$$\begin{aligned} |P(A\cap B) - P(A)P(B)| \le\ & P\left(X_{n+h} = \phi^{-1}(\phi(X_n) - h\ln p_2)\right) + \\ & \int_{-\infty}^{\infty}\int_{-\infty}^{\phi^{-1}(\phi(y)+h\ln p_2)} g(x)g(y)\,dx\,dy \quad + \\ & \int_{-\infty}^{\infty}\int_{\phi^{-1}(\phi(y)+h\ln p_2)}^{\infty} |g(x,y) - g(x)g(y)\,dx\,dy| \end{aligned} \qquad (4.3.12)$$

where $g(x, y)$ is the joint density function and g(x) is the marginal density function.

Now to find $P\left(X_{n+h} = \phi^{-1}\left(\phi(X_n) - \ln p_2^{\,h}\right)\right)$, we proceed as follows.

Using (4.2.1) and (4.3.10) we have,

$$X_{n+h} = \begin{cases} \varepsilon_{n+h} & w.p. \quad p_1 \\ \min\left(\phi^{-1}(\phi(X_{n+h-1}) - \ln p_2), \eta_{n+h}\right) & w.p. \quad (1-p_1) \end{cases}$$

Repeated substitution for X_n's gives,

$$X_{n+h} = \begin{cases} \varepsilon_{n+h} & w.p. \quad p_1 \\ \min\left(\phi^{-1}(\phi(\varepsilon_{n+h-1}) - \ln p_2), \eta_{n+h}\right) & w.p. \quad p_1(1-p_1) \\ \vdots & \vdots \quad \vdots \\ \min\left(\phi^{-1}\left(\phi(X_n) - \ln p_2{}^h\right), \phi^{-1}\left(\phi(\eta_{n+1}) - \ln p_2{}^{h-1}\right), \ldots, \eta_{n+h}\right) & w.p. \quad (1-p_1)^h \end{cases}$$

Therefore, $P\left(X_{n+h} = \phi^{-1}\left(\phi(X_n) - \ln p_2{}^h\right)\right)$

$$= (1-p_1)^h \left[P\begin{pmatrix} \phi^{-1}\left(\phi(X_n) - \ln p_2{}^h\right) < \phi^{-1}\left(\phi(\eta_{n+1}) - \ln p_2{}^{h-1}\right), \\ \cdots, \phi^{-1}\left(\phi(X_n) - \ln p_2{}^h\right) < \eta_{n+h}. \end{pmatrix} \right]$$

$$= (1-p_1)^h \int_{-\infty}^{\infty} \prod_{j=1}^{h} P\left(X_n < \phi^{-1}\left(\phi(\eta_{n+j}) + \ln p_2{}^j\right)\right)$$

$$= (1-p_1)^h \int_{-\infty}^{\infty} \frac{e^{\phi(x)}\phi'(x)}{\left(1 + p^{-h}e^{\phi(x)}\right)\left(1 + e^{\phi(x)}\right)} dx$$

$$= \frac{(1-p_1)^h \left(-p_2{}^h . \ln p_2{}^h\right)}{1 - p_2{}^h} > 0. \tag{4.3.13}$$

From (4.3.1), for $x > \phi^{-1}(\phi(y) + h \ln p_2)$

$$g(x,y) = \frac{\left(1 - p_2{}^h(1-p_1)\right)^h \phi'(x)\, e^{\phi(x)} \phi'(y)\, e^{\phi(y)}}{\left(1 + e^{\phi(x)}\right)^2 \left(1 + e^{\phi(y)}\right)^2} \tag{4.3.14}$$

Substituting (4.3.13), (4.3.14), g(x) and g(y) in (4.3.12) and after simplifications, we get

$$\Phi(h) = \frac{-2{p_2}^h \ln {p_2}^h}{1-{p_2}^h}(1-p_1)^h .$$

As $h \to \infty, \Phi(h) \to 0$ and hence the theorem. □

Now let us consider a particular case of (4.2.1) and we can see that it is a generalization of semi-logistic process by Jayakumar and Mathew (2004).

4.4 Generalized Semi-Logistic Process

Consider the process $\{X_n\}$ defined by

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p_1 \\ X_{n-1} - \ln p_2 & w.p. \quad p_2(1-p_1) \\ \min(X_{n-1} - \ln p_2, \varepsilon_n) & w.p. \quad (1-p_2)(1-p_1) \end{cases} \qquad (4.4.1)$$

where $\{\varepsilon_n\}$is a sequence of i.i.d. $SL(\beta, p_2)$ random variables, X_{n-1} *and* ε_n are independent random variables, $0 < p_1, p_2 < 1$ and $\beta > 0$. If $X_o \overset{d}{=} \varepsilon_1$ then the process $\{X_n\}$ is defined as semi-logistic process with parameter (β, p_1, p_2).

The following theorem gives a necessary and sufficient condition for the process (4.4.1) to be stationary.

Theorem 4.4.1

Let X_o be distributed as ε_1. Then the process $\{X_n\}$ in (4.4.1) is stationary SL (β, p_2) marginals if and only if $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables with common distribution $SL(\beta, p_2)$.

Proof: Similar to Proof of Theorem 4.2.1. □

Remark 4.4.1: If $p_1 = 0,$ then the process (4.4.1) is the semi-logistic process of Jayakumar and Mathew (2004).

Remark 4.4.2: All other results of the process (4.2.1) can be easily verified to be holding well in the case of SLP (β, p_2) by replacing $e^{\phi(x)}$ by $\psi(x)$.

Remark 4.4.3: The generalized moving average model can be constructed on similar lines as given below.

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p_1 \\ \phi^{-1}(\phi(\varepsilon_{n-1}) - \ln p_2) & w.p. \quad p_2(1-p_1) \\ \min\left(\phi^{-1}(\phi(\varepsilon_{n-1}) - \ln p_2), \varepsilon_n\right) & w.p. \quad (1-p_2)(1-p_1) \end{cases}$$

where $\phi(x) = \ln \frac{F(x)}{\bar{F}(x)}$ with $\phi(-\infty) = -\infty$, $\phi(\infty) = \infty, 0 < p_i < 1\ (i = 1,2)$ and $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables with distribution function F. This model is also easily tractable and all properties can be studied as the AR process.

4.5 Higher Order Processes

Our aim here is to extend the model (4.2.1) to higher order cases. We propose the model,

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p_1 \\ \min\left(\phi^{-1}(\phi(X_{n-1}) - \ln p_2), \varepsilon_n\right) & w.p. \quad p_2(1-p_1) \\ \min\left(\phi^{-1}(\phi(X_{n-2}) - \ln p_2), \varepsilon_n\right) & w.p. \quad (1-p_2)(1-p_1) \end{cases} \tag{4.5.1}$$

where $\phi(x) = \ln \dfrac{F(x)}{\overline{F}(x)}$, $\phi(-\infty) = -\infty$, $\phi(\infty) = \infty$, $0 < p_1, p_2 < 1$.

The distribution of ε_n 's is given in the following theorem.

Theorem 4.5.1

If X_n 's in (4.5.1) are stationary with survival function $\dfrac{1}{1+e^{\phi(x)}}$, $-\infty < x < \infty$, then,

$$\varepsilon_n = \begin{cases} X_n & w.p. \quad \dfrac{1-p_2}{1-p_2 p_1} \\ \phi^{-1}(\phi(X_n) - \ln p_2 p_1) & w.p. \quad \dfrac{p_2(1-p_1)}{1-p_2 p_1} \end{cases} \tag{4.5.2}$$

Proof:

Suppose X_n 's are stationary with survival function ,

$$\overline{F}(x) = \frac{1}{1+e^{\phi(x)}} \ .$$

Then from (4.5.1),

$$\overline{F}_X(x) = \overline{F}_{\varepsilon_n}(x)\left(p_1 + p_2(1-p_1)\frac{1}{1+p_2 e^{\phi(x)}} + (1-p_2)(1-p_1)\frac{1}{1+p_2 e^{\phi(x)}} \right)$$

That is, $\overline{F}_{\varepsilon_n}(x) = \dfrac{1+p_2 e^{\phi(x)}}{\left(1+e^{\phi(x)}\right)\left(1+p_1 p_2 e^{\phi(x)}\right)}$

On simplification, we get

$$\overline{F}_{\varepsilon_n}(x) = \frac{1-p_2}{1-p_2 p_1}\frac{1}{1+e^{\phi(x)}} + \frac{p_2(1-p_1)}{1-p_2 p_1}\frac{1}{1+p_1 p_2 e^{\phi(x)}} .$$

From this, it is clear that ε_n 's can be written as a mixture of two random variables such that,

$$\varepsilon_n = \begin{cases} X_n & w.p. \quad \dfrac{1-p_2}{1-p_2p_1} \\ \phi^{-1}\left(\phi(X_n) - \ln p_2p_1\right) & w.p. \quad \dfrac{p_2(1-p_1)}{1-p_2p_1} \end{cases}$$

Hence the theorem. □

To study the autocorrelation structure of (4.5.1), we ignore ϕ and take a particular case of (4.5.1).

4.5.1 Autocorrelation Structure

Consider a particular case of (4.5.1) with the following structure

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p_1 \\ \min(X_{n-1} - \ln p_2, \varepsilon_n) & w.p. \quad p_2(1-p_1) \\ \min(X_{n-2} - \ln p_2, \varepsilon_n) & w.p. \quad (1-p_2)(1-p_1) \end{cases} \tag{4.5.3}$$

Then,

$$E(X_nX_{n-2}) = p_1E(\varepsilon_nX_{n-2}) + p_2(1-p_1)E(\min(X_{n-1} - \log p_2, \varepsilon_n)X_{n-2}) +$$

$$(1-p_2)(1-p_1)E(\min(X_{n-2} - \log p_2, \varepsilon_n)X_{n-2})$$

$$= \quad p_1E(\varepsilon_n)E(X_{n-2})$$

$$+$$

$$p_2(1-p_1)\left[\iint\limits_{x\,y\;y-\ln p_2<z} \int (y-\ln p_2)x\,f(x)f(y)f(z)\,dx\,dy\,dz + \iint\limits_{x\,y\;y-\ln p_2>z} \int z\,x\,f(x)f(y)f(z)\,dx\,dy\,dz\right]$$

$$+$$

$$(1-p_2)(1-p_1)\left[\int_x \int_{x-\ln p_2<z} (x-\ln p_2) f(x) f(z)\, dx\, dz + \int_x \int_{x-\ln p_2>z} z\, f(x) f(z)\, dx\, dz\right].$$

From (4.5.3) the expression for $E(\varepsilon_n)$ can be calculated. Substituting this in the above equation and simplifying we get,

$$Cov(X_n, X_{n-2}) = p_2(1-p_1)\, Cov(\min(X_{n-1}-\ln p_2, \varepsilon_n)\, X_{n-2})$$

$$+ (1-p_2)(1-p_1)\, Cov(\min(X_{n-2}-\ln p_2, \varepsilon_n)\, X_{n-2}).$$

This implies the following relationship of the correlation functions,

$$\rho_h = p_2(1-p_1)\,\rho_{1,h} + (1-p_2)(1-p_1)\,\rho_{2,h} \quad , \; h \geq 1$$

where $\rho_h = Corr(X_n, X_{n-h})$, $\rho_{r,h} = Corr(\min(X_{n-r}-\ln p, \varepsilon_n), X_{n-h})$, $r = 1,2$, $h = 1,2$.

The second order moving average model corresponding to (4.5.1) can be obtained as

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p_1 \\ \min\left(\phi^{-1}(\phi(\varepsilon_{n-1})-\ln p_2), \varepsilon_n\right) & w.p. \quad p_2(1-p_1) \\ \min\left(\phi^{-1}(\phi(\varepsilon_{n-2})-\ln p_2), \varepsilon_n\right) & w.p. \quad (1-p_2)(1-p_1) \end{cases} \tag{4.5.4}$$

where ε_n's are given by,

$$\varepsilon_n = \begin{cases} X_n & w.p. \quad \dfrac{1-p_2}{1-p_2p_1} \\ \phi^{-1}(\phi(X_n) - \ln p_2p_1) & w.p. \quad \dfrac{p_2(1-p_1)}{1-p_2p_1} \end{cases}$$

4.6 Generalized Half Semi-Logistic Processes

We have defined half semi-logistic process in Chapter III. Now we present a generalized process called generalized half semi-logistic process with the structure,

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p_1 \\ X_{n-1} & w.p. \quad p_2(1-p_1) \\ \min(X_{n-1}, \varepsilon_n) & w.p. \quad (1-p_2)(1-p_1) \end{cases} \tag{4.6.1}$$

It can be readily verified that if X_n's are stationary $GHSL(\alpha,\beta,p_2)$ with survival function, $\overline{F}X_n(x) = \dfrac{2\alpha}{\psi(x)+2\alpha-1}$ then the ε_n's are i.i.d. with survival function,

$$\overline{F}\varepsilon_n(x) = \frac{2\alpha\left(\dfrac{1-p_2+p_2p_1}{p_1}\right)}{\psi(x)+2\alpha\left(\dfrac{1-p_2+p_2p_1}{p_1}\right)-1}.$$

Note that when $p_2 = 0$, the above reduces to the $GHSL(\alpha,\beta,p_1)$ process defined by (3.3.1).

The autocorrelation structure of the process can be obtained as

$\mathrm{Corr}(X_n, X_{n-1})$ =

$$\frac{p_2(1-p_1)+(1-p_2)(1-p_1)E(X_{n-1})\left[\displaystyle\int_0^{X_{n-1}} P(\varepsilon_n > y)dy - \int_0^{\infty} P(X_{n-1} > y)P(\varepsilon_n > y)dy\right]}{V(X_n)}.$$

4.7 Estimation of the Parameters

Now let us estimate the parameters of the generalized process (4.2.1).

Note that $P(X_{n+1} > X_n) = \dfrac{1+p_2(1-p_1)}{2}$.

So we can estimate the parameters p_1, p_2 *and* ϕ using the procedure used in Chapter II.

Let $X_0, X_1, \cdots, X_N$ be a realization from the process (4.2.1).

Since we have one more parameter, p_2 we need another process defined as follows:

$$V_n = \begin{cases} 1 & \text{if} \quad \phi(X_{n+2}) - \phi(X_{n+1}) = \phi(X_{n+1}) - \phi(X_n) \\ 0 & \text{othewise} \end{cases}$$

Consider the process $\{U_n\}$ in (2.8.1).

Then $E(U_n) = P(X_{n+1} > X_n) = \dfrac{1+p_2(1-p_1)}{2}$.

Now $P(V_n = 1) = P(\phi(X_{n+2}) - \phi(X_{n+1}) = \phi(X_{n+1}) - \phi(X_n))$

$$= P\left(X_{n+2} = \phi^{-1}\left(\phi(X_n) - \ln p_2^{\,2}\right)\right).$$

Then using (4.3.13),

$$P(V_n = 1) = \frac{(1-p_1)^2}{1-p_2^{\,2}}\left(-p_2^{\,2} \ln p_2^{\,2}\right).$$

Hence $E(V_n) = \dfrac{(1-p_1)^2}{1-p_2^{\,2}}\left(-p_2^{\,2} \ln p_2^{\,2}\right)$.

Now the estimators $\hat{p}_1$ and $\hat{p}_2$ of p_1 and p_2 respectively, can be obtained by solving the equations,

$$\overline{U}_N = \frac{1}{N}\sum_{i=0}^{N-1} U_i = \frac{1+p_2(1-p_1)}{2} \text{ and } \overline{V}_{N-1} = \frac{1}{N-1}\sum_{i=0}^{N-2} V_i = \frac{(1-p_1)^2}{1-p_2^2}\left(-p_2^2 \ln p_2^2\right).$$

To estimate ϕ, we consider the random variable,

$$Z_n(t) = \begin{cases} 1 & if \quad X_n > t \\ 0 & otherwise \end{cases}$$

Since $\phi(x) = \ln \frac{F(x)}{\overline{F}(x)}$, we have $E(Z_n(t)) = P(X_n > t) = \frac{1}{1+e^{\phi(t)}}$.

Hence the estimator $\hat{\phi}(t)$ of $\phi(t)$ is given by,

$$\hat{\phi}(t) = \ln\left(\frac{1-\overline{Z}_{N+1}(t)}{\overline{Z}_{N+1}(t)}\right), \text{ where } \overline{Z}_{N+1}(t) = \frac{1}{N+1}\sum_{i=0}^{N} Z_i(t).$$

STUDY ON MINIFICATION PROCESSES

Thesis submitted to the

University of Calicut

for the award of the Degree of

DOCTOR OF PHILOSOPHY

under the Faculty of Science

by

KRISHNARANI, S.D.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALICUT
KERALA – 673 635
INDIA

December 2007

CHAPTER V

RANDOM COEFFICIENT MINIFICATION PROCESSES

CHAPTER V

RANDOM COEFFICIENT MINIFICATION PROCESSES

5.1 Introduction

The uses of logistic distribution have already been mentioned in the previous Chapters. It is mainly considered as a substitute for the normal distribution. From the similarity in shape of the normal and logistic distributions, the results of probit and logit analysis of the same data are usually similar. If a logistic distribution is used, in place of a normal distribution to represent the population tolerance distribution, then the analysis is carried out in terms of logits instead of probits.

Various non-Gaussian AR(1) models have been studied by many researchers with logistic marginals. In all these models, they have considered the innovation distribution as a non-Gaussian distribution. So all the models are of the form (1.3.1). Generalizations of these type of models are done by considering λ_i's to be random variables and the model can be written generally as

$$Y_n = V_n \, Y_{n-1} + \varepsilon_n \tag{5.1.1}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables. Here $\{V_n\}$ represents an i.i.d. sequence of random coefficients such that $E(V_n) = v$ and $\{V_n\}$ is independent of $\{\varepsilon_n\}$.

There are several approaches to construct these types of models. One may specify a known distribution for the innovation $\{\varepsilon_n\}$, then the stationary marginal density can be obtained. Else, one may give a known stationary marginal distribution for Y_n and attempt to find the distribution for $\{\varepsilon_n\}$. Lawrance and Lewis (1985) used the random coefficient autoregressive model to generate exponential random variables.

In the minification models given by (1.6.1) and (1.6.4), if the coefficients k or p is random, we call such a process as a random coefficient minification process. In this Chapter, we construct a general random coefficient minification process, such that most of the random coefficient minification processes are particular cases of this new model.

In Section 2, a new class of random coefficient first order minification processes is introduced and its properties are studied. In Section 3, a generalized random coefficient second order process is studied. Several examples are also given. A generalized process with marginals as distributions from Lehmann family is introduced and discussed in the last Section of this Chapter.

5.2 A New Class of Random Coefficient Minification Processes

Here we introduce a new class of random coefficient process, which generates several known minification models.

Let F(.) be a non-degenerate distribution function with $F(-\infty)=0$ and $F(\infty)=1$.

Consider the monotone transformation,

$$\phi(x) = \ln \frac{F(x)}{\bar{F}(x)} \tag{5.2.1}$$

where $\phi(-\infty) = -\infty, \phi(\infty) = \infty$ and $\bar{F}(x) = 1 - F(x)$.

Define the new process as,

$$X_n = \min\left(\phi^{-1}\left(\phi(X_{n-1}) - \ln V_n{}^{\beta}\right), \phi^{-1}\left(\phi(\varepsilon_n) - \ln V_n{}^{\beta}\right)\right), \quad n = 1,2,\ldots \tag{5.2.2}$$

where $\{V_n\}$ and $\{\varepsilon_n\}$ are i.i.d. sequence of random variables.

The following theorem gives the condition for stationarity of X_n's.

Theorem 5.2.1

Let $X_0 \overset{d}{=} \varepsilon_1$. Define

$$X_n = \min\left(\phi^{-1}\left(\phi(X_{n-1}) - \ln V_n{}^{\beta}\right), \phi^{-1}\left(\phi(\varepsilon_n) - \ln V_n{}^{\beta}\right)\right), \quad n = 1,2,\ldots$$

where $\{V_n\}$ and $\{\varepsilon_n\}$ are i.i.d. sequences of random variables such that V_n has the power function distribution $F_{V_n}(v) = v^{\beta}, \quad 0 < v < 1, \beta > 0$. Then the process $\{X_n\}$ is stationary if and only if ε_1 has distribution function F.

Proof:

Given $X_n = \min\left(\phi^{-1}\left(\phi(X_{n-1}) - \ln V_n{}^{\beta}\right), \phi^{-1}\left(\phi(\varepsilon_n) - \ln V_n{}^{\beta}\right)\right)$

Then $\bar{F}_{X_n}(x) = \bar{F}_{X_{n-1}}\left(\phi^{-1}\left(\phi(x) + \ln V_n{}^{\beta}\right)\right) \quad \bar{F}_{\varepsilon_n}\left(\phi^{-1}\left(\phi(x) + \ln V_n{}^{\beta}\right)\right)$

That is,

$$\bar{F}_{X_n}(x) = \int_0^1 \bar{F}_{X_{n-1}}\left(\phi^{-1}\left(\phi(x) + \ln v^{\beta}\right)\right)\bar{F}_{\varepsilon_n}\left(\phi^{-1}\left(\phi(x) + \ln v^{\beta}\right)\right) f_{V_n}(v)\, dv. \tag{5.2.3}$$

Assume that the process is stationary.

Since $X_0 \underline{\underline{d}}\ \varepsilon_1$ and $F_{V_n}(v) = v^{\beta}, 0 < v < 1, \beta > 0$,

$$\overline{F}_X(x) = \int_0^1 \overline{F}^2\left(\phi^{-1}\left(\phi(x) + \ln v^{\beta}\right)\right) \quad \beta v^{\beta-1}\, dv .$$

Therefore $e^{\phi(x)}\,\overline{F}_X(x) = \int_{-\infty}^{x} \overline{F}^2{}_X(t)\, e^{\phi(t)}\phi'(t)\, dt$.

Differentiating with respect to x,

$$\frac{\overline{F}'_X(x)}{\overline{F}^2{}_X(x)} + \frac{\phi'(x)}{\overline{F}_X(x)} = \phi'(x) \tag{5.2.4}$$

We know that every survival function can be written in the form,

$\overline{F}_X(x) = \dfrac{1}{1 + h(x)}$, where $h(x)$ is monotone increasing with $h(-\infty) = 0$ and $h(\infty) = \infty$.

Substituting this in (5.2.4) and simplifying we get,

$$\frac{h'(x)}{h(x)} = \phi'(x).$$

Integrating with respect to x,

$\ln h(x) = \phi(x)$. That is, $h(x) = e^{\phi(x)}$.

Therefore $\overline{F}_X(x) = \dfrac{1}{1 + e^{\phi(x)}}$. That is, X has distribution function F.

Hence ε_1 has distribution function F.

Conversely, assume that ε_1 follows F and $X_0 \underline{\underline{d}}\ \varepsilon_1$.

Then for n=1 (5.2.3) becomes,

$$\overline{F}X_1(x) = \int_0^1 \overline{F}X_0\left(\phi^{-1}\left(\phi(x) + \ln v^{\alpha}\right)\right) \quad \overline{F}\varepsilon_1\left(\phi^{-1}\left(\phi(x) + \ln v^{\alpha}\right)\right) f_{V_n}(v)\, dv$$

$$= \int_0^1 \overline{F}^{\,2}\left(\phi^{-1}\left(\phi(x) + \ln v^{\beta}\right)\right) \quad \beta\, v^{\beta-1}\, dv$$

$$= \int_0^1 \frac{1}{\left(1 + v^{\beta} e^{\phi(x)}\right)^2} \quad \beta\, v^{\beta-1}\, dv$$

$$= \frac{1}{1 + e^{\phi(x)}}.$$

That is, X_1 has distribution function F.

Assuming the distribution function of X_{n-1} as F, we can show that X_n also has distribution function F.

Therefore, by mathematical induction $\{X_n\}$ is stationary.

Hence the proof. □

Now we will look at some properties of (5.2.2).

The joint survival function of (X_n, X_{n+1}) is,

$$\overline{F}X_n,X_{n+1}(x,y) =$$

$$\int_0^1 P\left(X_n > \max\left(x, \phi^{-1}\left(\phi(y) + \ln v^{\beta}\right)\right)\right) P\left(\varepsilon_{n+1} > \phi^{-1}\left(\phi(y) + \ln v^{\beta}\right)\right) \beta\, v^{\beta-1} dv$$

$$= \int_0^1 \frac{1}{1 + e^{\max\left(\phi(x), \phi(y) + \ln v^{\beta}\right)}} \frac{1}{1 + v^{\beta} e^{\phi(y)}} \quad \beta\, v^{\beta-1}\, dv\,. \qquad (5.2.5)$$

$$P\left(X_{n+1} > X_n\right) = P\left(\varepsilon_{n+1} > \phi^{-1}\left(\phi(X_n) + \ln V_{n+1}{}^{\beta}\right)\right)$$

$$= \int_0^1 P\left(\varepsilon_{n+1} > \phi^{-1}\left(\phi(X_n) + \ln v^{\beta}\right)\right) \beta v^{\beta-1} \, dv$$

$$= \int_x \int_v P\left(\varepsilon_{n+1} > \phi^{-1}\left(\phi(x) + \ln v^{\beta}\right)\right) \beta v^{\beta-1} \, dv \quad dF(x)$$

$$= \int_x \phi'(x) \frac{\ln\left(1 + e^{\phi(x)}\right)}{\left(1 + e^{\phi(x)}\right)^2} dx. \tag{5.2.6}$$

The distribution of the minimum is,

$$\overline{F}_{T_n}(x) = P\left(\min(X_1, X_2, \ldots, X_n) > x\right)$$

$$= \frac{e^{-(n-1)\phi(x)}}{1 + e^{\phi(x)}} \ln\left(1 + e^{\phi(x)}\right)^{n-1}. \tag{5.2.7}$$

When N is a random variable with probability mass function, $P(N = n) = p_0 (1 - p_0)^{n-1}$, $n = 1, 2, \cdots$, the distribution of the geometric minimum is,

$$\overline{F}_{T_N}(x) = P\left(\min(X_1, X_2, \ldots, X_N) > x\right)$$

$$= \frac{p_0 \, e^{\phi(x)}}{\left(1 + e^{\phi(x)}\right)\left(e^{\phi(x)} - (1 - p_0) \ln\left(1 + e^{\phi(x)}\right)\right)}. \tag{5.2.8}$$

Following examples show that the process (5.2.2) is a generalization of several random coefficient autoregressive models.

Example 5.2.1

Let $\overline{F}(x) = \dfrac{1}{1 + e^{\beta x}}$, then (5.2.2) becomes

$$X_n = \min(X_{n-1} - \ln V_n, \varepsilon_n - \ln V_n) \tag{5.2.9}$$

where $\{V_n\}$ and $\{\varepsilon_n\}$ are i.i.d. sequence of random variables, V_n has the power function distribution $F_{V_n}(v) = v^{\beta}, \quad 0 < v < 1, \beta > 0$. Then the process $\{X_n\}$ is stationary logistic if and only if ε_1 is logistic.

Example 5.2.2

Consider the Pareto distribution with survival function, $\overline{F}(x) = \dfrac{1}{1+x^{\beta}}, x > 0, \beta > 0.$

Then (5.2.2) takes the form,

$$X_n = V_n^{-1} \min(X_{n-1}, \varepsilon_n) \tag{5.2.10}$$

Note that (5.2.10) is the random coefficient Pareto process when $\{\varepsilon_n\}$ are i.i.d. Pareto and V_n has the power function distribution $F_{V_n}(v) = v^{\beta}, \quad 0 < v < 1, \beta > 0$.

Example 5.2.3

When X_n's have uniform distribution with $\overline{F}(x) = 1 - x, \quad 0 < x < 1$ then (5.2.2) is of the structure,

$$X_n = \min\left(\frac{X_n}{V_n{}^{\beta} + (1 - V_n{}^{\beta})X_n}, \frac{\varepsilon_n}{V_n{}^{\beta} + (1 - V_n{}^{\beta})\varepsilon_n}\right). \tag{5.2.11}$$

We call this as the random coefficient uniform process where $\{\varepsilon_n\}$ is a sequence of i.i.d. uniform distributions and $\{V_n\}$ has the power function distribution with distribution function $F_{V_n}(v) = v^{\beta}, \quad 0 < v < 1, \beta > 0$.

Remark 5.2.1. A random coefficient moving average process can be defined with a structure similar to (5.2.2) as follows.

$$X_n = \min\left(\phi^{-1}\left(\phi(\varepsilon_n) - \ln V_n^{\beta}\right), \phi^{-1}\left(\phi(\varepsilon_{n-1}) - \ln V_n^{\beta}\right)\right) \quad n = 1,2,\ldots \qquad (5.2.12)$$

where $\{V_n\}$ and $\{\varepsilon_n\}$are i.i.d. sequence of random variables, $\{V_n\}$is independent of $\{\varepsilon_n\}$ with $F_{V_n}(v) = v^{\beta}, \quad 0 < v < 1, \beta > 0.$

All the results and properties of this process can be studied similar to the autoregressive process defined by (5.2.2).

5.3 A Generalized Random Coefficient Process

In Chapter IV, we have given generalized autoregressive minification processes. Here we define a generalized random coefficient minification process.

Definition 5.3.1. Let F(.) be a non-degenerate distribution function with $F(-\infty)=0$ and $F(\infty)=1$. Consider the monotone transformation,

$$\phi(x) = \ln \frac{F(x)}{\overline{F}(x)}$$

where $\phi(-\infty) = -\infty, \phi(\infty) = \infty$ and $\overline{F}(x) = 1 - F(x)$.

Define the new process as,

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p \\ \min\left(\phi^{-1}\left(\phi(X_{n-1}) - \ln V_n^{\beta}\right), \phi^{-1}\left(\phi(\varepsilon_n) - \ln V_n^{\beta}\right)\right) & w.p. \quad (1-p) \end{cases} \qquad (5.3.1)$$

where $\{V_n\}$ and $\{\varepsilon_n\}$ are i.i.d. sequence of random variables such that V_n has the power function distribution $F_{V_n}(v) = v^{\beta}, \quad 0 < v < 1, \beta > 0$ and $\{V_n\}$ independent of $\{\varepsilon_n\}$. Then $\{X_n\}$ is called a generalized random coefficient minification process.□

Next we propose a Theorem which gives the stationary distribution of X_n.

Theorem 5.3.1

Let $X_0 \underset{=}{d} \varepsilon_1$ and X_n be defined as (5.3.1) where $\{V_n\}$ and $\{\varepsilon_n\}$ are two i.i.d.. sequences of random variables such that V_n has a power function distribution with $F_{V_n}(v) = v^{\beta}, \quad 0 < v < 1, \beta > 0$. Then the process $\{X_n\}$ in (5.3.1) is stationary if and only if ε_1 has distribution function F.

Proof:

Let X_n be of the form (5.3.1). Then

$$\overline{F}X_n(x) = p\overline{F}\varepsilon_n(x) +$$

$$(1-p)\int_0^1 \overline{F}X_{n-1}\left(\phi^{-1}\left(\varphi(x) + \ln v^{\beta}\right)\right)\overline{F}\varepsilon_n\left(\phi^{-1}\left(\phi(x) + \ln v^{\beta}\right)\right)\beta\, v^{\beta-1}dv. \quad (5.3.2)$$

Given that $X_0 \underset{=}{d} \varepsilon_1$. Assume that $\{X_n\}$ is stationary.

Then (5.3.2) becomes,

$$\overline{F}X(x) = p\overline{F}X(x) + (1-p)\int_0^1 \overline{F}^2 X\left(\phi^{-1}\left(\phi(x) + \ln v^{\beta}\right)\right)\beta\, v^{\beta-1}\, dv.$$

That is, $\overline{F}_X(x) = p\overline{F}_X(x) + (1-p)\int_{-\infty}^{x} \overline{F}^2_X(t)\ e^{-\phi(x)}e^{\phi(t)}\phi'(t)dt$.

Hence, $e^{\phi(x)}\overline{F}_X(x) = \int_{-\infty}^{x} \overline{F}^2_X(t)\ e^{\phi(t)}\phi'(t)dt$.

Differentiating with respect to x we have,

$$\frac{\overline{F}'_X(x)}{\overline{F}^2_X(x)} + \frac{\phi'(x)}{\overline{F}_X(x)} = \phi'(x).$$

This is same as (5.2.4). We already obtained the solution of this equation as ,

$$\overline{F}_X(x) = \frac{1}{1+e^{\phi(x)}}.$$

Hence X has distribution function F. That is, ε_1 follows F.

Conversely, assume that ε_1 has distribution function F.

Therefore, $\overline{F}_{\varepsilon_1}(x) = \dfrac{1}{1+e^{\phi(x)}}$.

When n=1, (5.3.2) becomes

$$\overline{F}_{X_1}(x) = p\overline{F}_{\varepsilon_1}(x) + (1-p)\int_0^1 \overline{F}_{X_0}\left(\phi^{-1}\left(\varphi(x) + \ln v^{\beta}\right)\right)\overline{F}_{\varepsilon_1}\left(\phi^{-1}\left(\phi(x) + \ln v^{\beta}\right)\right)\beta\, v^{\beta-1}\, dv.$$

Assume $X_0 \underline{\underline{d}}\ \varepsilon_1$.

Then the above equation becomes,

$$\overline{F}_{X_1}(x) = p\frac{1}{1+e^{\phi(x)}} + (1-p)\int_0^1 \frac{1}{\left(1+v^{\beta}e^{\phi(x)}\right)^2}\ \beta\, v^{\beta-1}\, dv$$

$$= \frac{1}{1+e^{\phi(x)}} .$$

That is, X_1 follows F.

Assuming X_{n-1} follows F we can prove that X_n follows F.

Hence the process is stationary. □

Next we look at some of the properties of the process defined by (5.3.1).

The joint survival function is,

$$\overline{F}_{X_n, X_{n+1}}(x, y) = p \frac{1}{1+e^{\phi(x)}} \frac{1}{1+e^{\phi(y)}} +$$

$$(1-p) \int_0^1 \frac{1}{1+e^{\max\left(\phi(x), \phi(y)+\ln v^{\beta}\right)}} \frac{1}{\left(1+v^{\beta} e^{\phi(y)}\right)} \beta\, v^{\beta-1} dv . \quad (5.3.3)$$

$$P\left(X_{n+1} > X_n\right) = p \int_x P\left(\varepsilon_{n+1} > x\right) dF(x) + (1-p) \int_x \phi'(x) \frac{\ln\left(1+e^{\phi(x)}\right)}{\left(1+e^{\phi(x)}\right)^2} dx$$

$$= \frac{p}{2} + (1-p) \int_x \phi'(x) \frac{\ln\left(1+e^{\phi(x)}\right)}{\left(1+e^{\phi(x)}\right)^2} dx . \quad (5.3.4)$$

Following are some examples of the generalized random coefficient process.

Example 5.3.1

When $\overline{F}(x) = \frac{1}{1+e^{\beta x}}$, we have (5.3.1) as

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p \\ \min(X_{n-1} - \ln V_n, \varepsilon_n - \ln V_n) & w.p. \quad (1-p) \end{cases} \tag{5.3.5}$$

This is a random coefficient logistic process when $X_0 \underline{\underline{d}}\ \varepsilon_1$, where $\{V_n\}$ and $\{\varepsilon_n\}$ are two i.i.d.. sequences of random variables such that V_n has a power function distribution with $F_{V_n}(v) = v^{\beta}, \quad 0 < v < 1, \beta > 0$ and ε_n's are logistic.

Example 5.3.2

If $\overline{F}(x) = \dfrac{1}{1+x^{\beta}} \quad x > 0, \beta > 0$, then (5.3.1) is transformed into generalized random coefficient Pareto process with the structure,

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p \\ V_n^{-1} \min(X_{n-1}, \varepsilon_n) & w.p. \quad (1-p) \end{cases} \tag{5.3.6}$$

where $X_0 \underline{\underline{d}}\ \varepsilon_1, \{V_n\}$ and $\{\varepsilon_n\}$ are two i.i.d.. sequences of random variables such that V_n has a power function distribution with $F_{V_n}(v) = v^{\beta}, \quad 0 < v < 1, \beta > 0$ and ε_n's follow Pareto distribution.

Thus we can construct two parameter random coefficient minification models with any given distribution as marginals.

Remark 5.3.1: Note that when p=0 we have the model (5.2.2).

Naturally we may construct the corresponding moving average process with the structure,

$$X_n = \begin{cases} \varepsilon_n & w.p. \quad p \\ \min\left(\phi^{-1}\left(\phi(\varepsilon_n) - \ln V_n^{\beta}\right), \phi^{-1}\left(\phi(\varepsilon_{n-1}) - \ln V_n^{\beta}\right)\right) & w.p. \quad (1-p) \end{cases} \tag{5.3.7}$$

The model (5.3.1) can be extended to second order. We propose the second order model as,

$$X_n = \begin{cases} \min\left(\phi^{-1}\left(\phi(X_{n-1}) - \ln V_n^{\beta}\right), \phi^{-1}\left(\phi(\varepsilon_n) - \ln V_n^{\beta}\right)\right) & w.p. \quad 1-p_2 \\ \min\left(\phi^{-1}\left(\phi(X_{n-2}) - \ln V_n^{\beta}\right), \phi^{-1}\left(\phi(\varepsilon_n) - \ln V_n^{\beta}\right)\right) & w.p. \quad p_2 \end{cases} \tag{5.3.8}$$

where $\{V_n\}$ and $\{\varepsilon_n\}$are i.i.d. sequence of random variables such that V_n has the power function distribution $F_{V_n}(v) = v^{\beta}, \quad 0 < v < 1, \beta > 0$ and $\{V_n\}$independent of $\{\varepsilon_n\}$.

Then it can be verified that if $\{X_n\}$ is stationary with survival function $\dfrac{1}{1+e^{\phi(x)}}$, then ε_n follows F.

We can extend the above defined MA(1) model and AR(2) model to higher orders.

In the next Section we introduce a process, which can generate autoregressive minification process of first order with its stationary distribution as the one from Lehmann family.

5.4 A Generalized Random Coefficient Autoregressive Lehmann Process

Let $\overline{G}(.)$ be a survival function with Lehmann structure given by,

$$\overline{G}(x) = \frac{1}{[1+e^{\phi(x)}]^{\gamma+1}}. \qquad (5.4.1)$$

Consider the process with the structure

$$X_n = \min\left(\phi^{-1}\left(\phi(X_{n-1}) - \ln V_n{}^{\beta}\right), \varepsilon_n\right) \qquad n = 1,2,\ldots \qquad (5.4.2)$$

where $\{V_n\}$ and $\{\varepsilon_n\}$are two independent sequences of i.i.d. random variables such that V_n has the distribution $F_{V_n}(v) = v^{\beta\gamma}, \quad 0 < v < 1, \quad \beta, \gamma > 0.$

The process (5.4.2) is called the generalized random coefficient Lehmann process. We denote the distribution with the survival function (5.4.1) by $\mathrm{L}(\gamma)$.

Next we discuss the stationarity of the process (5.4.2).

Theorem 5.4.1

Let $\{X_n\}$ be defined by (5.4.2). Then if X_0 follows $L(\gamma)$ and ε_n has distribution function F, then the process $\{X_n\}$ is stationary with $\mathrm{L}(\gamma)$ marginals.

Proof:

Assuming the structure of $\{X_n\}$ given by (5.4.2) the survival function is,

$$\overline{G}_{X_n}(x) = \overline{F}_{\varepsilon_n}(x) \int_0^1 \overline{G}_{X_{n-1}}\left(\phi^{-1}\left(\phi(x) + \ln v^{\beta}\right)\right) \beta\gamma\, v^{\beta\gamma - 1} dv \,. \qquad (5.4.3)$$

Suppose X_0 follows $L(\gamma)$ and ε_n has distribution F.

Then for n=1, (5.4.3) becomes,

$$\overline{G}_{X_1}(x) = \overline{F}_{\varepsilon_1}(x) \int_0^1 \overline{G}_{X_0}\left(\phi^{-1}\left(\phi(x) + \ln v^{\beta}\right)\right) \beta\gamma\, v^{\beta\gamma - 1} dv$$

$$= \frac{1}{[1 + e^{\phi(x)}]^{r+1}} \,. \qquad (5.4.4)$$

That is, X_1 follows $L(\gamma)$.

Similarly we can prove that if X_{n-1} follows $L(\gamma)$, then X_n follows $L(\gamma)$.

Hence it follows by mathematical induction that $\{X_n\}$ is stationary with $L(\gamma)$ marginals. □

Theorem 5.4.2

Let $\{X_n\}$ be defined as (5.4.2), where $\{V_n\}$ and $\{\varepsilon_n\}$ are two independent sequences of i.i.d. random variables such that V_n has the distribution function $F_{V_n}(v) = v^{\beta\gamma}$, $0 < v < 1$, $\beta, \gamma > 0$. Suppose that the process is stationary. Then X_n follows $L(\gamma)$ if and only if ε_n follows F.

Proof:

Suppose that $\{X_n\}$ is stationary. Then from (5.4.3), we have

$$\overline{G}_X(x) = \overline{F}_{\varepsilon_n}(x)\int_0^1 \overline{G}_X\left(\phi^{-1}\left(\phi(x)+\ln v^{\beta}\right)\right)\beta\gamma\, v^{\beta\gamma-1}dv$$

$$= \overline{F}_{\varepsilon_n}(x)\int_{-\infty}^{x} \overline{G}_X(t)\,\gamma\, e^{\gamma\phi(t)}e^{-\gamma\phi(x)}\phi'(t)dt.$$

Therefore, $$\frac{e^{\gamma\phi(x)}\overline{G}_X(x)}{\overline{F}_{\varepsilon_n}(x)} = \int_{-\infty}^{x} \overline{G}_X(t)\,\gamma\, e^{\gamma\phi(t)}\phi'(t)dt. \qquad (5.4.5)$$

Differentiating both sides with respect to x, and simplifying we get,

$$\frac{\overline{G}'_X(x)}{\overline{G}_X(x)} - \frac{\overline{F}'_{\varepsilon_n}(x)}{\overline{F}_{\varepsilon_n}(x)} = -\gamma\phi'(x)\left(1-\overline{F}_{\varepsilon_n}(x)\right). \qquad (5.4.6)$$

Integrating the above now yields

$$\ln\overline{G}_X(x) - \ln\overline{F}_{\varepsilon_n}(x) = -\gamma\int\phi'(x)[1-\overline{F}_{\varepsilon_n}(x)]dx.$$

But since ε_n follows F, we have $\overline{F}_{\varepsilon_n}(x) = \dfrac{1}{1+e^{\phi(x)}}$. Then the above equation becomes,

$$\ln\overline{G}_X(x) = \ln\left(\frac{1}{[1+e^{\phi(x)}]^{r+1}}\right).$$

That is, $\overline{G}_X(x) = \left(\dfrac{1}{1+e^{\phi(x)}}\right)^{\gamma+1}$ and so X_n has distribution $L(\gamma)$.

Conversely assume $\{X_n\}$ is stationary and X_n has distribution $L(\gamma)$. Then it follows that,

$$\left(\frac{1}{1+e^{\phi(x)}}\right)^{\gamma+1} = \overline{F}_{\varepsilon_n}(x)\int_0^1\left(\frac{1}{1+v^{\alpha}e^{\phi(x)}}\right)^{\gamma+1}\beta\gamma v^{\beta\gamma-1}dv.$$

Therefore, $\overline{F}_{\varepsilon_n}(x)=\dfrac{1}{1+e^{\phi(x)}}$.

That is, ε_n has distribution function F.

Hence the proof of the Theorem. □

Also we derive the following expressions concerning the generalized random coefficient Lehmann autoregressive process.

$$\overline{F}_{X_n,X_{n+1}}(x,y)=P\left(X_n>x,\phi^{-1}(\phi(X_n)-\ln V_n{}^{\beta})>y,\varepsilon_n>y\right)$$

$$=P(\varepsilon_n>y)P\left(X_n>\max(x,\phi^{-1}(\phi(y)+\ln V_n{}^{\beta})\right)$$

$$=\frac{1}{1+e^{\phi(y)}}\int_0^1\left(\frac{1}{1+e^{\max\left(\phi(x),\phi(y)+\ln v^{\beta}\right)}}\right)^{\gamma+1}\beta\gamma\, v^{\beta\gamma-1}dv. \qquad (5.4.7)$$

$$P\left(X_{n+1}>X_n\right)=P\left(\varepsilon_n>X_n\right)$$

$$=\int_x P\left(\varepsilon_n>x\right)f_X(x)\,dx$$

$$=\frac{\gamma+1}{\gamma+2}. \qquad (5.4.8)$$

We can see that several minification models with marginals from the Lehmann family can be deduced as special cases from the process defined in (5.4.2). For instance we have the following generalized models already reported in the literature.

Example 5.4.1

On taking $\overline{F}(x)=\dfrac{1}{1+e^{\beta x}}$, (5.4.2) takes the form,

$$X_n=\min(X_{n-1}-\ln V_n,\varepsilon_n) \tag{5.4.9}$$

which generates generalized logistic marginals where $\{V_n\}$ and $\{\varepsilon_n\}$are two independent sequences of i.i.d. random variables such that V_n has the distribution $F_{V_n}(v)=v^{\beta\gamma},\quad 0<v<1,\quad \beta,\gamma>0$ and $\{\varepsilon_n\}$ follows logistic distribution.

Example 5.4.2

When $\overline{F}(x)=\dfrac{1}{1+x^{\beta}}\quad x>0,\beta>0$, we have

$$X_n=\min\left(V_n^{-1}X_{n-1},\varepsilon_n\right) \tag{5.4.10}$$

which generates generalized Pareto marginals where $\{\varepsilon_n\}$ has the Pareto survival function and $\{V_n\}$ has the power function distribution with distribution function $F_{V_n}(v)=v^{\beta\gamma},\quad 0<v<1,\quad \beta,\gamma>0$.

Example 5.4.3

If $\overline{F}(x) = 1 - x, \; 0 < x < 1$, then (5.4.2) becomes,

$$X_n = \min\left(\frac{X_{n-1}}{V_n{}^{\beta} + (1 - V_n{}^{\beta}) X_{n-1}}, \varepsilon_n \right) \tag{5.4.11}$$

which generates generalized uniform marginals when $\{\varepsilon_n\}$ is a sequence of uniform random variables and $\{V_n\}$ has the power function distribution with distribution function $F_{V_n}(v) = v^{\beta\gamma}, \quad 0 < v < 1, \quad \beta, \gamma > 0$.

CHAPTER VI

BIVARIATE LOGISTIC MINIFICATION PROCESSES

CHAPTER VI

BIVARIATE LOGISTIC MINIFICATION PROCESSES♠

6.1 Introduction

In many socio-economic contexts, the data are usually multivariate in nature. Several components like income, expenditure, area of land holdings etc. are taken into consideration. Another example for multivariate data is the collection of mean monthly temperatures recorded at scattered locations. Similarly the set of signals recorded by an array of seismometers in the aftermath of an earthquake or nuclear explosion is a multivariate time series. So to model these types of data we need multivariate time series processes.

Some additive and minification bivariate models are available in the literature. Block et al. (1988) introduced additive first order autoregressive bivariate exponential and geometric processes and studied their properties. Dewald et al. (1989) introduced an additive first order autoregressive bivariate exponential process. A bivariate integer valued moving average model is proposed by Quoreshi (2006) for modeling of financial count data. Bivariate semi α-Laplace distributions and processes are introduced in Kuttikrishnan and Jayakumar (2007). Ristic and Popovic (2003) introduced bivariate uniform autoregressive processes. Krishnarani and Jayakumar (2007c) introduced a bivariate semi-logistic distribution and developed autoregressive models with bivariate semi-logistic distribution as

---

♠ This Chapter is based on Krishnarani and Jayakumar (2007c).

marginals. Some bivariate minification models were already been given in Balakrishna and Jayakumar (1996, 1997), Alice and Jose (2004) and Ristic (2006). The structure of all these processes is given in the introductory Section 1.8.

Several univariate minification processes have been discussed in the previous Chapters. Now we introduce bivariate forms of these processes. In the next Section, we introduce a bivariate semi-logistic distribution and obtain some characterizations of the distribution. In Section 3, bivariate half semi-logistic distribution is introduced and studied. First order autoregressive minification processes with bivariate semi-logistic distributions as stationary distributions are discussed in Section 4. Section 5 deals with a bivariate logistic process as a particular case. A general process useful in generating any first order autoregressive minification process is given in Section 6. In the last Section bivariate half semi-logistic process is discussed.

6.2 Bivariate Semi-Logistic Distribution

A random vector (X,Y) has bivariate semi-logistic distribution if its survival function is of the form

$$\bar{F}(x,y)=\frac{1}{1+\psi(x,y)}, \quad -\infty<x,y<\infty, \tag{6.2.1}$$

where $\psi(x,y)$ satisfies the functional equation,

$$\psi(x,y)=\frac{1}{p}\psi\left(\frac{1}{\beta_1}\ln p+x,\frac{1}{\beta_2}\ln p+y\right). \tag{6.2.2}$$

The solution of this functional equation is $\psi(x,y) = e^{\beta_1 x} h_1(x) + e^{\beta_2 y} h_2(y)$ where $h_1(x)$ and $h_2(y)$ are periodic functions in x,y with periods $\frac{1}{\beta_1}\ln p$ and $\frac{1}{\beta_2}\ln p$ respectively. We denote this distribution as BVSL (β_1, β_2, p).

Now we obtain a characterization of BVSL (β_1, β_2, p) through geometric minimization.

Theorem 6.2.1

Let $\{(X_i, Y_i), i \ge 1\}$ be a sequence of i.i.d. bivariate random vectors with common survival function $\overline{F}(x,y)$ and N be a geometric random variable with $P(N=n) = p(1-p)^{n-1}$, where N is independent $(X_i, Y_i), i \ge 1$.. Define $U_N = \min(X_1, X_2, \ldots X_N)$ and $V_N = \min(Y_1, Y_2, \ldots Y_N)$. Then the random vectors $\left(U_N - \frac{1}{\beta_1}\ln p, V_N - \frac{1}{\beta_2}\ln p\right)$ and (X_i, Y_i) are identically distributed if and only if (X_i, Y_i) has BVSL (β_1, β_2, p) distribution.

Proof:

Suppose (X_i, Y_i) follows BVSL (β_1, β_2, p).

Then $\overline{G}(x,y) = P\left(U_N - \frac{1}{\beta_1}\ln p > x, V_N - \frac{1}{\beta_2}\ln p > y\right)$

$$= \frac{p\overline{F}\left(x + \frac{1}{\beta_1}\ln p, y + \frac{1}{\beta_2}\ln p\right)}{1-(1-p)\overline{F}\left(x + \frac{1}{\beta_1}\ln p, y + \frac{1}{\beta_2}\ln p\right)} \tag{6.2.3}$$

$$= \bar{F}(x,y).$$

Therefore $\left(U_N - \frac{1}{\beta_1}\ln p, V_N - \frac{1}{\beta_2}\ln p\right)$ and (X_i, Y_i) are identically distributed.

To prove the converse, suppose $\bar{G}(x,y) = \bar{F}(x,y)$. (6.2.4)

Note that a bivariate survival function can be written in the form

$$\frac{1}{1+\phi(x,y)}, -\infty < x, y < \infty \tag{6.2.5}$$

where $\phi(x,y)$ is a non-decreasing function in both x and y with

$$\lim_{x\to\infty} \lim_{y\to\infty} \phi(x,y) = \infty, \qquad \lim_{x\to-\infty} \lim_{y\to-\infty} \phi(x,y) = 0$$

$$\lim_{x\to-\infty} \phi(x,y) = \phi_1(y) \quad \text{and} \quad \lim_{y\to-\infty} \phi(x,y) = \phi_2(x).$$

Then using (6.2.3) and (6.2.4) and (6.2.5) we have,

$$\frac{1}{1+\phi(x,y)} = \frac{p\dfrac{1}{1+\phi(x+\frac{1}{\beta_1}\ln p, y+\frac{1}{\beta_2}\ln p)}}{1-(1-p)\dfrac{1}{1+\phi(x+\frac{1}{\beta_1}\ln p, y+\frac{1}{\beta_2}\ln p)}}.$$

Therefore, we get $\phi(x,y) = \frac{1}{p}\,\phi(x+\frac{1}{\beta_1}\ln p, y+\frac{1}{\beta_2}\ln p)$.

That is, $\phi(x,y) = \psi(x,y)$ and (X_i, Y_i) follows BVSL (β_1, β_2, p). □

Let $\{N_k, k \geq 1\}$ be a sequence of geometric random variables with parameters p_k, $0 \leq p_k \leq 1$. Define

$$\overline{F}_k(x,y) = P\left(U_{N_{k-1}} > x, V_{N_{k-1}} > y\right) \quad , \qquad k = 1,2,\ldots$$
$$= \frac{p_{k-1}\overline{F}_{k-1}(x,y)}{1 - q_{k-1}\overline{F}_{k-1}(x,y)} \tag{6.2.6}$$

where $\overline{F}_{k-1}(x,y)$ is the survival function of the minimum of geometric (p_{k-1}) random variables.

Theorem 6.2.2

Let $\{(X_i, Y_i), i \geq 1\}$ be a sequence of i.i.d. bivariate random vectors with common survival function $\overline{F}(x,y)$. Define $\overline{F_1} = \overline{F}$ and $\overline{F_k}$ as the survival function of the geometric (p_{k-1}) minimum of i.i.d. random vectors with common survival function $\overline{F}_{k-1}$, k= 1, 2, ... Then,

$$\overline{F}_k\left(x + \sum_{j=1}^{k-1}\frac{1}{\beta_1}\ln p_j,\, y + \sum_{j=1}^{k-1}\frac{1}{\beta_2}\ln p_j\right) = \overline{F}(x,y) \tag{6.2.7}$$

if and only if (X_1, Y_1) has BVSL (β_1, β_2, p) distribution.

Proof:

Suppose (X_1, Y_1) has BVSL (β_1, β_2, p) distribution and let

$$\overline{F}_k(x,y) = \frac{1}{1 + \phi_k(x,y)} \;. \tag{6.2.8}$$

Substituting in (6.2.6) and simplifying we get ,

$$\phi_k(x,y) = \frac{1}{\prod_{j=1}^{k-1} p_j}\phi_1(x,y).$$

Therefore,

$$\phi_k\left(x+\frac{1}{\beta_1}\sum_{j=1}^{k-1}\ln p_j, y+\frac{1}{\beta_2}\sum_{j=1}^{k-1}\ln p_j\right)=\frac{1}{\prod_{j=1}^{k-1}p_j}\phi_1\left(x+\frac{1}{\beta_1}\sum_{j=1}^{k-1}\ln p_j, y+\frac{1}{\beta_2}\sum_{j=1}^{k-1}\ln p_j\right). \quad (6.2.9)$$

Replacing ϕ by ψ and assuming that ψ satisfies (6.2.2), we have

$$\overline{F}_k\left(x+\sum_{j=1}^{k-1}\frac{1}{\beta_1}\ln p_j, y+\sum_{j=1}^{k-1}\frac{1}{\beta_2}\ln p_j\right)=\overline{F}(x,y).$$

To prove the converse part, assume that (6.2.7) is true.

Then using (6.2.8) and (6.2.9) we have

$$\phi(x,y)=\frac{1}{\prod_{j=1}^{k-1}p_j}\phi_1\left(x+\frac{1}{\beta_1}\sum_{j=1}^{k-1}\ln p_j, y+\frac{1}{\beta_2}\sum_{j=1}^{k-1}\ln p_j\right).$$

That is (X_1, Y_1) has BVSL (β_1, β_2, p) distribution.

This completes the proof. □

6.2.1 Some Examples

Now we give some examples of bivariate semi-logistic distribution.

Let (X,Y) be a random vector with survival function

$$\overline{F}(x,y)=\frac{1}{1+e^{\beta_1 x}+e^{\beta_2 y}}\quad,\quad -\infty<x,y<\infty\,,\beta_1,\beta_2>0.$$

Note that this is a special case of BVSL (β_1, β_2, p) since $\psi(x,y)= e^{\beta_1 x}+e^{\beta_2 y}$ satisfies (6.2.2). We call this distribution as bivariate logistic distribution denoted by BVL (β_1, β_2).

Another example is,

$$\bar{F}(x,y)=\frac{1}{1+e^{\beta_1 x}+e^{\beta_2 y}+e^{\min(\beta_1 x,\beta_2 y)}}\quad,\beta_1,\beta_2>0,\ 0<x,y<\infty$$

where $\psi(x,y)=e^{\beta_1 x}+e^{\beta_2 y}+e^{\min(\beta_1 x,\beta_2 y)}$ satisfies (6.2.2).

A bivariate semi-logistic distribution in the Marshall-Olkin form is

$\bar{G}(x,y)=\dfrac{\alpha}{\psi(x,y)+\alpha}$, denoted by MOBVSL (α,β_1,β_2,p).

In the next Section we introduce bivariate half semi-logistic distribution.

6.3 Bivariate Half Semi-Logistic Distribution

Let us introduce a bivariate half semi-logistic distribution with survival function

$$\bar{F}(x,y)=\frac{2}{1+\psi(x,y)} \tag{6.3.1}$$

where $\psi(x,y)$ satisfies the functional equation ,

$$\psi(x,y)=\frac{1}{p}\psi\left(\frac{1}{\beta_1}\ln p+x,\frac{1}{\beta_2}\ln p+y\right),\ 0<p<1,\ \beta_1,\beta_2>0,\ 0<x,y<\infty.$$

The solution of this functional equation is $\psi(x,y)=e^{\beta_1 x}h_1(x)+e^{\beta_2 y}h_2(y)$, where

$$h_1(x)=h_1\left(x+\frac{1}{\beta_1}\ln p\right)\ \text{and}\ h_2(y)=h_2\left(y+\frac{1}{\beta_2}\ln p\right),\ 0<x,y<\infty.$$

We denote this distribution as BVHSL (β_1,β_2,p).

The corresponding Marshall-Olkin form is

$$\bar{F}(x,y)=\frac{2\alpha}{\psi(x,y)+2\alpha-1},\qquad \alpha>0, \tag{6.3.2}$$

denoted by MOBVHSL ($\alpha, \beta_1, \beta_2, p$).

As an example, we have a bivariate half logistic distribution denoted by BVHL (β_1, β_2) with survival function

$$\overline{F}(x,y) = \frac{2}{1 + e^{\beta_1 x} + e^{\beta_2 y} - e^{\min(\beta_1 x, \beta_2 y)}} \quad ,0 < x, y < \infty, \beta_1, \beta_2 > 0. \qquad (6.3.3)$$

The corresponding Marshall-Olkin form is,

$$\overline{F}(x,y) = \frac{2\alpha}{e^{\beta_1 x} + e^{\beta_2 y} - e^{\min(\beta_1 x, \beta_2 y)} + 2\alpha - 1}, \alpha > 0, 0 < x, y < \infty, \beta_1, \beta_2 > 0. \quad (6.3.4)$$

When $\alpha = \frac{1}{2}$, (6.3.4) becomes a bivariate exponential distribution.

6.4 Autoregressive Bivariate Semi-Logistic Processes

Here we give first order stationary bivariate semi-logistic process. The following theorem gives a necessary and sufficient condition for a first order autoregressive minification process to have BVSL distribution as stationary distribution.

Theorem 6.4.1

Let (X_0, Y_0) follows BVSL (β_1, β_2, p). Define $\{(X_n, Y_n), n \geq 1\}$ as follows:

$$X_n = \begin{cases} X_{n-1} - \dfrac{1}{\beta_1}\ln p & w.p. \quad p \\ \min\left(X_{n-1} - \dfrac{1}{\beta_1}\ln p, \varepsilon_n \right) & w.p. \quad (1-p) \end{cases}$$

and (6.4.1)

$$Y_n = \begin{cases} Y_{n-1} - \dfrac{1}{\beta_2}\ln p & w.p. \quad p \\ \min(Y_{n-1} - \dfrac{1}{\beta_2}\ln p, \eta_n) & w.p. \quad (1-p) \end{cases}$$

Then $\{(X_n, Y_n)\}$ defines a stationary first order BVSL (β_1, β_2, p) process if and only if $\{(\varepsilon_n, \eta_n)\}$ has BVSL (β_1, β_2, p) distribution.

Proof:

Suppose (X_0, Y_0) follows BVSL (β_1, β_2, p) and (ε_n, η_n) follows BVSL (β_1, β_2, p).

Then, $\overline{F}_{X_1,Y_1}(x,y) = \overline{F}_{X_0,Y_0}(x + \dfrac{1}{\beta_1}\ln p, y + \dfrac{1}{\beta_2}\ln p)\left(p + (1-p)\overline{F}_{\varepsilon_1,\eta_1}(x,y)\right)$

$$= \frac{1}{1+\psi(x,y)}.$$

Therefore, $(X_1, Y_1) \underline{\underline{d}}$ BVSL (β_1, β_2, p).

Assuming $(X_{n-1}, Y_{n-1}) \underline{\underline{d}}$ BVSL (β_1, β_2, p) we can prove that

$(X_n, Y_n) \underline{\underline{d}}$ BVSL (β_1, β_2, p).

Then by mathematical induction, it follows that $\{(X_n, Y_n)\}$ follows BVSL (β_1, β_2, p). Converse part easily follows. □

Theorem 6.4.2

If (X_0, Y_0) has an arbitrary bivariate distribution with survival function $\overline{G}_0(x,y)$ and $\{(\varepsilon_n, \eta_n)\}$ is a sequence of i.i.d. BVSL (β_1, β_2, p) random vectors then $\{(X_n, Y_n)\}$ converges in distribution to BVSL (β_1, β_2, p).

Proof:

$$\overline{G}_n(x,y) = P(X_n > x, Y_n > y)$$

$$= P\left(X_{n-1} > x + \frac{1}{\beta_1}\ln p, Y_{n-1} > y + \frac{1}{\beta_2}\ln p\right)\left[\frac{1+p\psi(x,y)}{1+\psi(x,y)}\right]$$

$$= P\left(X_0 > x + \frac{n}{\beta_1}\ln p, Y_0 > y + \frac{n}{\beta_2}\ln p\right)\left[\frac{1+p^n\psi(x,y)}{1+\psi(x,y)}\right].$$

As $n \to \infty$, $\quad \overline{G}_n(x,y) = \dfrac{1}{1+\psi(x,y)}$.

This implies $\{(X_n, Y_n)\}$ converges in distribution to BVSL (β_1, β_2, p). □

To study the properties, let us consider (6.4.1) with bivariate logistic process as stationary distribution.

6.5 A Bivariate Logistic Process

We denote the random vector with survival function

$$\overline{F}(x,y) = \frac{1}{1+e^{\beta_1 x}+e^{\beta_2 y}}, \quad \beta_1, \beta_2 > 0, \quad 0 < x, y < \infty$$

by BVL(β_1, β_2).

Let (X_0, Y_0) follows BVL(β_1, β_2) and $\{(\varepsilon_n, \eta_n)\}$ follows i.i.d. BVL (β_1, β_2). Then from theorem (6.4.1) it is immediate that the process (6.4.1) is stationary with marginal BVL (β_1, β_2).

Now we study some properties of the BVL autoregressive process.

$P(X_{n+1} > x, Y_{n+1} > y / X_n = x_n, Y_n = y_n)$

$$= \begin{cases} \dfrac{1 + pe^{\beta_1 x} + pe^{\beta_2 y}}{1 + e^{\beta_1 x} + e^{\beta_2 y}} & if \quad x + \dfrac{1}{\beta_1} \ln p < x_n, y + \dfrac{1}{\beta_2} \ln p < y_n \\ 0 & otherwise \end{cases}$$

Consider the autocovariance matrix

$$\sum = \begin{bmatrix} \operatorname{cov}(X_n, X_{n+h}) & \operatorname{cov}(X_n, Y_{n+h}) \\ \operatorname{cov}(X_{n+h}, Y_n) & \operatorname{cov}(Y_n, Y_{n+h}) \end{bmatrix}.$$

The elements of this matrix can be calculated using the following steps.

$$P(X_{n+h} \le x / X_n = y) = \begin{cases} (1 - p^h) \dfrac{e^{\beta_1 x}}{1 + e^{\beta_1 x}} & if \quad x + \dfrac{h}{\beta_1} \ln p < y \\ 1 & if \quad x + \dfrac{h}{\beta_1} \ln p > y \end{cases}$$

$E(X_{n+h} / X_n) =$

$$\beta_1 (1 - p^h) \int_0^{y - \frac{h}{\beta_1} \ln p} \frac{x e^{\beta_1 x}}{(1 + e^{\beta_1 x})^2} dx + \left(y - \frac{h}{\beta_1} \ln p\right) \left(\frac{1 + e^{\beta_1 y}}{1 + p^{-h} e^{\beta_1 y}} \right).$$

$E(X_n X_{n+h}) = E(X_n E(X_{n+h} / X_n))$.

$$Cov(X_n, X_{n+h}) = E\left(X_n \left((1-p^h) \int_0^{p^{-h} e^{\beta_1 x}} \frac{\ln u}{(1+u)^2} du + E\left(X_n (y - \frac{h}{\beta_1} \ln p) \left(\frac{1+e^{\beta_1 y}}{1+p^{-h} e^{\beta_1 y}} \right) \right) \right) \right)$$

$$= \left(\frac{1-p^h}{\beta_1} \right) \Phi[1,0,1,-1] - \frac{1}{\beta_1^2} \Phi[1,1,0,-1]$$

where, $\Phi[p,q,r,-1] = \int_0^1 \left(\ln\left(\frac{t}{1-t} \right) \right)^p \left(\ln\left(p^{-h} \frac{t}{1-t} \right) \right)^q t^r \left(1 - t(1-p^{-h}) \right)^{-1} dt$.

Similarly we get,

$$Cov(Y_n, Y_{n+h}) = \left(\frac{1-p^h}{\beta_2} \right) \Phi[1,0,1,-1] - \frac{1}{\beta_2^2} \Phi[1,1,0,-1].$$

$$Cov(X_n, Y_{n+h}) = \frac{1}{\beta_1 \beta_2} \left(2p^h \Psi[1,1,1,3] + \Psi[1,0,1,2] - p^h \Psi[0,0,1,2] \right)$$

where, $\Psi[p,q,r,s] = \int_0^1 \int_0^1 \frac{\ln\left(\frac{t_1}{1-t_1} \right) \ln\left(\frac{t_2}{1-t_2} \right) \left(1 - t_2(1-p^h) \right)^p (1-t_1)^q (1-t_2)^r}{\left(1 - t_1(1-p^h) - t_1 t_2 p^h \right)^s} dt_1 dt_2.$

Thus we have all the elements of the covariance matrix.

Also, $P(X_n > x, X_{n+h} > y) = \begin{cases} \overline{F_X}(y) & if \quad x > y + \frac{h}{\beta_1} \ln p \\ \dfrac{\overline{F_X}(x) \overline{F_X}(y)}{\overline{F_X}(y + \frac{h}{\beta_1} \ln p)} & if \quad x < y + \frac{h}{\beta_1} \ln p \end{cases}$

We can prove the following asymptotic property of extremes ,

$$P(\min(X_1,...X_n) > x - \frac{1}{\beta_1}\ln n, \min(Y_1,...,Y_n) > y - \frac{1}{\beta_2}\ln n) \to e^{-\left(e^{\beta_1 x} + e^{\beta_2 y}\right)}$$

as $n \to \infty$.

6.6 A Generalized Bivariate First Order Autoregressive Minification Process

In this Section we define a generalized bivariate first order autoregressive process, which is a generalization of many first order bivariate autoregressive processes.

Let F and G be a non-degenerate distribution functions with $F(-\infty) = G(-\infty) = 0$ and $F(\infty) = G(\infty) = 1$.

Let $\phi_1(x) = \ln \frac{F(x)}{\bar{F}(x)}$ and $\phi_2(y) = \ln \frac{G(y)}{\bar{G}(y)}$.

Define the general bivariate first order minification autoregressive process $\{(X_n, Y_n)\}$ as follows:

$$X_n = \begin{cases} \phi_1^{-1}\left(\phi_1(X_{n-1}) - \frac{1}{\beta_1}\ln p\right) & w.p. \quad p \\ \min\left(\phi_1^{-1}\left(\phi_1(X_{n-1}) - \frac{1}{\beta_1}\ln p\right), \varepsilon_n\right) & w.p. \quad (1-p) \end{cases}$$

and

$$Y_n = \begin{cases} \phi_2^{-1}\left(\phi_2(Y_{n-1}) - \dfrac{1}{\beta_2}\ln p\right) & w.p. \quad p \\ \min\left(\phi_2^{-1}\left(\phi_2(Y_{n-1}) - \dfrac{1}{\beta_2}\ln p\right), \eta_n\right) & w.p. \quad (1-p) \end{cases}$$

where $\{(\varepsilon_n, \eta_n)\}$ is a bivariate random vector independent of $\{(X_i, Y_i)\}$ for i< n.

This bivariate process is useful in constructing stationary bivariate first order autoregressive processes with required bivariate distribution as marginal.

Example 6.6.1

If $\overline{F}(x) = \dfrac{1}{1+e^{\beta_1 x}}$ and $\overline{G}(y) = \dfrac{1}{1+e^{\beta_2 y}}$ then

$$X_n = \begin{cases} X_{n-1} - \dfrac{1}{\beta_1}\ln p & w.p. \quad p \\ \min\left(X_{n-1} - \dfrac{1}{\beta_1}\ln p,\ \varepsilon_n\right) & w.p. \quad (1-p) \end{cases}$$

and

$$Y_n = \begin{cases} Y_{n-1} - \dfrac{1}{\beta_2}\ln p & w.p. \quad p \\ \min\left(Y_{n-1} - \dfrac{1}{\beta_2}\ln p,\ \eta_n\right) & w.p. \quad (1-p) \end{cases}$$

This is the bivariate process with $\text{BVL}(\beta_1, \beta_2)$ as stationary marginal distribution if and only if $(X_0, Y_0) \stackrel{d}{=} (\varepsilon_n, \eta_n)$ and has the distribution $BVL(\beta_1, \beta_2)$.

Example 6.6.2.

If $\overline{F}(x) = \dfrac{1}{1+x^{\beta_1}}$ and $\overline{G}(y) = \dfrac{1}{1+y^{\beta_2}}$, then

$$X_n = \begin{cases} p^{-\frac{1}{\beta_1}} X_{n-1} & w.p. \quad p \\ \min\left(p^{-\frac{1}{\beta_1}} X_{n-1},\ \varepsilon_n\right) & w.p. \quad (1-p) \end{cases}$$

and

$$Y_n = \begin{cases} p^{-\frac{1}{\beta_2}} Y_{n-1} & w.p. \quad p \\ \min\left(p^{-\frac{1}{\beta_2}} Y_{n-1},\ \eta_n\right) & w.p. \quad (1-p) \end{cases}.$$

This is the bivariate Pareto process where $(X_0, Y_0) \underline{\underline{d}} (\varepsilon_1, \eta_1)$ and $\{(\varepsilon_n, \eta_n)\}$'s are i.i.d. bivariate Pareto distribution with survival function,

$$\overline{F}(x, y) = \frac{1}{1+x^{\beta_1}+y^{\beta_2}}, \ 0 < x, y < \infty, \quad \beta_1, \beta_2 > 0.$$

Example 6.6.3.

If $\overline{F}(x) = 1 - x$ and $\overline{G}(y) = 1 - y$ then

$$X_n = \begin{cases} \dfrac{X_{n-1}}{p+(1-p)X_{n-1}} & \text{w.p.} \quad \text{p} \\ \min\left(\dfrac{X_{n-1}}{p+(1-p)X_{n-1}},\ \varepsilon_n\right) & \text{w.p.} \quad (1-\text{p}) \end{cases}$$

and

$$Y_n = \begin{cases} \dfrac{Y_{n-1}}{p+(1-p)Y_{n-1}} & \text{w.p. } p \\ \min\left(\dfrac{Y_{n-1}}{p+(1-p)Y_{n-1}}, \eta_n\right) & \text{w.p. } (1-p) \end{cases}$$

defines bivariate uniform process when $(X_0, Y_0) \underline{\underline{d}} (\varepsilon_1, \eta_1)$ and $\{(\varepsilon_n, \eta_n)\}$'s are i.i.d. uniform on the region $\left(0, \dfrac{1}{1-\sqrt{2p}}\right) \times \left(0, \dfrac{1}{1-\sqrt{2p}}\right)$ if $p < 1/2$ and on the region $\left(0, \dfrac{1}{1+\sqrt{2p}}\right) \times \left(0, \dfrac{1}{1+\sqrt{2p}}\right)$ if $p > 1/2$, with survival function $\overline{F}_{\varepsilon,\eta}(x,y) = \dfrac{1-x-y+xy-2pxy}{1+xy}$.

6.7 Bivariate Half Semi-Logistic Process

In this Section we define first order autoregressive minification models with bivariate half semi-logistic and bivariate half logistic random vectors.

Define $\{(X_n, Y_n)\}$ as follows:

Let $X_n = \begin{cases} \varepsilon_n & w.p. \quad \alpha \\ \min(X_{n-1}, \varepsilon_n) & w.p. \quad 1-\alpha \end{cases}$

and (6.7.1)

$$Y_n = \begin{cases} \eta_n & w.p. \quad \alpha \\ \min(Y_{n-1}, \eta_n) & w.p. \quad 1-\alpha \end{cases}$$

where $\{(\varepsilon_n, \eta_n)\}$'s are i.i.d. random vectors independent of $\{(X_i, Y_i)\}, i < n$..

Theorem 6.7.1

Assuming that $(X_0, Y_0) \sim$ BVHSL (β_1, β_2, p), the process $\{(X_n, Y_n)\}$ in (6.7.1) is stationary if and only if $\{(\varepsilon_n, \eta_n)\}$ follows MOBVHSL $(\alpha, \beta_1, \beta_2, p)$ with survival function (6.3.2).

Proof follows easily. □

Some of the properties of this process are given below.

$$P(X_{n+1} > x, Y_{n+1} > y / X_n = x_n, Y_n = y_n) = \begin{cases} \dfrac{2\alpha}{\psi(x,y) + 2\alpha - 1} & \text{if} \quad x_n > x, y_n > y \\ \dfrac{2\alpha^2}{\psi(x,y) + 2\alpha - 1} & \text{if} \quad x_n < x, y_n < y \end{cases}$$

$$P\left((X_n, Y_n) > (x,y), (X_{n+1}, Y_{n+1}) > (x', y')\right) = \begin{cases} \overline{F}_{X,Y}(x,y)\overline{F}_{\varepsilon,\eta}(x', y') \\ \qquad \text{if } x > x', y > y' \\ \overline{F}_{\varepsilon,\eta}(x', y')\left(\alpha \overline{F}_{X,Y}(x,y) + (1-\alpha)\overline{F}_{X,Y}(x', y')\right) \\ \qquad \text{if } x < x', y < y' \end{cases}$$

If $\psi(x,y) = e^{\beta_1 x} + e^{\beta_2 y} - e^{\min(\beta_1 x, \beta_2 y)}$ then we have $\overline{F}(x,y)$ as in (6.3.3). Thus we get a bivariate half logistic process.

Remark 6.7.1: Multivariate extensions of the models studied in the foregoing sections can be done on similar lines. In the multivariate extension many of the desirable properties of the bivariate model can be seen to be true. Thus it turns out to be a useful model for non-negative multivariate data as well.

CHAPTER VII

SUMMARY AND CONCLUDING REMARKS

CHAPTER VII

SUMMARY AND CONCLUDING REMARKS

7.1 Summary of the Thesis

Minification models have gained wide spread applications in many fields. Its initial study was motivated by hydrological considerations. Another case of interest concerns time series of wind velocity magnitudes. Also in reliability studies, sequences of times between failures that are correlated can be modeled with minification processes. Rapid development in this area has been amply supported by very interesting and fruitful applications in several real life situations. This Thesis mainly deals with some minification models that generate some useful probability distributions. It is organized into seven chapters as follows.

Chapter I give a brief account of the developments in the subject matter together with its historical perceptive.

In Chapter II, a new class of autoregressive processes with minification structure is constructed, which can generate any autoregressive minification model with a given distribution as marginal. A necessary and sufficient condition for the process to be stationary is given and properties are studied. Several examples are given. The new process so defined is extended to higher orders. The corresponding moving average and autoregressive moving average processes are also given. We develop another class of autoregressive process, called autoregressive Lehmann process, which can generate any distribution from the Lehmann family. Properties of the same are discussed. Estimation of the parameters is also done.

In Chapter III, we introduce generalized half semi-logistic distribution generated through the Marshall-Olkin form. The autoregressive minification process with generalized half semi-logistic distribution as marginals are constructed and properties are given. It is also shown that Pareto, Weibull, exponential and folded logistic belong to this class of distributions.

Chapter IV deals with a generalized autoregressive minification process. This newly introduced process is a generalization of the processes we discussed in the previous Chapters. A necessary and sufficient condition for this new process to be stationary is given. Properties are studied and examples are given. Higher order processes are constructed. A generalized half semi-logistic process and generalized semi-logistic process are discussed as particular cases. Estimation of the parameters is also done.

Random coefficient first order autoregressive minification models is constructed in Chapter V. Again a generalized random coefficient first order process is introduced and properties are studied. Also random coefficient autoregressive Lehman process is constructed and studied.

In Chapter VI, bivariate semi-logistic and half semi-logistic distributions are introduced and studied. The first order minification processes with these distributions as marginals are also discussed. Further a general process useful in generating any bivariate autoregressive minification processes with a given distribution as marginal is introduced and studied.

Chapter VII gives an over all summary of the Thesis and concluding remarks. A fairly exhaustive list of references on the topic of interest is given at the end of the Thesis.

7.2 Concluding Remarks

For modeling of time series data, it has been customary to use autoregressive models of appropriate orders. The classical analysis of time series rests heavily on Gaussian assumption, but there are many situations where the data shows a tendency to follow asymmetric and heavy tailed distributions, which cannot be modeled by Gaussian distributions. Hence in recent years there are many models introduced for explaining time series data using non-Gaussian distributions. Further, most of the models employed in analysing Gaussian time series are linear in nature. However in recent years several non-linear models with non-Gaussian marginal distributions are found to be more suitable than linear Gaussian models in certain situations. One of the important non- linear models used to generate a sequence of random variables is the minification model and these models possess most of the properties of the additive autoregressive models. The existence of such models and their properties can be easily studied using the survival function of the underlying random variable.

Several autoregressive minification models yielding useful stationary marginal distributions have been studied in this Thesis. There are many important distributions like Weibull and extreme value, which are commonly used for modeling run off series, which cannot be generated, with the usual additive autoregressive models. So the minification models suggested in this Thesis can be used as an alternative to generate those marginal distributions.

Specifically, in this Thesis, a new class of first order autoregressive minification process is constructed using the log odds function, which generalizes autoregressive minification processes of first order. Extension to bivariate case is also carried out. Also another minification structure, using the log odds function, generating any distribution from the Lehmann family is developed.

The class of generalized half semi-logistic distribution generated through the Marshall- Olkin form is introduced in this work. The well-known distributions such as exponential, Weibull, Pareto, half logistic etc. belong to this class. The autoregressive minification processes with this new class of distributions as marginals are constructed. Generalization of the autoregressive minification process with two parameters is also studied and its particular cases like half semi-logistic processes and generalized semi-logistic processes are investigated.

Random coefficient autoregressive minification models are constructed using the log odds function and their properties are studied. Special emphasis is given for the autoregressive Lehmann process under the same set up.

Finally, in order to model data that are bivariate in nature, a bivariate logistic minification process has been introduced. To model data having periodic fluctuations, bivariate semi-logistic and half semi-logistic processes are introduced and studied.

REFERENCES

REFERENCES

1. Ali, M.M., Mikhail, N.N. and Haq, M.S. (1978). A class of bivariate distributions including the bivariate logistic. *J. Multi. Analysis* **8**, 405-412.
2. Alice, T. and Jose, K.K. (2004). Bivariate semi-Pareto minification processes. *Metrika* **59**, 305-313.
3. Arnold, B.C. (1989). A logistic process constructed using geometric minimization. *Statist. Probab. Lett.* **7**, 253-257.
4. Arnold, B.C. (1993). Logistic process involving Markovian minimization. *Commun. Statist.- Theor. Meth.* **22**, 1699-1707.
5. Arnold, B.C. (2001). Pareto processes. *Handbook of Statistics* **19**, Eds. C.R. Rao and D.N. Shanbhag, Elsevier B.V., Amsterdam, 1-33.
6. Arnold B.C. and Hallett T.J. (1989). A characterization of the Pareto process among stationary stochastic processes of the form $X_n = c\min(X_{n-1}, Y_n)$. *Statist. Probab. Lett.* **8**, 377-380.
7. Arnold, B.C. and Robertson, C.A. (1989). Autoregressive logistic process. *J. Appl. Prob.* **26**, 524-531.
8. Balakrishnan, N. (1992). *Handbook of the logistic distribution.* Statistics: Textbooks and Monographs **123**. Marcel Dekker Inc., New York.
9. Balakrishna, N. (1998). Estimation for the semi-Pareto process. *Commun. Statist. - Theor. Meth.* **27**, 2307-2323.
10. Balakrishna, N. and Jacob, T.M. (2003). Parameter estimation in minification processes. *Commun. Statist. - Theor. Meth.* **32**, 2139-2152.
11. Balakrishna, N. and Jayakumar, K. (1996). Bivariate autoregressive minification process. *Journal of Applied Statistical Sciences* **5**, 129-141.

12. Balakrishna, N. and Jayakumar, K. (1997). Bivariate semi-Pareto distributions and processes. *Statistical Papers* **38**. 149-165.

13. Basu, A.K. and Das, J.K. (1993). On random coefficient autoregressive models with infinite variance. *J. Indian Stat. Assoc.* **31**, 5-19.

14. Billingsley, P. (1986). *Probability and measure* (2nd Ed.), John Wiley & Sons, New York.

15. Block, H.W., Lanberg, N.A. and Stoffer, D.S. (1988). Bivariate exponential and geometric autoregressive and autoregressive moving average models. *Adv. Appl. Prob.* **20**, 798-821.

16. Cooray, K., Gunasekhara, A. and Ananda, M. M. A. (2006). The folded logistic distributions. *Commun. Statist. - Theor. Meth.* **35,** 385-393.

17. Davis, R.A. and Resnick, S.I. (1989). Basic properties and prediction of max-ARMA processes. *Adv. Appl. Prob.* **21**, 781-803.

18. Dewald, L.S. and Lewis, P.A.W. (1985). A new Laplace second-order autoregressive time series model- NLAR(2). *IEEE Trans. Inf. Theor.* **31**, 645-651.

19. Dewald, L.S., Lewis, P.A.W. and McKenzie, E.D. (1989). A bivariate first order autoregressive time series model in exponential variables (BEAR(1)). *Management Sci.* **35**, 1236-1246.

20. Dziubdziela, W. (1997). A note on the characterization of some minification processes. *Applicationes Mathematicae 24.* 425-428.

21. Gaver, D.P. and Lewis, P.A.W. (1980). First order autoregressive gamma sequences and point processes. *Adv. Appl. Prob.* **12**, 727-745.

22. Grunwald, G.K., Hyndman, R.J., Tedesco, L. and Tweedie, R.L. (2000). Non-Gaussian conditional linear AR(1) models. *Aust. N. Z. J. Stat.* **42**, 479-495.

23. Gumbel, E.J. (1961). Bivariate logistic distributions. *J. Amer. Statist. Assoc.* **56,** 335-349.

24. Jacobs, P.A. and Lewis, P.A.W. (1977). A mixed autoregressive moving average exponential sequence and point process (EARMA 1,1). *Adv. Appl. Prob.* **9**, 87-104.

25. Jayakumar, K. and Pillai, R.N. (1993). The first order autoregressive Mittag-Leffler process. *J. Appl. Prob.* **30**, 462-466.

26. Jayakumar, K. and Pillai, R.N. (2002). A class of stationary Markov processes. *Appl. Math. Lett.* **15**, 513-519.

27. Jayakumar, K. and Mathew, T. (2002). Burr processes. *Far East J. Theo. Stat.* **8**, 99-113.

28. Jayakumar, K. and Mathew, T. (2004). Semi-Logistic distributions and processes. *Stoch. Model. and Appl.* **7**, 20-32.

29. Johnson, N.L., Kotz, S. and Balakrishnan, N. (2004). *Continuous Univariate Distributions I, II.* John & Sons (Asia) Pte. Ltd., Singapore.

30. Kalamkar, V.A. (1995). Minification processes with discrete marginals. *J. Appl. Prob.* **32**, 692-706.

31. Karlsen, H. and Tjøstheim, D. (1988). Consistent estimates for the NEAR(2) and NLAR(2) time series models. *J.R. Statist. Soc.B* **48**, 313-320.

32. Krishnarani, S.D. and Jayakumar, K. (2007a). A class of autoregressive minification processes. Accepted for publication in *Statist. Probab. Lett.*.

33. Krishnarani, S.D. and Jayakumar, K. (2007b). Generalized folded logistic distributions and processes. Paper presented in the IISA Joint Statistical Meeting and International Conference on Statistics, Probability and Related Areas, Jan. 2-5, 2007, Cochin, India.

34. Krishnarani, S.D. and Jayakumar, K. (2007c). On bivariate logistic distributions and processes. Accepted for publication in *Far East J. Theo. Stat.*.

35. Kuttikrishnan, A.P. and Jayakumar, K. (2007). Bivariate semi α-Laplace distribution and processes. *Statistical Papers* **49,** 303-313.

36. Lawrance, A.J. and Lewis, P.A.W. (1977). An exponential moving average sequence and point process (EMA1). *J. Appl. Prob.* **14**, 98-113.

37. Lawrance, A.J. and Lewis, P.A.W. (1980). The exponential autoregressive-moving average EARMA(p,q) process. *J.R. Statist. Soc.B.* **42**, 150-161.

38. Lawrance, A.J. and Lewis, P.A.W. (1981). A new autoregressive time series model in exponential variables (NEAR(1)). *Adv. Appl. Prob.* **13**, 826-845.

39. 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**, 165-202.

40. Lehmann, E.L. (1953). The power of rank tests. *Ann. Math. Statist.* **24**, 23-43.

41. Lewis, P.A.W. and McKenzie, E.D. (1991). Minification processes and their transformations. *J. Appl. Prob.* **28**, 45 – 57.

42. Littlejohn, R.P. (1992). Discrete minification processes and reversibility. *J. Appl. Prob.* **29**, 82-91.

43. Liu, J (1990). Estimation of some bilinear time series in case of infinite variance. *Commun. Statist. - Stoch. Models* **16**, 649-665.

44. Marshall, A.W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull processes. *Biometrika* **84**, 641-652.

45. McKenzie, E. (1986). Autoregressive moving average process with negative binomial and geometric marginal distributions. *Adv. Appl. Prob.* **18**, 679-705.

46. McKenzie, E. (2003). Discrete variate time series. *Handbook of Statistics* **21**, Eds. C.R.Rao and D.N.Shanbhag, Elsevier B.V., Amsterdam, 573-606.

47. Nicholls, D.F. and Quinn, B.C. (1982). *Random coefficient autoregressive models- An introduction.* Lecture Notes in Statistics **11**, Springer Verlag, New York.

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

49. Pillai, R.N., Jose, K.K. and Jayakumar, K. (1995). Autoregressive minification processes and the class of distributions of universal geometric minima. *J. Indian Stat. Assoc.* **33**, 53-61.

50. Quoreshi, S. (2006). Bivariate time series modeling of financial count data. *Commun. Statist. –Theor. Meth.* **35**, 1343-1358.

51. Ristic, M.M. (2006). Stationary bivariate minification processes. *Statist. Probab. Lett.* **76,** 439–445.

52. Ristic, M.M. (2007). A generalized Semi-Pareto minification process. *Statistical Papers* **49**, 343-351.

53. Ristic, M.M. and Popovic, B.C. (2000a). Parameter estimation for uniform autoregressive processes, *Novi Sad J. Math.* **30**, 89-95.

54. Ristic, M.M. and Popovic, B.C. (2000b). A new uniform AR(1) time series model. *Publications De L'Institut Mathematique,Nouvelle serie,* **68**, 145-152.

55. Ristic, M.M. and Popovic, B.C. (2003). A bivariate uniform autoregressive process. *Ann. Inst. Statist. Math.* **55**, 797-802.

56. Satterthwaite, S.P. and Hutchinson, T.P. (1978). A generalization of Gumbel's bivariate logistic distribution. *Metrika* **25**,163-170.

57. Sim, C.H. (1986). Simulation of Weibull and gamma autoregressive stationary processes. *Commun. Statist. - Simul. Comput.* **15**,1141-1146.

58. Sim, C.H. (1987). A mixed gamma ARMA(1,1) model for river flow time series. *Water Resources Research* **23**, 32-36.

59. Sim, C.H. (1990). First order autoregressive models for gamma and exponential processes. *J. Appl. Prob.* **27**, 325-332.

60. Sim, C.H. (1993). First order autoregressive logistic process. *J. Appl. Prob.* **30**, 467-470.

61. Smith, R.L. (1986). Maximum likelihood estimation for the NEAR(2) model. *J.R. Statist. Soc. A* **48** , 251-257.

62. Stout, W.F (1974). *Almost sure convergence.* Academic Press, New York.

63. Tavares, L.V. (1977). The exact distribution of extremes of a non-Gaussian process. *Stoch. Proc. Appl.* **5**, 151-156.

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

65. Tong (1990). *Nonlinear time series: A dynamical system approach.* Oxford University Press, Oxford.

66. Wang, Y., Hossain, A.M. and Zimmer, W.J. (2003). Monotone log odds rate distributions in reliability analysis. *Commun. Statist. – Theor. Meth.* **32,** 2227-2244.

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

****