CHAPTER 6

TIME SERIES AND STOCHASTIC PROCESSES

[*This chapter is based on the lectures of K.K. Jose, Department of Statistics, St. Thomas College, Palai, Kerala, India at the 5th SERC School*]

6.0. Introduction

A time series is a record of values of any fluctuating quantity measured at different points of time. To describe the statistical nature of an observed time series it is necessary to regard it as a member of an abstract ensemble of functions called stochastic processes. A stochastic process is a statistical phenomenon that evolves in time according to probabilistic laws. More precisely, a stochastic process $\{X(t); t \in T\}$ is a rule for assigning a function $X(t, e)$ which, for every fixed t belonging to T, is a measurable function of e $\in \Omega$ with respect to a probability space (Ω, F, P). In analyzing a time series we regard it as a realization of a stochastic process. The process is called discrete parameter or continuous parameter according as t takes discrete or continuous set of values. There are two approaches to time series analysis, namely frequency domain analysis and time domain analysis. Complimentary to these techniques, very recently wavelet analysis is also being used as a tool for time series analysis.

6.1. Stationary Processes

One of the characteristic features that distinguishes time series data from other types of statistical data is the fact that, the values of the series at different time instants will be correlated. Hence a basic problem in time series analysis is to study the pattern of the correlation between values at different time instants and try to construct statistical models which explain the correlation structure of the series. The auto covariance function (abbreviated as $ACVF$) of a stochastic process $\{X(t)\}$ is denoted by $R(s, t)$ and is defined as

$$R(s,t) = \text{Cov}[X(s), X(t)]. \qquad (6.1.1)$$

6.1.1. Stationary time series

Stationarity is a basic assumption in classical time series analysis. It means, in effect, that the main statistical properties of the series remain unchanged over time. More precisely, a process $\{X(t)\}$ is said to be completely stationary or strict sense stationary (abbreviated as SSS) if the process $X(t)$ and $X(t+c)$ have the same statistics for any c. That is for any set of time points $t_1, t_2, \cdots, t_n$ and any integer c, the joint probability distribution of $[X(t_1), X(t_2), \cdots, X(t_n)]$ is identical with that of $[X(t_1+c), X(t_2+c), \cdots, X(t_n+c)]$. Less stringently, a process $\{X(t)\}$ is said to be covariance stationary (second order stationary) or wide sense stationary (abbreviated as WSS) if the mean and variance of $X(t)$ remain constant over time and the autocovariance between any two values depends only on the time difference and not on their individual locations. That is,

$$\begin{aligned}
&\text{(i)} \quad E[X(t)] = \mu, \quad \text{independent of } t \\
&\text{(ii)} \quad \text{Var}[X(t)] = \sigma_x^2, \quad \text{independent of } t \\
&\text{(iii)} \quad \text{Cov}[X(t), X(t+s)] = R(s).
\end{aligned} \qquad (6.1.2)$$

It may be noted that SSS implies WSS, but the converse is not true always.

6.1.2. Properties of the autocovariance function

An immediate consequence of stationarity is that the $ACVF$ can be written as

$$R(s) = \text{Cov}[X(t), X(t+s)].$$

Obviously $R(o) = \sigma_x^2$. Then the autocorrelation function of a WSS process can be defined as

$$\rho(s) = \frac{R(s)}{R(0)}. \qquad (6.1.3)$$

The basic properties of the $ACVF$ are

(i) $R(0) = \sigma_x^2$
(ii) $R(s) = R(-s)$
(iii) $|R(s)| \leq R(0)$ (6.1.4)
(iv) The autocovariance matrix is positive semi definite
(v) If $\{X(t)\}$ is continuous, then $R(s)$ is a continuous function in s

Here property (iv) leads to the concept of power spectrum, making use of Bochner's theorem discussed in 6.1.7.

6.1.3. Ergodicity

Ergodicity plays the central role in the Wiener-Khintchine theory of WSS processes. A stochastic process $X(t)$ is ergodic if its ensemble averages equal appropriate time averages. That is with probability 1, any statistic of $X(t)$ can be determined from a single sample $X(t; \mathrm{e})$. Ergodic property allows the substitution of time averages for ensemble averages. (Papoulis, 1985).

Example 6.1.1. Consider a stochastic process $\{x(t)\}$ with constant mean $E[x(t)] = \mu$. We form the time average

$$\mu_T = \frac{1}{2T}\int_{-T}^{T} X(t)\mathrm{d}t.$$

Clearly $E(\mu_T) = \frac{1}{2T}\int_{-T}^{T} E[X(t)]\mathrm{d}t = \mu$. Now we say the process $\{X(t)\}$ is mean-ergodic if, with probability 1, $\mu_T \to \mu$ as $T \to \infty$.

This is true iff $\sigma_T^2 = Var[X(t)] \to 0$ as $T \to \infty$.

6.1.4. Gaussian processes

Most results in classical time series analysis are based on the assumption that the underlying process is Gaussian. A stochastic process $\{X(t)\}$ is called a Gaussian process if for all $t_1, t_2, \cdots, t_n$ the set of random variables $[X(t_1), X(t_2), \cdots, X(t_n)]$ has a multivariate normal distribution. Gaussian processes are characterized by the mean and covariance function. Hence for such processes WSS implies SSS and vice versa.

6.1.5. White noise process

A process $\{\varepsilon(t)\}$ is called a white noise if its values $\varepsilon(t_i)$ and $\varepsilon(t_j)$ are uncorrelated for every t_i and $t_j \neq t_i$. If the random variables $\varepsilon(t_i)$ and $\varepsilon(t_j)$ are not only uncorrelated but also independent, then $\{\varepsilon(t)\}$ is called a strictly white noise process. For a white noise process $\varepsilon(t)$ with variance σ_ε^2 the $ACVF$ is

$$R(s) = \begin{cases} \sigma_\varepsilon^2, \, s = 0 \\ 0, \text{otherwise.} \end{cases} \tag{6.1.5}$$

It will be assumed that the mean of a white noise process is identically zero, unless otherwise stated. White noise processes are important in time series analysis due to their simple structure.

6.1.6. Harmonic analysis

One of the basic problems in time series analysis is to represent a time series as a sum of harmonics of sines and cosines. For that, we consider a process of the form

$$X(t) = \sum_{j=1}^{n} (A_j \cos \lambda_j t + B_j \sin \lambda_j t) \quad t = 0, \pm 1, \cdots \tag{6.1.6}$$

where λ_j's are constants and A_j's and B_j's are random variables such that

$$\begin{aligned} E(A_j) &= E(B_j) = 0 \\ E(A_j^2) &= E(B_j^2) = \sigma_j^2, \quad j = 1, 2, \cdots, n. \\ E(A_i A_j) &= E(B_i B_j) = 0, \quad i \neq j, \quad i, j = 1, 2, \cdots, n. \\ E(A_i B_j) &= 0, \quad i, j = 1, 2, \cdots, n. \end{aligned} \tag{6.1.7}$$

Then the process $X(t)$ is wide sense stationary, since $E(X(t)) = 0$ and

$$R(s) = E(X(t)X(t+s)) = \sum_{j=1}^{n} \sigma_j^2 \cos \lambda_j s \tag{6.1.8}$$

If we assume that the A_j's and B_j's are normally distributed, then $X(t)$ is a Gaussian process and the process will be stationary in the strict sense also.

It may be noted that for any stationary stochastic process we can construct a process of the type given by (6.1.7). Now from (6.1.8) it follows that

$$R(0) = \sum_{j=1}^{n} \sigma_j^2 \tag{6.1.9}$$

This representation of the variance has a number of features that promote a clear understanding of the nature of the mean square variation in the process. The additivity of the variance components and the absence of covariance terms permit us to determine the relative importance of each periodic component. The orthogonality of the components allows us to consider the mean-square variation as being generated by independent sources, under the assumption of normality.

The approximation of $X(t)$ in (6.1.7) can be improved by taking n large enough. It may be noted that this representation is only a particular case of the fundamental result in Fourier analysis that any analytic function may be approximated to any given degree of accuracy by a linear combination of sines and cosines. The process $X(t)$ given by (6.1.7) can also be represented in the from

$$X(t) = \sum_{j=1}^{n} R_j \cos(\lambda_j t - \theta_j) \quad t = 0, \pm 1, \cdots$$

where

$$R_j = A_j^2 + B_j^2 \geq 0, \quad \theta_j = \tan^{-1}(B_j/A_j)$$
$$0 < \theta_j < \pi \;\text{ if } B_j > 0, \quad \pi < \theta_j < 2\pi \;\text{ if } B_j < 0.$$

Further if A_j and B_j are normally distributed, then R_j^2/σ_j^2 is a chi-square variable with 2 degrees of freedom.

6.1.7. Bochner's theorem

A function $r(t)$ is non-negative definite if and only if it can be represented in the form

$$r(t) = \int_{-\infty}^{\infty} e^{itw} dF(w) \tag{6.1.10}$$

where $F(w)$ is real, never-decreasing and bounded.

This theorem is extremely important in the spectral analysis of time series, since it guarantees the existence of the spectral distribution function associated with any covariance sequence.

Exercises 6.1.

6.1.1. Give an example of a strictly stationary process.

6.1.2. Give an example of a covariance stationary process.

6.1.3. If $\{X_n; n \geq 1\}$ is a set of uncorrelated random variables with mean 0 and variance 1, show that $\{X_n\}$ is a strictly stationary process.

6.1.4. Let $X(t) = A_1 + A_2 t$ where A_1, A_2 are independent random variables with $E(A_i) = a_i$, $\text{Var}(A_i) = \sigma_i^2$, $i = 1, 2$. Show that $\{X(t)\}$ is not stationary.

6.1.5. Let $X(t) = A \cos wt + B \cos wt$ where A and B are uncorrelated random variables with means 0 and variances 1 and w is a positive constant. Examine whether $\{X(t)\}$ is covariance stationary.

6.1.6. Consider a Poisson process $\{N(t)\}$ where $P(N(t) = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$; $n = 0, 1, 2, \cdots$. Show that $\{N(t)\}$ is not stationary.

6.1.7. Let $X(t) = X_0(-1)^{N(t)}$ where X_0 is a random variable such that $P(X_0 = 1) = P(X_0 = -1) = \frac{1}{2}$ and $\{N(t)\}$ is a Poisson process as given in Exercise 6.1.6. Examine whether $\{X(t)\}$ is covariance stationary.

6.1.8. Consider an autoregressive process $\{X_n\}$ where $X_n = \rho X_{n-1} + \epsilon_n; n \geq 1, |\rho| < 1$ and $X_0 = \epsilon_0$ where $\epsilon_0, \epsilon_1, \epsilon_2, \cdots$ are uncorrelated random variable with $E(\epsilon_n) = 0$ for $n \geq 0$. Show that $\{X_n\}$ is covariance stationary.

6.2. The Spectral Distribution

The spectral distribution function (integrated spectrum) and the spectral density function (spectrum) are the two principal functions of interest in the frequency domain analysis of time series. Of these, the spectral distribution function exists always by virtue of the Bochner's Theorem. Consider a WSS stochastic process

$X(t)$ with auto covariance function $R(s)$, which is assumed to be continuous. Then by Bochner's Theorem, there exists a bounded non-decreasing function of a real variable w, denoted by $F(w)$ defined for $-\infty < w < \infty$ such that

$$R(s) = \int_{-\infty}^{\infty} \mathrm{e}^{iws} \mathrm{d}F(w), -\infty < s < \infty \tag{6.2.1}$$

This result was established by Khintchine in 1933. Here $F(w)$ satisfies $F(-\infty) = 0$, $F(+\infty) = R(0) = \sigma_x^2$. Similarly, Wold established that in the discrete parameter case there exists a non-decreasing bounded function $F(w)$ defined for $-\pi \leq w \leq \pi$ such that

$$R(s) = \int_{-\pi}^{\pi} \mathrm{e}^{iws} \mathrm{d}F(w), \quad s = 0, \pm 1, \cdots, F(-\pi) = 0, F(\pi) = \sigma_x^2. \tag{6.2.2}$$

The function $F(w)$ is called the spectral distribution function or integrated spectrum of the process.

6.2.1. Decomposition of the spectral distribution function

The Spectral distribution function $F(w)$ can be uniquely written as the sum

$$F(w) = F_1(w) + F_2(w) + F_3(w) \tag{6.2.3}$$

where $F_1(w), F_2(w)$ and $F_3(w)$ are each nondecreasing and

(i) $F_1(w)$ is a pure step function

(ii) $F_2(w)$ is an absolutely continuous function which can be written as the integral of a non-negative function $f(w)$ and

(iii) $F_3(w)$ is a singular continuous function which may be increasing, although $F_3'(w) = 0$ almost everywhere. (Parzen, 1961).

6.2.2. The spectral density function

The Spectral density function or the spectrum denoted by $f(w)$ is defined as the derivative of the absolutely continuous part of $F(w)$. Then from (6.2.1) it follows that

$$R(s) = \int_{-\infty}^{\infty} \mathrm{e}^{iws} f(w) \mathrm{d}w. \tag{6.2.4}$$

This shows that the autocovariance function $R(s)$ is the Fourier transform of the spectral density function $f(w)$. Then by inverse Fourier transform

$$f(w) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iws}R(s)ds. \tag{6.2.5}$$

Obviously it follows that the integrated spectrum can also be defined as

$$F(w) = \int_{-\infty}^{w} f(\theta)d\theta. \tag{6.2.6}$$

In the discrete parameter case, the spectral density function is defined as the discrete Fourier transform of the autocovariance function $R(s)$. That is,

$$f(w) = \frac{1}{2\pi}\sum_{s=-\infty}^{\infty} R(s)e^{-iws}, \quad -\pi \le w \le \pi \tag{6.2.7}$$

provided

$$\sum_{s=-\infty}^{\infty} |R(s)| < \infty.$$

Then

$$R(s) = \int_{-\pi}^{\pi} e^{iws}F(w)dw$$

and

$$F(w) = \int_{-\pi}^{w} f(\theta)d\theta. \tag{6.2.8}$$

Anderson (1971) considers real discrete parameter stationary processes in which case we have

$$f(w) = \frac{1}{2\pi}\sum_{s=-\infty}^{\infty} R(s)\cos ws \tag{6.2.9}$$

$$R(s) = \int_{-\pi}^{\pi} \cos ws\ f(w)dw = \int_{-\pi}^{\pi} \cos ws\ dF(w)$$

and

$$F(w) = \frac{R(0)}{2\pi}(w+\pi) + \frac{1}{\pi}\sum_{s=1}^{\infty}\left(\frac{R(s)\sin ws}{s}\right). \tag{6.2.10}$$

The spectral density function is an even, non-negative, integrable function. Now we consider discrete parameter, real valued times series only. Then from (6.2.9) we get

$$R(0) = \sigma_x^2 = \int_{-\pi}^{\pi} f(w)\mathrm{d}w. \tag{6.2.11}$$

In engineering contexts, σ_x^2 represents the total power contained in the signal $X(t)$ and the integral on the right-hand side of (6.2.11) represents a frequency decomposition of the total power. In other words, $f(w)\mathrm{d}w$ represents the power contained in the frequency band $(w, w+\mathrm{d}w)$ of the process $X(t)$. The spectral density function $f(w)$ can be normalized by dividing by $R(0)$. Hence the function

$$g(w) = \frac{f(w)}{R(0)} \tag{6.2.12}$$

is similar to the $p.d.f.$ in probability theory. The function $g(w)$ is obtained by taking the Fourier transform of the autocorrelation function $\rho(s)$. Hence for continuous parameter process,

$$g(w) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-iws}\rho(s)\mathrm{d}s. \tag{6.2.13}$$

In the discrete case,

$$g(w) = \frac{1}{2\pi}\sum_{s=-\infty}^{\infty} \mathrm{e}^{-iws}\rho(s). \tag{6.2.14}$$

The function $g(w)$ is called the normalized spectral density function.

Example 6.2.1. Consider a white noise process $\{\varepsilon(t)\}$ with zero mean and variance σ_ε^2, then the $ACVF$ is

$$R(s) = \begin{cases} \sigma_\varepsilon^2 & ; s = 0 \\ 0 & ; s \neq 0 \end{cases} \tag{6.2.15}$$

and therefore $f(w) = \frac{\sigma_\varepsilon^2}{2\pi}$. Hence a white noise process has a constant spectrum.

Example 6.2.2. If $X(t)$ is a binary transmission process, then its *ACVF* is

$$R(s) = 1 - \frac{|s|}{T}, \quad |s| \le T, \tag{6.2.16}$$

and therefore

$$f(w) = \frac{2\sin^2(wT/2)}{\pi T w^2}. \tag{6.2.17}$$

$R(s)$ and $f(w)$ are given in Figure (6.2.1).

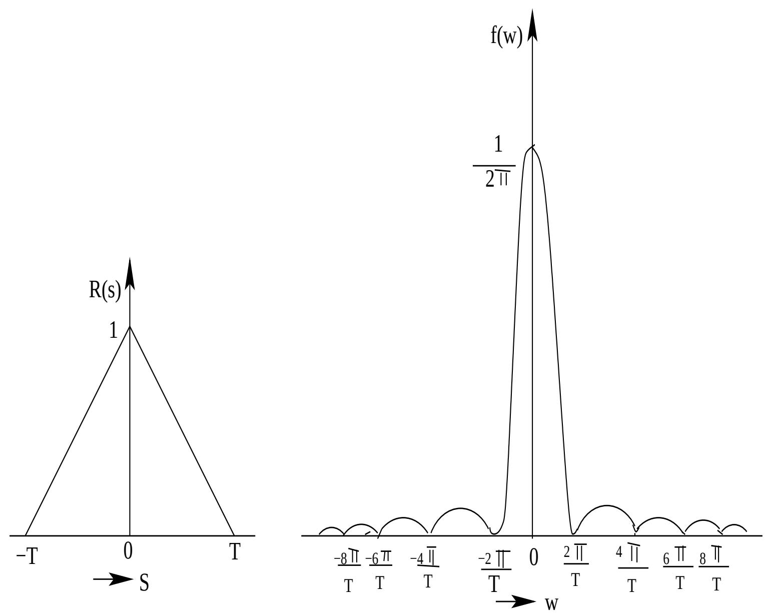

Figure 6.2.1

Example 6.2.3. If $X(t)$ is a random telegraph signal with *ACVF*

$$R(s) = e^{-a|s|}, -\infty < s < \infty.$$

Then

$$f(w) = \frac{a}{\pi(a^2 + w^2)}, -\infty < w < \infty.$$

Here $R(s)$ has the shape of a double sided negative exponential (Laplace) curve while $f(w)$ has the shape of a Cauchy distribution curve, as shown in Figure (6.2.2).

These illustrative examples point out the fact that by observing the shape of the *ACVF* we can infer the form of the spectral density function and vice versa. This analysis helps to infer the nature of the original process (Papoulis, 1985).

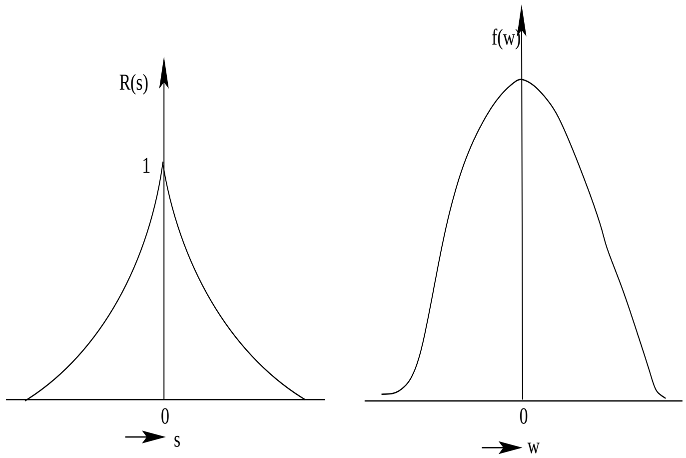

Figure 6.2.2

6.2.3. Spectral representation of WSS process

The spectral representation of a stochastic process enables us to give an interpretation of the spectral distribution function in terms of the variances of random amplitudes of trigonometric functions composing the process. The spectral representation is given in terms of stochastic integrals, which correspond to the integrals of spectral distribution functions in the spectral representation of the covariance sequence given by (6.2.2). Corresponding to (6.2.2) there exists a spectral representation for any zero mean discrete parameter process $X(t)$, in the form

$$X(t) = \int_{-\pi}^{\pi} \mathrm{e}^{iwt} \mathrm{d}Z(w) \quad t = 0, \pm 1, \pm 2, \cdots \tag{6.2.18}$$

where $Z(w)$ is a process with orthogonal increments, and related to $F(w)$ by

$$E(|\mathrm{d}Z(w)|^2) = \mathrm{d}F(w). \tag{6.2.19}$$

Expression (6.2.18) shows that any discrete parameter stationary process can be represented as a sum of sinusoids involving a continuous range of frequencies over $(-\pi, \pi)$. Expression (6.2.19) gives a physical interpretation of $\mathrm{d}F(w)$. It represents the contribution to the total power from the components of $X(t)$ with frequencies between w and $w + \mathrm{d}w$. Thus the spectrum $f(w)$ represents the distribution of the power density over frequency. For a zero mean continuous parameter process $X(t)$, there exists a similar spectral representation of the form

$$X(t) = \int_{-\infty}^{\infty} \mathrm{e}^{itw} \mathrm{d}Z(w) \tag{6.2.20}$$

where $\{\mathrm{d}Z(w)\}$ is an orthogonal process over $(-\infty, \infty)$.

6.2.4. Linear representations for stationary processes

Apart from the spectral representation, there is an alternative linear representation of stationary processes involving orthogonal variables. This enables us to represent any stationary process with a continuous spectrum as a linear combination of white noise processes. Consider a stationary process $X(t)$ with an absolutely continuous spectrum, so that its spectral density function $f(w)$ exists for all w. Hence from (6.2.19) we have

$$E(|\mathrm{d}Z(w)|^2) = f(w)\mathrm{d}w \tag{6.2.21}$$

where $\mathrm{d}Z(w)$ is as given in (6.2.19). Since $f(w) \geq 0$ for all w, we can find some function $\psi(w)$ such that

$$f(w) = |\psi(w)|^2 = \psi(w)\, \psi^*(w). \tag{6.2.22}$$

Assuming that $\mathrm{Var}(X(t)) < \infty, \psi(w)$ is quadratically integrable and hence can be expanded in a Fourier series of the form

$$\psi(w) = \sum_{k=-\infty}^{\infty} a_k \mathrm{e}^{-iwk}. \tag{6.2.23}$$

Assuming that $f(w)$ is strictly positive over $(-\pi, \pi)$ we get $\psi(w) > 0$ for all w. Now let

$$\mathrm{d}U(w) = \frac{\mathrm{d}Z(w)}{\psi(w)}. \tag{6.2.24}$$

Then (6.2.18) can be rewritten as

$$X(t) = \int_{-\pi}^{\pi} \mathrm{e}^{iwt} \psi(w) \mathrm{d}U(w). \tag{6.2.25}$$

Using (6.2.23) we get

$$X(t) = \sum_{k=-\infty}^{\infty} a_k \left(\int_{-\pi}^{\pi} \mathrm{e}^{i(t-k)w} \mathrm{d}U(w) \right) = \sum_{k=-\infty}^{\infty} a_k \varepsilon(t-k) \tag{6.2.26}$$

where

$$\varepsilon(t) = \int_{-\pi}^{\pi} e^{itw} dU(w). \tag{6.2.27}$$

Obviously $\varepsilon(t)$ is a white noise process. Expression (6.2.26) shows that any stationary process $X(t)$ can be represented as a linear combination of white noises. If we impose the condition that $\log f(w)$ is integrable, that is,

$$\int_{-\pi}^{\pi} \log f(w) d(w) > -\infty \tag{6.2.28}$$

then it is possible to find a function $\psi(w)$ satisfying (6.2.22), having a one sided Fourier series expansion of the form

$$\psi(w) = \sum_{k=0}^{\infty} b_k e^{-iwk}. \tag{6.2.29}$$

Then we can express $X(t)$ in the form

$$X(t) = \sum_{k=0}^{\infty} b_k \varepsilon(t-k). \tag{6.2.30}$$

It may be noted that this representation forms a special case of the Wold Decomposition Theorem.

6.2.5. Wold decomposition theorem

Let $X(t), t = 0, \pm 1, \pm 2, \cdots$ be a zero mean WSS stochastic process. Then $X(t)$ can be expressed as the sum of two zero mean WSS processes $U(t)$ and $V(t)$ in the form

$$X(t) = U(t) + V(t) \tag{6.2.31}$$

such that

(i) the process $U(t)$ is uncorrelated with $V(t)$;

(ii) $U(t)$ has a representation as a linear combination of white noises in the form

$$U(t) = \sum_{k=0}^{\infty} a_k \varepsilon(t-k), a_0 = 1, \sum_{k=0}^{\infty} a_k^2 < \infty; \tag{6.2.32}$$

(iii) the process $V(t)$ is completely determined by linear functions of its past values.

Exercises 6.2.

6.2.1. Define a spectral density function. How does it resemble the probability density function?

6.2.2. If the covariance function $R(s) = e^{-as}, s > 0$ find the expression for the spectral density function.

6.2.3. Compare the properties of spectral distribution function and cumulative distribution function in probability theory.

6.2.4. Describe how spectral density function and covariance function are related. Can you develop an inversion theorem in this context?

6.2.5. Explain Wold decomposition theorem. Can you state an analogous theorem in Probability Theory.

6.2.6. Suppose that $X(t)$ is a random telegraph signal with covariance function $R(s) = e^{-2\lambda|s|}$. Obtain the expression for the spectral density function.

6.2.7. Suppose that the covariance function of a process $X(t)$ is given by

$$R(s) = \begin{cases} 1 - \frac{|s|}{T} & ; \ |s| < T \\ 0 & ; \ |s| > T \end{cases}$$

Find the expression for the spectral density function.

6.2.8. If $X(t)$ is a sum of Poisson impulses such that $X(t) = \sum_i \delta(t - t_i)$. Obtain the covariance function and spectral density function.

6.2.9. If the covariance function is

$$R(s) = \frac{\sin(as)}{\pi s}$$

obtain the spectral density function.

6.2.10. Find the spectral density function if the covariance function is (i) $R(s) = e^{-\alpha s^2}$ and (ii) $R(s) = e^{-\alpha s^2} \cos(w_0 s)$.

6.3. Linear Systems and Linear Filters

The term linear filter is an abbreviated form of the title time invariant linear transformation. A linear filter is characterized by three basic properties namely, scale preservation, superposition and time invariance. Consider two stochastic processes $X(t)$ and $Y(t)$. Then a linear filter transforms the process, $X(t)$ called input into the process $Y(t)$ called output. This defines a linear system given by the relation

$$Y(t) = \sum_{k=0}^{\infty} a_k X(t-k) \tag{6.3.1}$$

where a_k is a sequence of constants called response function coefficients. If $X(t)$ is a zero mean stationary process with a bounded spectral density function $f_x(w)$ and $\sum_{k=0}^{\infty} a_k^2 < \infty$, then $Y(t)$ is a zero mean stationary process with spectral density function given by

$$f_y(w) = f_x(w)|A(\mathrm{e}^{-iw})|^2 \tag{6.3.2}$$

where

$$A(z) = \sum_{k=0}^{\infty} a_k z^k. \tag{6.3.3}$$

The function $A(z)$ is called the transfer function of the filter.

Using the concept of filters, any stationary process $X(t)$ can be considered as the output of a linear filter with white noise input. Then the spectral density $f_x(w)$ of $X(t)$ can be given as

$$f_x(w) = \frac{\sigma_\varepsilon^2}{2\pi}|B(\mathrm{e}^{-iw})|^2 \tag{6.3.4}$$

where $B(z) = \sum_{k=0}^{\infty} b_k z^k$. Then from (6.2.29), we have

$$\psi(w) = B(\mathrm{e}^{-iw}) \text{ and } f_x(w) = \frac{\sigma_\varepsilon^2}{2\pi}\Big|\sum_{k=0}^{\infty} b_k \mathrm{e}^{-iwk})\Big|^2. \tag{6.3.5}$$

6.3.1. Linear models

Linear models and filters occupy a central role in time series analysis in the time domain. The most commonly used linear models are the Moving Average (MA) model, Auto Regressive (AR) model and the mixed Auto Regressive Moving Average ($ARMA$) model. Box and Jenkins (1970) have considered a more general model called Auto Regressive Integrated Moving Average ($ARIMA$) model. All these linear models are generally reffered to as the Box-Jenkins models. A detailed description of such models and their analysis are given in 3rd and 4th SERC School Notes , CMS publication series No.32&33.

6.3.2. Moving average model $MA(p)$

By a moving average model of order p, denoted by $MA(p)$, we mean a linear filter of the form

$$X(t) = \varepsilon(t) + b_1\varepsilon(t-1) + \cdots + b_p\varepsilon(t-p) \tag{6.3.6}$$

where $\varepsilon(t)$ is a white noise process. In operator form, this can be written as

$$X(t) = P(B)\ \varepsilon(t)$$

where

$$P(z) = 1 + b_1 z + b_2 z^2 + \cdots + b_p z^p \tag{6.3.7}$$

and

$$BX(t) = X(t-1), \cdots, B^i X(t) = X(t-i).$$

$MA(p)$ can be considered as an approximation of a more general form of $X(t)$ in the form

$$X(t) = \sum_{k=o}^{\infty} b_k \varepsilon(t-k) \tag{6.3.8}$$

by a finite parameter model. Then the spectral density function of $X(t)$ is given by

$$f(w) = \frac{\sigma_\varepsilon^2}{2\pi}\left|1 + \sum_{k=1}^{p} b_k e^{-iwk}\right|^2. \tag{6.3.9}$$

Example 6.3.1. An $MA(1)$ model is given by $X(t) = b_1\varepsilon(t-1) + \varepsilon(t)$.

6.3.3. Autoregressive model $AR(q)$

An autoregressive model of order q, ($AR(q)$), is a filter of the form

$$X(t) + a_1 X(t-1) + \cdots + a_q X(t-q) = \varepsilon(t). \qquad (6.3.10)$$

In operator form

$$Q(B)X(t) = \varepsilon(t)$$

where

$$Q(z) = 1 + a_1 z + a_2 z^2 + \cdots + a_q z^q. \qquad (6.3.11)$$

The condition for the invertibility of $X(t)$, so as to express $X(t)$ as a linear combination of present and past values is that $Q(z)$ has no zeros inside and on the unit circle. This condition is of little consequence other than it guarantees a unique correspondence between the autocovariance structure and AR model. If $Q(z)$ is not invertible then there may be different moving average representations. Hence invertibility is a mathematical nicety, rather than a necessity (McIntire, 1977). If $X(t)$ satisfies (6.3.10) then its spectral density function is given by

$$f_x(w) = \frac{\sigma_\varepsilon^2}{2\pi}\Big|1 + \sum_{k=1}^{q} a_k e^{-iwk})\Big|^{-2}. \qquad (6.3.12)$$

Example 6.3.2. An $AR(1)$ process is given by

$$X(t) = aX(t-1) + \varepsilon(t)$$

This can be expressed as an infinite order moving average process.

6.3.4. Autoregressive moving average model $ARMA(q, p)$

A stochastic process $X(t)$ is called an $ARMA(q, p)$ model if it satisfies the difference equation.

$$X(t) + a_1 X(t-1) + \cdots + a_q X(t-q) = \varepsilon(t) + b_1 \varepsilon(t-1) + \cdots + b_p \varepsilon(t-p). \qquad (6.3.13)$$

In operator form this can be given as

$$Q(B)X(t) = P(B)\,\varepsilon(t) \qquad (6.3.14)$$

where $Q(z)$ and $P(z)$ are defined by (6.3.7) and (6.3.11). The model given by (6.3.14) possesses a stationary solution only if $Q(z)$ has no zeros inside or on the unit circle.

An $ARMA(q, p)$ model is characterized by a rational transfer function of the form

$$R(z) = \frac{1 + b_1 z + b_2 z^2 + \cdots + b_p z^p}{1 + a_1 z + a_2 z^2 + \cdots + a_q z^q} = \frac{P(z)}{Q(z)}. \tag{6.3.15}$$

In engineering literature, MA models are known as all zero models while AR models are known as all-pole models. $ARMA$ models are referred to as pole-zero models. If the process $X(t)$ satisfies (6.3.13), its spectral density function is given by

$$f(w) = \frac{\sigma_\varepsilon^2}{2\pi} \frac{|1 + \sum_{k=1}^{p} b_k \mathrm{e}^{-iwk}|^2}{|1 + \sum_{k=1}^{q} a_k \mathrm{e}^{-iwk}|^2}. \tag{6.3.16}$$

An $ARMA(1, 1)$ process is having the general structure

$$X(t) = aX(t-1) + b\varepsilon(t-1) + \varepsilon(t).$$

6.3.5. Autoregressive integrated moving average model $ARIMA(q, r, p)$

Box and Jenkins (1970) consider a slightly more general model called $ARIMA$ (q, r, p) in which case $Q(B)$ in (6.3.14) contains a factor of the form $(1 - B)^r$. Then the transfer function $R(z)$ has a r^{th} order pole at $z = 1$. Consequently $X(t)$ turns out to be a nonstationary process. However, its r^{th} difference $\Delta^r X(t)$ would be stationary. When $r = 0$, we get the $ARMA(q, p)$ model. The $ARIMA(q, r, p)$ model can be given in the operator form as

$$\phi(B)(1 - B)^r X(t) = \theta(B)\varepsilon(t). \tag{6.3.17}$$

This can be rewritten as

$$\phi(B)Y(t) = \theta(B)\varepsilon(t),\ Y(t) = \Delta^r X(t),\ \Delta = 1 - B. \tag{6.3.18}$$

Then the reduced process $Y(t)$ will be a stationary $ARMA$ model.

6.3.6. Processes with rational spectrum

Definition 6.3.1. A zero mean stationary process $X(t)$ is said to have a rational spectrum if it has a spectral density of the form

$$f(w) = R(e^{-iw}) \tag{6.3.19}$$

where R is a rational function.

The following theorem characterizes an *ARMA* model as one having a rational spectrum.

Theorem 6.3.1. *A zero mean stationary process $X(t)$ is an ARMA process having a spectral density iff $X(t)$ has rational spectrum.*

A process $X(t)$ with rational spectrum has a spectral density of the form

$$f_x(w) = \frac{\sum_{k=0}^{p} c_k e^{-iwk}}{\sum_{k=0}^{q} d_k e^{-iwk}} \tag{6.3.20}$$

where the numerator and denominator have no factors in common. In order that $f_x(w)$ to be integrable, it is necessary that the denominator does not vanish for $-\pi < w \leq \pi$. The importance of this type of models results from the fact that any continuous spectral density function can be approximated arbitrarily closely by a rational function such as (6.3.15) by proper choice of p, q and the other parameters. A model with a rational spectral density will almost invariably lead to a better fit with fewer parameters (Koopmans, 1974). This property is referred to as parsimony by Box and Jenkins (1970).

Exercises 6.3.

6.3.1. Define an *AR*(1) model. Obtain its covariance function and spectral density function, clearly stating the assumptions.

6.3.2. Show that an *AR*(1) model can be reduced to an infinite order *MA* model.

6.3.3. What do you mean by Box-Jenkins models?

6.3.4. Derive the covariance function of a *MA*(1) model. Obtain its spectral density function.

6.3.5. What is an *ARMA*(1, 1) model. Derive its autocovariance function and spectral density function.

6.3.6. Describe how various *ARMA* models can be obtained from moving average models through a linear transfer function on $X(t)$.

6.4. Mittag-Leffler Processes

The function

$$E_\alpha(z) = \sum_{k=0}^{\infty}\left[\frac{z^k}{\Gamma(1+ak)}\right]$$

was first introduced by Mittag-Leffler in 1903 (Erdélyi, et al (1955)). It was subsequently investigated by Wiman, Pollard, Humbert, Aggarwal and Feller. Many properties of the function follow from Mittag-Leffler integral representation

$$E_\alpha(z) = \frac{1}{2\pi i}\int_C \frac{y^{\alpha-1}e^y}{y^\alpha - z}dy.$$

where the path of integration C is a loop which starts and ends at $-\infty$ and encircles the circular disc $|y| \leq z^{\frac{1}{\alpha}}$. Figure (6.4.1) gives the graph of $E_\alpha(z)$ for various values of α.

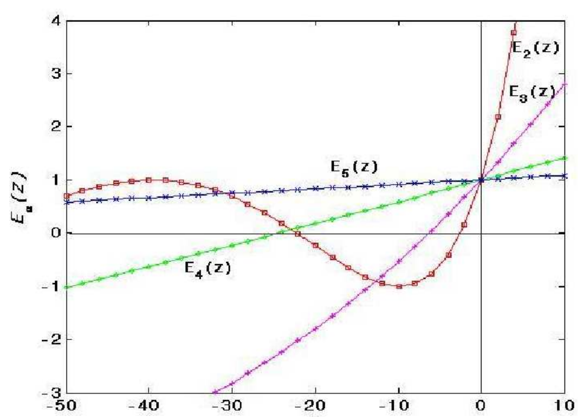

Figure 6.4.1. Graph of $E_\alpha(z)$

Recently Mittag-Leffler functions and distributions have received considerable attention of mathematicians, statisticians and scientists in physical and chemical sciences. Pillai (1990a) introduced the Mittag-Leffler distribution in terms of a Mittag-Leffler function. He proved that $F_\alpha(x) = 1 - E_\alpha(-x^\alpha), 0 < \alpha \leq 1$ are distribution functions, having the Laplace transform $\psi(s) = (1 + s^\alpha)^{-1}, s > 0$ which is completely monotone for $0 < \alpha \leq 1$. He called $F_\alpha(x)$, for $0 < \alpha \leq 1$, a Mittag-Leffler distribution. The Mittag-Leffler distribution is a generalization of the exponential distribution, since for $\alpha = 1$, we get exponential distribution. Pillai (1990a) has shown that $F_\alpha(x)$ is geometrically infinitely divisible (g.i.d.) and is in the domain of attraction of stable laws. It can be used in reliability modeling as an alternative for exponential lifetime distribution.

A discrete version of the Mittag-Leffler distribution was introduced by Pillai and Jayakumar (1995). Lin (1998a, b) obtained various characterizations of the Mittag-Leffler distributions and referred to them as positive Linnik laws. In Physics, Haubold and Mathai (2000) derived a closed form representation of the fractional kinetic equation and thermonuclear function in terms of Mittag-Leffler function. Saxena et al. (2004a, b) extended the result and derived the solutions of a number of fractional kinetic equations in terms of generalized Mittag-Leffler functions and obtained the solution of a unified form of generalized fractional kinetic equations, which provides the unification and extension of the earlier results. These results are useful in explaining various fundamental laws of Physics like Maxwell's equations, Schroedinger's equation, Newton's laws of motion and Einstein's equations for geodesics. Such behaviors occur frequently in biology, chemistry, thermodynamical and statistical analysis. In all such situations the solutions can be expressed in terms of generalized Mittag-Leffler functions. Weron and Kotulski (1996) use Mittag-Leffler distribution in explaining cole-cole relaxation.

Modelling and forecasting of Wolf's sunspot numbers has been one of the celebrated problems in time series modeling. Many Gaussian and non-Gaussian as well as linear and non-linear time series models have been developed by various researchers. Gaver and Lewis (1980) derived the exponential solution of first order autoregressive equation $X_n = \rho X_{n-1} + \xi_n, n = 0, \pm 1, \pm 2, \ldots$ where$\{\xi_n\}$ is a sequence of independently and identically distributed random variables when $0 \leq \rho < 1$. Jayakumar and Pillai (1993) developed a first order autoregressive process with Mittag-Leffler marginal distribution. Fujita (1993) discussed a generalization of the results of Pillai. Jose and Seetha Lekshmi (1997) developed geometric exponential distribution and Seetha Lekshmi and Jose (2002,2003) extended the results to obtain geometric Mittag-Leffler distributions. Jayakumar and Ajitha (2003) obtained various results on geometric Mittag-Leffler distributions. Seetha Lekshmi and Jose (2004) extended the concept to obtain geometric α-Laplace processes.

6.4.1. Generalized Mittag-Leffler distribution

In this section we introduce a new class of distributions called generalized Mittag-Leffler distribution denoted by $GMLD(\alpha, \beta)$. A random variable with support over $(0, \infty)$ is said to follow the generalized Mittag-Leffler distribution with parameters α and β if its Laplace transform is given by

$$\psi(s) = E[\mathrm{e}^{-sX}] = (1 + s^{\alpha})^{-\beta}; 0 < \alpha \leq 1, \beta > 0, s > 0. \tag{6.4.1}$$

The cumulative distribution function (c.d.f.) corresponding to (6.4.1) is given by

$$F_{\alpha,\beta}(x) = P[X \leq x] = \sum_{k=0}^{\infty} \frac{(-1)^k \Gamma(\beta + k) x^{\alpha(\beta+k)}}{k! \Gamma(\beta) \Gamma(1 + \alpha(\beta + k))}. \tag{6.4.2}$$

It easily follows that when β=1, we get Pillai's Mittag-Leffler distribution (see Pillai (1990a)). When α=1, we get the gamma distribution. When α=1, β=1 we get the exponential distribution. This family may be regarded as the positive counterpart of Pakes generalized Linnik distribution characterized by the characteristic function

$$\phi_X(t) = E(\mathrm{e}^{itX}) = (1 + |t|^{\alpha})^{-\beta}; 0 < \alpha \leq 2, \beta > 0, -\infty < t < \infty, \tag{6.4.3}$$

(see Pakes (1998)). The Pakes generalized Linnik distribution consists of variuos families of distributions like the Laplace distribution, α-Laplace distributin of Pillai (1985), generalized Laplace distribution of Mathai (1993) etc. More details are available in Seetha Lekshmi and Jose (2006). Now we shall discuss some properties of generalized Mittag-Leffler distributions.

Theorem 6.4.1. *Let U_α follow the positive stable distribution with Laplace transform $\psi(s) = \exp(-s^{\alpha}), s > 0, 0 < \alpha \leq 1$. Let V_β be independent of U_α, and follow a gamma distribution with parameter β having Laplace transform $\phi_{V_\beta}(s) = \left(\frac{1}{1+s}\right)^{\beta}, \beta > 0$. Then $X_{\alpha,\beta} = U_\alpha V_\beta^{\frac{1}{\alpha}}$ follows generalized Mittag-Leffler distribution GMLD(α, β).*

Proof 6.4.1. The Laplace transform of $X_{\alpha,\beta}$ is

$$\begin{aligned}\psi_{X_{\alpha,\beta}}(s) &= E(\mathrm{e}^{-sV_\beta^{\frac{1}{\alpha}}U_\alpha}) = \int_0^\infty \psi_{U_\alpha}(sV_\beta^{\frac{1}{\alpha}})\mathrm{d}F(v) \\ &= \int_0^\infty \mathrm{e}^{-s^\alpha V_\beta}\mathrm{d}F(v) = \left(\frac{1}{1+s^\alpha}\right)^\beta.\end{aligned}$$

Theorem 6.4.2. *The generalized Mittag-Leffler distribution is a mixture of gamma densities.*

Proof 6.4.2. Using Theorem 6.4.1, we have the c.d.f. of $X_{\alpha,\beta}$ as

$$F_{\alpha,\beta}(x) = \int_0^\infty S_{\alpha,v}(x)\mathrm{d}F_\beta(v) \tag{6.4.4}$$

where $S_{\alpha,v}(x)$ is the c.d.f. of a distribution with Laplace transform $\exp(-vs^\alpha)$ and $F_\beta(v)$ is the c.d.f. of gamma distribution with p.d.f. $f(y) = \frac{1}{\Gamma(\beta)}y^{\beta-1}\mathrm{e}^{-y}$. We can rewrite (6.4.4) as

$$F_{\alpha,\beta}(x) = \int_0^\infty G_\beta((x|y)^\alpha)\mathrm{d}S_{\alpha,1}(y)$$

Hence the p.d.f. of $X_{\alpha,\beta}$ is

$$f_{\alpha,\beta}(x) = \frac{\mathrm{d}}{\mathrm{d}x}F_{\alpha,\beta}(x) = \int_0^\infty \frac{\alpha}{\Gamma(\beta)}\frac{x^{\alpha\beta-1}}{y^{\alpha\beta}}\mathrm{e}^{-(x|y)^\alpha}\mathrm{d}S_{\alpha,1}(y).$$

This shows that $f_{\alpha,\beta}$ is a mixture of generalized gamma densities. For $\beta = 1$, $f_{\alpha,\beta}$ reduces to a mixture of Weibull densities.

Remark 6.4.1. Lin (1998b) has shown that $F_{\alpha,\beta}(x)$ is slowly varying at infinity, for $\alpha \in (0,1]$ and $\beta > 0$.

Remark 6.4.2. Pillai (1990a) had obtained the fractional moments for $\beta = 1$, as

$$E(X_\alpha^r) = \frac{\Gamma(1-r/\alpha)\Gamma(1+r/\alpha)}{\Gamma(1-r)}; 0 < r < \alpha \le 1.$$

In a similar manner Lin (1998b) obtained the fractional moments for $X_{\alpha,\beta}$ for $0 < \alpha \le 1$ and $\beta > 0$ as

$$E(X_{\alpha,\beta}^r) = \begin{cases}\frac{\Gamma(1-r/\alpha)\Gamma(\beta+r/\alpha)}{\Gamma(1-r)\Gamma(\beta)}, & -\alpha\beta < r < \alpha \\ \infty, r \le -\alpha\beta \text{ or } r \ge \alpha. \end{cases}$$

Definition 6.4.1. A probability distribution on $R_+ = (0, \infty)$ is said to be in class L if its Laplace transform $\psi(s)$ satisfies

$$\psi(s) = \psi(ks)\psi_k(s), s \in R, k \in (0, 1) \tag{6.4.5}$$

with ψ_k a Laplace transform. Such distributions are called self-decomposable.

Theorem 6.4.3. *The generalized Mittag-Leffler distribution belongs to the class L of selfdecomposable distributions.*

Proof 6.4.3. The proof follows due to the relation

$$(1 + s^\alpha)^{-\beta} = (1 + k^\alpha s^\alpha)^{-\beta} \left(k^\alpha + (1 - k^\alpha) \frac{1}{1 + s^\alpha} \right)^\beta . \tag{6.4.6}$$

Theorem 6.4.4. *The GMLD (α, β) is geometrically infinitely divisible for $0 < \alpha \leq 1, 0 < \beta \leq 1$.*

Proof 6.4.4. From Pillai and Sandhya (1990), a distribution is g.i.d. if and only if its Laplace transform is of the form $\psi(s) = \frac{1}{1+\phi(s)}$ where $\phi(s)$ has complete monotone derivative (c.m.d.) and $\phi(0) = 0$. Now for GMLD (α, β), the Laplace transform is $\psi(s) = \frac{1}{1+\phi(s)}$ where $\phi(s) = (1 + s^\alpha)^\beta - 1$. This has c.m.d. if and only if $0 < \alpha \leq 1$ and $0 < \beta \leq 1$.

Remark 6.4.3. Also being a mixture of gamma random variables, it is g.i.d. provided $0 < \beta \leq 1$. By Pillai and Sandhya (1990), GMLD is a distribution with c.m.d. Hence it is infinitely divisible.

6.4.2. Processes with GMLD(α, β) marginals

Now we shall construct a first order time series model with GMLD marginals. The generalized Mittag-Leffler first order autoregressive process GMLAR(1) is constituted by $\{X_n; n \geq 1\}$ where X_n satisfies the equation

$$X_n = \rho X_{n-1} + \xi_n, 0 < \rho < 1 \tag{6.4.7}$$

where $\{\xi_n\}$ is a sequence of independently and identically distributed random variables such that X_n is stationary Markovian with generalized Mittag-Leffler marginal distribution. In terms of Laplace transform (6.4.7) can be given as

$$\psi_{X_n}(s) = \psi_{\xi_n}(s)\psi_{X_{n-1}}(\rho s). \tag{6.4.8}$$

Assuming stationarity we have,

$$\begin{aligned}\psi_\xi(s) &= \frac{\psi_X(s)}{\psi_X(\rho s)} = \frac{(1+\rho^\alpha s^\alpha)^\beta}{(1+s^\alpha)^\beta} \\ &= \left[\frac{1+(\rho s)^\alpha}{1+s^\alpha}\right]^\beta = \left[\rho^\alpha + (1-\rho^\alpha)\frac{1}{1+s^\alpha}\right]^\beta. \end{aligned} \tag{6.4.9}$$

We can regard the innovations $\{\xi_n\}$ as the β-fold convolutions of random variables ϵ_n's such that

$$\epsilon_n = \begin{cases} 0 & \text{with probability } \rho^\alpha \\ M_n & \text{with probability } 1-\rho^\alpha \end{cases}$$

where M_n's are independently and identically distributed Mittag-Leffler random variables. Mittag-Leffler random variables can be generated easily by using the following result. Let E be distributed as exponential with unit mean and let U_α be distributed as positive stable with Laplace transform e^{-s^α}. Then $X = U_\alpha E^{\frac{1}{\alpha}}$ will be distributed as Mittag-Leffler with Laplace transform $(1+s^\alpha)^{-1}$ (see, Kozubowski and Rachev(1999)).

Jayakumar et al (1995) developed an algorithm to generate Linnik random variables. In a similar manner, the GMLAR(1) process can be generated by using computers. If $X_0 \stackrel{d}{=} GMLD(\alpha,\beta)$, then the process is strictly stationary. It is sufficient to verify that $X_n \stackrel{d}{=} GMLD(\alpha,\beta)$ for every n. An inductive argument can be presented as follows. Suppose $X_{n-1} \stackrel{d}{=} GMLD(\alpha,\beta)$. Then from (6.4.8) we have,

$$\psi_{X_n}(s) = \left[\frac{1+(\rho s)^\alpha}{1+s^\alpha}\right]^\beta \left[\frac{1}{1+(\rho s)^\alpha}\right]^\beta = \left[\frac{1}{1+s^\alpha}\right]^\beta.$$

Hence the process is strictly stationary and Markovian provided X_0 is distributed as GMLD.

Remark 6.4.4. If X_0 is distributed arbitrarily, then also the process is asymptotically Markovian with generalized Mittag-Leffler marginal distribution.

Proof 6.4.5. This follows from the following lines. We have

$$X_n = \rho X_{n-1} + \xi_n = \rho^n X_0 + \sum_{k=0}^{n-1} \rho^k \xi_{n-k}.$$

Writing in terms of Laplace transform,

$$\begin{aligned}\psi_{X_n}(s) &= \psi_{X_0}(\rho^n s)\prod_{k=0}^{n-1}\psi_{\xi}(\rho^k s)\\ &= \psi_{X_0}(\rho^n s)\prod_{k=0}^{n-1}\left[\frac{1+(\rho^{k+1}s)^{\alpha}}{1+(\rho^k s)^{\alpha}}\right]^{\beta} \longrightarrow (1+s^{\alpha})^{-\beta} \text{ as } n\to\infty.\end{aligned}$$

Hence it follows that even if X_0 is arbitrarily distributed the process is asymptotically stationary Markovian with generalized Mittag-Leffler marginals. We therefore have the following theorem.

Theorem 6.4.5. *The first order autoregressive process* $X_n = \rho X_{n-1} + \xi_n, \rho \in (0,1)$ *is strictly stationary Markovian with generalized Mittag-Leffler marginal distribution as in (6.4.1) if and only if the* $\{\xi_n\}$ *are distributed independently and identically as the* β *-fold convolution of the random variable* $\{\epsilon_n\}$ *where*

$$\epsilon_n = \begin{cases} 0 & \textit{with} \quad \textit{probability} \quad \rho^{\alpha} \\ M_n & \textit{with} \quad \textit{probability} \quad 1-\rho^{\alpha} \end{cases}$$

where $\{M_n\}$ *are independently and identically distributed Mittag-Leffler random variables provided* $X_0 \stackrel{d}{=} GMLD(\alpha,\beta)$ *and independent of* ξ_n.

Remark 6.4.5. The model is defined for all values of ρ such that $\rho \in (0,1)$.The autocorrelation is given by $\rho(r) = Cor(X_n, X_{n-r}) = \rho^{|r|}, r = 0, \pm 1, \pm 2, \ldots$

6.4.3. Distribution of sums and bivariate distribution of (X_n, X_{n+1})

We have,

$$X_{n+j} = \rho^j X_n + \rho^{j-1}\xi_{n+1} + \rho^{j-2}\xi_{n+2} + \ldots + \xi_{n+j};\ j = 0,1,2,\ldots$$

Hence

$$\begin{aligned}T_r &= X_n + X_{n+1} + \ldots + X_{n+r-1}\\ &= \sum_{j=0}^{r-1}[\rho^j X_n + \rho^{j-1}\xi_{n+1} + \rho^{j-2}\xi_{n+2} + \ldots + \xi_{n+j}]\\ &= X_n\left[\frac{1-\rho^r}{1-\rho}\right] + \sum_{j=1}^{r-1}\xi_{n+j}\left[\frac{1-\rho^{r-j}}{1-\rho}\right].\end{aligned}$$

Therefore the distribution of the sums T_r is uniquely determined by the Laplace transform

$$\begin{aligned}\psi_{T_r}(t) &= \psi_{X_n}\left(\frac{1-\rho^r}{1-\rho}t\right)\prod_{j=1}^{r-1}\psi_{\xi}\left(\frac{1-\rho^{r-j}}{1-\rho}s\right) \\ &= \frac{1}{\left[1+\left(\frac{1-\rho^r}{1-\rho}t\right)^{\alpha}\right]^{\beta}}\prod_{j=1}^{r-1}\left[\rho^{\alpha}+(1-\rho^{\alpha})\frac{1}{\left[1+\left(\frac{1-\rho^{r-j}}{1-\rho}s\right)^{\alpha}\right]}\right]^{\beta}.\end{aligned}$$

The distribution of T_r can be obtained by inverting the above expression. Next, the joint distribution of contiguous observations (X_n, X_{n+1}) can be given in terms of bivariate Laplace transform as,

$$\begin{aligned}\psi_{X_n,X_{n+1}}(s_1,s_2) &= E[\exp(-s_1X_n - s_2X_{n+1})] \\ &= E[\exp(-s_1X_n - s_2(\rho X_n+\xi_n))] \\ &= E[\exp(-(s_1+\rho s_2)X_n - s_2\xi_{n+1})] \\ &= \psi_{\xi_n}(s_2)\psi_{X_n}(s_1+\rho s_2) \\ &= \left[\frac{1+(\rho s_2)^{\alpha}}{1+s_2^{\alpha}}\right]^{\beta}\left[\frac{1}{1+(s_1+\rho s_2)^{\alpha}}\right]^{\beta}.\end{aligned}$$

Since this expression is not symmetric in s_1 and s_2, it follows that the GMLAR(1) process is not time reversible.

6.4.4. Geometric generalized Mittag-Leffler distribution

Geometric generalized Mittag-Leffler distribution is introduced here and some of its properties are studied.

Definition 6.4.2. A random variable X on $R_+ = (0,\infty)$ is said to follow geometric generalized Mittag-Leffler distribution and written as $X \overset{d}{=} GGMLD(\alpha,\beta)$ if it has the Laplace transform

$$\psi(s) = \frac{1}{1+\beta\ln(1+s^{\alpha})}, 0<\alpha\le 1, \beta>0, s>0. \tag{6.4.10}$$

Remark 6.4.6. Geometric generalized Mittag-Leffler distribution is geometrically infinitely divisible.

Theorem 6.4.6. *Let $X_1, X_2, \ldots$ be independently and identically distributed geometric Mittag-Leffler random variables and let $N(p)$ be geometric with mean $\frac{1}{p}$, $P[N(p) = k]=p(1-p)^{k-1}$, $k = 1, 2, \ldots, 0 < p < 1$. Define $Y = X_1 + X_2 + \ldots + X_{N(p)}$, then $Y \stackrel{d}{=} GGMLD(\alpha,\beta)$ where $\beta = \frac{1}{p}$.*

Proof 6.4.6. The Laplace transform of Y is

$$\psi_Y(s) = \sum_{k=1}^{\infty}[\psi_X(s)]^k p(1-p)^{k-1} = \frac{1}{1+\frac{1}{p}\ln(1+s^\alpha)}.$$

Hence $Y \stackrel{d}{=} GGMLD(\alpha, \frac{1}{p})$.

Theorem 6.4.7. *Geometric generalized Mittag-Leffler distribution is the limit of geometric sum of $GML(\alpha, \frac{\beta}{n})$ random variables.*

Proof 6.4.7. $(1+s^\alpha)^{-\beta} = \{1 + (1+s^\alpha)^{\frac{\beta}{n}} - 1\}^{-n}$ is the Laplace transform of a probability distribution since generalized Mittag-Leffler distribution is infinitely divisible. Hence by Lemma 3.2 of Pillai (1990b)

$$\psi_n(s) = \{1 + n[(1+s^\alpha)^{\frac{\beta}{n}} - 1]\}^{-n}$$

is the Laplace transform of a geometric sum of independently and identically distribute generalized Mittag-Leffler random variables. Then,

$$\begin{aligned}\psi(s) &= \lim_{n\to\infty}\psi_n(s) = \{1 + \lim_{n\to\infty} n[(1+s^\alpha)^{\frac{\beta}{n}} - 1]\}^{-1}\\ &= [1+\beta\ln(1+s^\alpha)]^{-1}.\end{aligned}$$

Theorem 6.4.8. *If W_β and V_α are independent random variables such that W_β has geometric gamma distribution with Laplace transform $\frac{1}{1+\beta\ln(1+s)}$ and V_α has a positive stable distribution having Laplace transform e^{-s^α}, then $W_\beta^{\frac{1}{\alpha}}V_\alpha = U_{\alpha,\beta}$ and $U_{\alpha,\beta} \stackrel{d}{=} GGMLD(\alpha,\beta)$.*

Proof 6.4.8. The Laplace transform of $U_{\alpha,\beta}$ is

$$\begin{aligned}\psi_{U_{\alpha,\beta}}(s) &= E(\mathrm{e}^{-sW_\beta^{\frac{1}{\alpha}}V_\alpha}) = \int_0^\infty \psi_{V_\alpha}(sW_\beta^{\frac{1}{\alpha}})\mathrm{d}F(w)\\ &= \int_0^\infty \mathrm{e}^{-W_\beta s^\alpha}\mathrm{d}F(w) = \frac{1}{1+\beta\ln(1+s^\alpha)}.\end{aligned}$$

6.4.5. Geometric generalized Mittag-Leffler processes

In this section, we develop a first order new autoregressive process with geometric generalized Mittag-Leffler marginals.

Consider an autoregressive structure given by,

$$X_n = \begin{cases} \xi_n & \text{with probability } p \\ X_{n-1} + \xi_n & \text{with probability } (1-p) \end{cases} \tag{6.4.11}$$

where $0 < p < 1$. Now we shall construct an AR (1) process with stationary marginal distribution as geometric generalized Mittag-Leffler distribution $GGMLD(\alpha, \beta)$.

Theorem 6.4.9. *Consider a stationary autoregressive process $\{X_n\}$ with structure given by (6.4.11). A necessary and sufficient condition that $\{X_n\}$ is stationary Markovian with geometric generalized Mittag-Leffler marginal distribution is that $\{\xi_n\}$ is distributed as geometric Mittag-Leffler provided X_0 is distributed as geometric generalized Mittag-Leffler.*

Proof 6.4.9. Let us denote the Laplace transform of X_n by $\psi_{X_n}(s)$ and that of ξ_n by $\psi_{\xi_n}(s)$, equation (6.4.11) in terms of Laplace transform becomes,

$$\psi_{X_n}(s) = p\psi_{\xi_n}(s) + (1-p)\psi_{X_{n-1}}(s)\psi_{\xi_n}(s).$$

On assuming stationarity, it reduces to the form,

$$\psi_\xi(s) = \frac{\psi_X(s)}{p + (1-p)\psi_X(s)}.$$

Writing

$$\psi_X(s) = \frac{1}{1 + \beta \ln(1 + s^\alpha)} \text{ and solving we get,}$$

$$\psi_\xi(s) = \frac{1}{1 + \beta p \ln(1 + s^\alpha)}.$$

Hence it follows that $\xi_n \stackrel{d}{=} GGMLD(\alpha, p\beta)$.

The converse can be proved by the method of mathematical induction. Now assume that $X_{n-1} \stackrel{d}{=} GGMLD(\alpha,\beta)$. Then

$$\begin{aligned}\psi_{X_{n-1}}(s) &= \psi_{\xi_n}(s)\left[p+(1-p)\psi_{X_{n-2}}(s)\right]\\ &= \frac{1}{1+p\beta\ln(1+s^{\alpha})}\left[p+(1-p)\frac{1}{1+\beta\ln(1+s^{\alpha})}\right]\\ &= [1+\beta\ln(1+s^{\alpha})]^{-1}.\end{aligned}$$

The rest follows easily.

Remark 6.4.7. Note that X_n and ξ_n belong to the same family of distributions.

6.4.6. The joint distribution of X_n and X_{n-1}

Consider the autoregressive structure given in (6.4.11). It can be rewritten as

$$\begin{aligned}X_n &= I_n X_{n-1} + \xi_n \text{ where}\\ P[I_n = 0] &= 1 - P[I_n = 1] = p, 0 < p < 1.\end{aligned}$$

Then the joint Laplace transform of (X_n, X_{n-1}) is given by,

$$\begin{aligned}\psi_{X_n,X_{n-1}}(s_1,s_2) &= E(\mathrm{e}^{-s_1X_{n-1}-s_2X_n}) = E(\mathrm{e}^{-s_1X_{n-1}-s_2(I_nX_{n-1}+\xi_n)})\\ &= E(\mathrm{e}^{(-s_1-s_2I_n)X_{n-1}})\psi_{\xi_n}(s_2)\\ &= \frac{1}{1+\beta\mathrm{p}\ln(1+s_2^{\alpha})}\left[\frac{p}{1+\beta\ln(1+s_1^{\alpha})}+\frac{1-p}{1+\beta\ln(1+(s_1+s_2)^{\alpha})}\right]\end{aligned}$$

This shows that the process is not time reversible.

6.4.7. Generalizations

Now we construct a k^{th} order autoregressive process. Lawrance and Lewis (1982) constructed higher order analogs of the autoregressive equations (6.4.11) with structure,

$$X_n = \begin{cases} \xi_n & \text{with probability } p\\ X_{n-1}+\xi_n & \text{with probability } p_1\\ X_{n-2}+\xi_n & \text{with probability } p_2\\ \vdots & \\ X_{n-k}+\xi_n & \text{with probability } p_k\end{cases} \tag{6.4.12}$$

where $p_1 + p_2 + \ldots + p_k = 1 - p, 0 \le p_i, p \le 1, i = 1, 2, \ldots, k$ and ξ_n is independent of $\{X_n, X_{n-1}, \ldots\}$. In terms of Laplace transform, (6.4.12) can be given as

$$\psi_{X_n}(s) = p\psi_{\xi_n}(s) + p_1\psi_{X_{n-1}}(s)\psi_{\xi_n}(s) + p_2\psi_{X_{n-2}}(s)\psi_{\xi_n}(s) + \ldots + p_k\psi_{X_{n-k}}(s)\psi_{\xi_n}(s).$$

Assuming stationarity we get,

$$\psi_{\xi}(s) = \frac{\psi_X(s)}{p + (1-p)\psi_X(s)}.$$

This shows that the results developed in Section 6.4.5 can be applied in this case also. This gives rise to the k^{th} order geometric generalized Mittag-Leffler autoregressive processes.

6.4.8. Applications

Mittag-Leffler distributions can be used as waiting-time distributions as well as first-passage time distributions for certain renewal processes. Pillai (1990b) developed renewal processes with geometric exponential as waiting-time distribution. In a similar manner renewal processes with generalized Mittag-Leffler and geometric generalized Mittag-Leffler waiting-times can be constructed. Mittag-Leffler functions are also used for computation of the change of the chemical composition in stars like the Sun. Recent investigations have proved that Mittag-Leffler functions are certain aspects of the solution of the solar neutrino problems in astrophysics, which can be expressed in terms of special functions like G-, H- and Wright functions. The fundamental laws of physics are written as equations for time evolution of a certain quantity Z(t), which could be Maxwell's equations or Schrödinger's equation, Newton's laws of motion or Einstein's equations for geodesics. In thermodynamical or statistical applications, one is interested in mean values of a quantity Z(t). Tsallis (1988) generalized the entropic functional of Boltzmann-Gibbs statistical mechanics that leads to q-exponential distributions. Tsallis used the mathematical simplicity of kinetic-type equations to emphasize the natural outcome of this distribution that corresponds exactly to the solution of the kinetic equation of non-linear type; the solution has power-law behavior. Saxena et al (2004a, b) showed that the fractional generalization of the linear kinetic-type equation also leads to power-law behavior. In both cases, solutions can be expressed in terms of generalized Mittag-Leffler functions.

Exercises 6.4.

6.4.1. Define Mittag-Leffler distribution. Obtain the Laplace transform. How is it related to exponential distribution.

6.4.2. Describe how you can develop a generalized Mittag-Leffler distribution and an $AR(1)$ process with this as marginal distribution.

6.4.3. What is meant by geometric infinite divisibility? Give an example of such a distribution.

6.4.4. Extend the concept of generalized Mittag-Leffler distribution to the real line to develop a generalized α-Laplace distribution.

6.4.5. Define Mathai's generalized Laplacian distribution. Construct an $AR(1)$ process with this as stationary marginal distribution.

6.4.6. Describe how an exponential autoregressive process can be developed.

6.4.7. Derive the innovation structure of a Laplacian autoregressive process.

6.4.8. Explain how a generalized α-Laplace distribution can be defined. Describe how it reduces to various special cases in literature.

6.4.9. Construct $AR(1)$ models with geometric and negative binomial stationary marginal distributions.

6.4.10. Develop an $AR(1)$ process with stationary marginal distribution as that of a convolution of normal and generalized Laplacian random variables.

6.5. Min Geometric Generalized Pareto Processes

Pareto and generalized Pareto distributions fit well for a wide variety of socio economic variables with heavy tails. Yeh et al (1988) developed Pareto processes with Pareto type III marginal distributions described in Arnold (1983). Pillai (1991) developed semi-Pareto distribution and processes in which Pareto distribution is a particular case. These distributions are used for modeling heavy tailed data.

The concept of geometric infinite divisibility (GID) was introduced by Klebanov at al (1984) while solving Zolotarev's problem. A lot of work is being done by many authors for the last two decades in the field of GID and geometric stable distributions (see Pillai (1990), Pillai and Sandhya (1990), Mohan et al (1993), Kozubowski and Rachev (1999), Jose and Seetha Lekshmi (1999), Seetha Lekshmi et al (2003a) and Seetha Lekshmi and Jose (2002, 2003b and 2004 a and b etc.) The concept of max geometric infinite divisibility (maxGID) was introduced by Rachev and Resnick (1991) as a ramification of GID. Seetha Lekshmi et al (2005) introduced min geometric infinite divisibility (minGID) and developed min geometric stable distributions. More details are available in the 4th SERC School Notes, CMS Publication Series No 33.

6.5.1. Min infinite divisibility

In this section we consider some properties of min infinite divisibility and min geometric infinite divisibility.

Definition 6.5.1. A survival function $\bar{F}$ is said to be min infinitely divisible if $\bar{F}^{\frac{1}{n}}$ is a survival function for all $n > 1$ or equivalently $\bar{F}^t$ is a survival function for all $t \geq 0$. A random variable X is said to be min geometrically infinitely divisible if for every $\theta \in (0, 1)$ there exists a sequence of independently and identically distributed (i.i.d.) random variables $X_1^{(\theta)}, X_2^{(\theta)}, \ldots$ such that $Y \stackrel{d}{=} \Lambda_{j=1}^{N(\theta)} X_j^{(\theta)}$ where $N(\theta)$ is geometrically distributed as $P[N(\theta) = k] = \theta(1-\theta)^{k-1},\ 0 \leq \theta \leq 1,\ k = 1, 2, \ldots$ where $X, N(\theta)$ and $X_j^{(\theta)}$ for $j = 1, 2, \ldots, N(\theta)$ are independent random variables ($\stackrel{d}{=}$ denotes equality in distributions, Λ denotes minimum). In terms of survival function it is equivalent to

$$\bar{F}_X(x) = \frac{\theta \bar{F}_\theta(s)}{1 - (1-\theta)\bar{F}_\theta(x)}$$

where $\bar{F}_X(x)$ and $\bar{F}_\theta(x)$ denote respectively the survival function of X and $X_j^{(\theta)}$. Clearly this is a survival function and a distribution with this survival function is known as min geometrically infinitely divisible distribution. For more details see Seetha Lekshmi et al (2005).

6.5.2. Min geometric generalized Pareto distribution

Definition 6.5.2. A random variable X on $(0,\infty)$ is said to follow generalized Pareto distribution if its survival function $\bar{F}_X(x)$ is of the form

$$\bar{F}_X(x) = \frac{1}{(1+x^\alpha)^\beta},\ \alpha > 0,\ \beta > 0. \tag{6.5.1}$$

It may be denoted by $GEPD(\alpha,\beta)$.

Definition 6.5.3. A random variable X on $(0,\infty)$ is said to follow min geometric generalized Pareto distribution and write $X \stackrel{d}{=} MGGEP(\alpha,\beta)$ if its survival function $\bar{F}_X(x)$ is of the form

$$\bar{F}_X(x) = \frac{1}{1+\beta\ln(1+x^\alpha)},\ \alpha > 0, \beta > 0. \tag{6.5.2}$$

Remark 6.5.1. Obviously $MGGEP(\alpha,\beta)$ distribution belongs to the class of min GID distributions and $MGGEP(\alpha,1)$ is the min geometric Pareto distribution.

Theorem 6.5.1. *Let $X_1, X_2, \ldots$ be i.i.d $MGP(\alpha)$ random variables and let $N(\theta)$ be geometric with mean $\frac{1}{\theta}$ such that $P[N(\theta)=k] = \theta(1-\theta)^{k-1},\ k = 1,2,\ldots 0 < \theta <$ 1. Define $Y = \Lambda_{j=1}^{N(\theta)} X_j$. Then $Y \stackrel{d}{=} MGGEP\left(\alpha, \frac{1}{\theta}\right)$.*

Proof 6.5.1.

$$\begin{aligned}\bar{F}_Y(x) &= P[\Lambda_{j=1}^{N(\theta)} X_j \geq x] = \sum_{k=1}^{\infty} P[\Lambda_{j=1}^{N(\theta)} X_j \geq x]\theta(1-\theta)^{k-1} \\ &= \frac{1}{1+\frac{1}{\theta}\log(1+x^\alpha)}\end{aligned}$$

and hence $Y \stackrel{d}{=} MGGEP(\alpha, \frac{1}{\theta})$.

To prove the Theorem 3.2 we require the following Lemma.

Lemma 6.5.1. *Let $\psi(x) = \left[\frac{1}{F(x)} - 1\right]$. Then $\frac{1}{1+a\psi(x)}$ is the survival function of geometric minimum for any $a > 1$.*

Proof 6.5.2.

$$\begin{aligned}[1+a\psi(x)]^{-1} &= \frac{1}{a}\left[\frac{1}{a}+\psi(x)\right]^{-1} = \frac{1}{a}\left[\frac{1}{a}+\left[\frac{1}{\bar{F}(x)}-1\right]\right]^{-1} \\ &= \frac{1}{a}\bar{F}(x)\left[1-\left(1-\frac{1}{a}\right)\bar{F}(x)\right]^{-1} = \sum_{k=1}^{\infty}\frac{1}{a}\left(1-\frac{1}{a}\right)^{k-1}[\bar{F}(x)]^{k} \\ &= P[\Lambda_{j=1}^{N(\theta)}X_j \geq x]\end{aligned}$$

where $N(\theta)$ is a geometric random variable with $E[N(\theta)] = \frac{1}{\theta}$; X_j's are i.i.d. with survival function $[1+\psi(x)]^{-1}$, $j = 1, 2, \ldots N(\theta)$. Now we consider the following theorem:

Theorem 6.5.2. *Generalized min geometric Pareto distribution is the limit distribution of geometric minimum of generalized Pareto GEP$\left(\alpha, \frac{\beta}{n}\right)$ variables.*

Proof 6.5.3. $[1+x^{\alpha}]^{-\frac{\beta}{n}} = [1+[1+x^{\alpha}]^{\frac{\beta}{n}}-1]^{-1}$ is the survival function of a probability distribution. Since generalized Pareto distribution is min ID. Therefore by Lemma 6.5.1, we have

$$\bar{F}_n(t) = \{1+n[(1+x^{\alpha})^{\frac{\beta}{n}}-1]\}^{-1} \tag{6.5.3}$$

is the survival function of geometric minimum of i.i.d random variables distributed as generalized Pareto $\left(\alpha, \frac{\beta}{n}\right)$ random variables. Taking the limit as $n \to \infty$ on both sides of (6.5.3),

$$\begin{aligned}\bar{F}(t) &= \lim_{n\to\infty}\bar{F}_n(t) = [1+\lim_{n\to\infty}[n(1+x^{\alpha})^{\frac{\beta}{n}}-1]^{-1} \\ &= [1+\beta\ln(1+x^{a})]^{-1}.\end{aligned}$$

In the next section we can extent the idea of min geometric generalized Pareto distribution to a larger class called min geometric generalized semi-Pareto distribution.

6.5.3. Min geometric generalized semi-Pareto distribution

Definition 6.5.4. A random variable X with positive support has semi-Pareto distribution SP (α, p) if its survival function is given by $\bar{F}_X(x) = \frac{1}{1+\psi(x)}$, where $\psi(x)$ satisfies the functional equation

$$p\psi(x) = \psi(p^{\frac{1}{\alpha}}x) \; 0 < p < 1, \; \alpha > 0. \tag{6.5.4}$$

This definition is analogous to the semi-stable law defined by Lev́y (see Pillai (1971)). It can be shown that $\psi(x) = x^{\alpha}h(x)$ where $h(x)$ is periodic $\ln x$ with period $\frac{-2\pi\alpha}{\ln p}$. (For proof see Kagan et al (1973)). The semi-Pareto distribution can be viewed as a more general class which includes the Pareto type III distribution when $\psi(x) = cx^{\alpha}$. The semi-Pareto distribution was introduced by Pillai (1991).

Definition 6.5.5. A random variable X on $(0, \infty)$ is said to follow generalized semi Pareto distribution if its survival function $\bar{F}_X(x)$ is of the form $\bar{F}_X(x) = \frac{1}{[1+\psi(x)]^k}$, $k > 0$ where $\psi(x)$ satisfies (6.5.4). It may be denoted by $GESP(\alpha, p, k)$.

Definition 6.5.6. A random variable X on $(0, \infty)$ is said to follow generalized min geometric semi Pareto distribution write $X \stackrel{d}{=} MGGESP(\alpha, p, k)$ if its survival function $\bar{F}(x)$ is of the form

$$\bar{F}(x) = \frac{1}{1 + k\ln(1 + \psi(x))}, \; k > 0 \tag{6.5.5}$$

where $\psi(x)$ satisfies (6.5.4).

Remark 6.5.2. $MGGESP(\alpha, p, k)$ distribution belongs to the class of min GID distributions.

Remark 6.5.3. If $k = 1$ this distribution is called min geometric semi-Pareto distribution and shall be denoted by $MGSP(\alpha, p)$.

Theorem 6.5.3. *Let $X_1, X_2, \ldots$ be identically and independently distributed MGSP (α, p) random variables and $N(\theta)$ be geometric with mean $\frac{1}{\theta}$ such that $P[N(\theta) = k] = \theta(1-\theta)^{k-1}, k = 1, 2, \ldots 0 < \theta < 1$. Define $Y = \Lambda_{j=1}^{N(\theta)} X_j$. Then $Y \stackrel{d}{=} MGGESP (\alpha, p, \frac{1}{\theta})$.*

Theorem 6.5.4. *The generalized min geometric semi-Pareto distribution is the limit distribution of geometric minimum of generalized semi Pareto GESP$\left(\alpha, p, \frac{k}{n}\right)$ variables.*

6.5.4. First order autoregressive minification process

The analysis of time series in the classical set up is based on the assumption that an observed time series is a realization from a Gaussian sequence. But there are many situations where the naturally occurring series shows a tendency to heavy tailed or asymmetric, so that this series cannot be modeled using Gaussian distributions. Hence a number of non-Gaussian time series models have been introduced by different researchers during last two decades.

The study of autoregressive minification process began with the pioneering work of Tavares (1980). In his work the observations are generated by the equation

$$X_n = k \min(X_{n-1}, \varepsilon_n), n \geq 0,$$

where $k > 1$ is a constant and $\{\varepsilon_n\}$ is an innovation process of independently and identically distributed random variables chosen to ensure that $\{X_n\}$ is a stationary Markov process with marginal distribution function $\bar{F}_{X_0}$. Because of its structure, the process $\{X_n\}$ is called minification process. Yeh et al (1988) developed an autoregressive minification process with Pareto marginal distribution and having the form

$$X_n = \begin{cases} kX_{n-1} & \text{with probability} \quad p \\ k\min(X_{n-1}, \varepsilon_n) & \text{with probability} \quad 1-p \end{cases}.$$

Pillai (1991) generalized this to obtain the semi Pareto processes. Alice and Jose (2003, 2004) developed minification process with Marshall-Olkin Pareto marginal distributions. These minification processes are found to be useful for modeling in various fields such as hydrological studies, reliability studies etc. In this section we introduce a new minification process $\{X_n\}$ having structure given by

$$X_n = \begin{cases} \varepsilon_n & \text{with probability} \quad \rho \\ \min(X_{n-1}, \varepsilon_n) & \text{with probability} \quad 1-\rho \end{cases} ; 0 < \rho < 1 \tag{6.5.6}$$

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. In terms of survival function, (6.5.6) can be rewritten as

$$\bar{F}_{X_n}(x) = \rho \bar{F}_{\varepsilon_n}(x) + (1-\rho) F_{X_{n-1}}(x) \bar{F}_{\varepsilon_n}.$$

On assuming stationarity, it reduces to the form

$$\bar{F}_{\varepsilon}(x) = \frac{\bar{F}_X(x)}{\rho + (1-\rho)\bar{F}_X(x)}. \tag{6.5.7}$$

6.5.5. Semi-Pareto autoregressive minification process

Consider an autoregressive minification structure given by (6.5.6). Substituting

$$\bar{F}_X(x) = \frac{1}{1 + k\ln(1 + \psi(x))} \text{ in (6.5.7)}$$

we get $\bar{F}_\varepsilon(x) = \frac{1}{1+k\rho\ln(1+\psi(x))}$. Hence it follows that $\varepsilon_n \stackrel{d}{=} GEMGSP(\alpha, p, k\rho)$. If ε_n's are i.i.d $MGGESP(\alpha, p, \rho)$, then

$$\begin{aligned} \bar{F}_X(x) &= \frac{\rho}{1 + k\rho\ln(1 + \psi(x))} + \frac{1-\rho}{1 + k\rho\ln(1 + \psi(x))}\frac{1}{1 + \ln(1 + \psi(x))} \\ &= \frac{1}{1 + k\ln(1 + \psi(x))}. \end{aligned}$$

If $X_{n-1} \stackrel{d}{=} MGSP(\alpha, p)$ then we get $X_n \stackrel{d}{=} MGSP(\alpha, p)$. Hence $\{X_n\}$ is strictly stationary with min geometric semi Pareto marginals. Hence we have the following theorem.

Theorem 6.5.5. *Let $\{X_n\}$ be an autoregressive process with minification structure given by (6.5.3) where $\{\varepsilon_n\}$ is a sequence of i.i.d random variables independent of X_n. Then $\{X_n\}$ is a stationary Markovian ARM(1) process with min geometric semi Pareto marginals if and only if $\{\varepsilon_n\}$'s are distributed as GEMSP(α, p, ρ) distribution.*

Proof 6.5.4. Consider the structure as given in (6.5.6). Substituting

$$\bar{F}_X(x) = \frac{1}{1 + \ln(1 + \psi(x))}$$

in (6.5.6), we get

$$\bar{F}_\varepsilon(x) = \frac{1}{1 + \rho\ln(1 + \psi(x))}.$$

Hence it follows that $\varepsilon_n \stackrel{d}{=} GEMGSP(\alpha, p, \rho)$. If ε_n's are i.i.d$MGSP(\alpha)$, then

$$\begin{aligned} \bar{F}_X(x) &= \frac{\rho}{1 + \rho\ln(1 + \psi(x))} + \frac{1-\rho}{1 + \rho\ln(1 + \psi(x))}\frac{1}{1 + \ln(1 + \psi(x))} \\ &= \frac{1}{1 + \ln(1 + \psi(x))}. \end{aligned}$$

If $X_{n-1} \stackrel{d}{=} MGSP(\alpha, p)$ then we get $X_n \stackrel{d}{=} MGSP(\alpha, p)$. Hence $\{X_n\}$ is strictly stationary with min geometric semi-Pareto marginals. Hence the theorem.

6.5.6. Extension to higher order processes

The first order min geometric semi-Pareto autoregressive minification model discussed above can be extended to a kth order min geometric semi-Pareto autoregressive process with minification structure denoted by $MGSPARM$ (k). Let $\{\varepsilon_n\}$ be a sequence of independently and identically distributed $MGSPD(\alpha)$ variables. Then $MGSPARM(k)$ process can be constructed by $\{X_n\}$ satisfying the structural relationship

$$X_n = \begin{cases} \varepsilon_n & \text{with probability} & \rho \\ \min(X_{n-1}, \varepsilon_n) & \text{with probability} & \rho_1 \\ \min(X_{n-2}, \varepsilon_n) & \text{with probability} & \rho_2 \\ \vdots & \vdots & \vdots \\ \min(X_{n-k}, \varepsilon_n) & \text{with probability} & \rho_k \end{cases}$$

$0 < \rho < 1,\ \rho_1 + \rho_2 + \cdots + \rho_k = 1 - \rho$. Then $\{X_n\}$ has stationary marginal distribution as $MGSPD(\alpha)$ if and only if $\{\varepsilon_n\}$ is distributed as $GEMGSPD(\alpha, k)$ distributions.

6.6. Case Study-Applications in Financial Modelling

6.6.1. Generalized Pareto distribution

In this section we conduct a more detailed study on the properties of the generalized Pareto distribution given by (6.5.1) and denoted by $GEPD(\alpha,\beta)$. The density function of $GEPD(\alpha,\beta)$ is given by

$$f(x) = \frac{\alpha\beta x^{\alpha-1}}{(1 + x^{\alpha})^{\beta+1}},\ x > 0$$

A plot of the generalized Pareto distribution for different values of α and β is presented in Figure 6.6.1

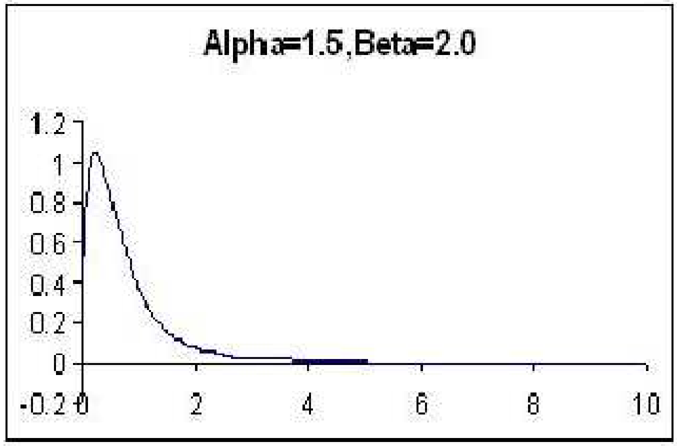

Figure 6.6.1

Certain properties of this distribution are given below

$$E(X^s) = \frac{\Gamma\left(1+\frac{s}{\alpha}\right)\Gamma\left(\beta-\frac{s}{\alpha}\right)}{\Gamma(\beta)}$$

$$E(X) = \frac{\Gamma\left(1+\frac{1}{\alpha}\right)\Gamma\left(\beta-\frac{1}{\alpha}\right)}{\Gamma(\beta)}$$

$$V(X) = \frac{\left[\Gamma\left(\frac{\alpha+2}{\alpha}\right)\Gamma\left(\frac{\alpha\beta-2}{\alpha}\right)\Gamma(\beta)-\left(\Gamma\left(\frac{\alpha+1}{\alpha}\right)\right)^2\left(\Gamma\left(\frac{\alpha\beta-1}{\alpha}\right)\right)^2\right]}{(\Gamma(\beta))^2}$$

$$\text{Median }(X) = \left[\left[\frac{1}{2}\right]^{-\frac{1}{\beta}}-1\right]^{\frac{1}{\alpha}}$$

$$\text{Mode }(X) = \begin{cases} \left[\frac{\alpha-1}{\alpha\beta+1}\right]^{\frac{1}{\alpha}} & \text{if } \alpha > 1 \\ 0 & \text{otherwise.} \end{cases}$$

6.6.2. An application

Daily observations of U.S. dollar-Indian rupee foreign exchange rate are considered. The data consists of 1897 observations starting from 05-04-1999 to 15-12-2006. The data are collected from the website of Reserve Bank of India. Time series plot of the data is provided in the Figure 6.6.2

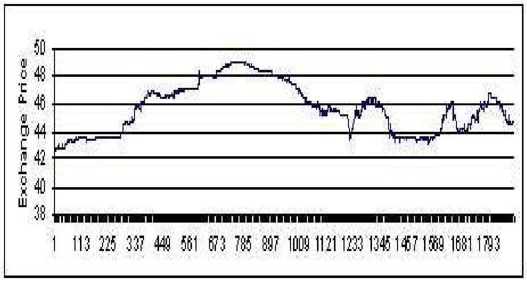

Figure 6.6.2 Time series of daily exchange rate of US dollar-Indian rupee

From the graph it can be seen that the exchange rate shows an increase initially, then decrease and then a slow increase. The first order autocorrelation of the series $\{X_n\}$ is obtained as $\rho_0 = 0.998665079$. To make the series stationary, by taking the first order autocorrelated difference of the $\{X_n\}$ is taken. The new series obtained is referred to as $\{Y_n\}$ where $Y_n = X_n - \rho_0 X_{n-1}$. This series is standardized by subtracting its mean and dividing by its standard deviation. The resulting series is made positive by considering its modulus. The autocorrelation of the resulting series is insignificant. Each observation in the series is multiplied by 10. The maximum value of the series is found to be 74.90.The observations are classified into 50 classes of equal width (1.5). Histogram is constructed with mid values of the classes along X-axis and normed frequency along Y- axis. A plot of the histogram is given in the following Figure 6.6.3

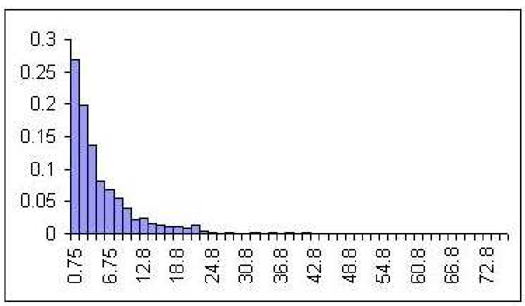

Figure 6.6.3 Histogram of the observed series

The histogram resembles the shape of generalized Pareto distribution presented in Figure 6.6.1. Therefore, we may assume that generalized Pareto distribution may be a good fit to the data set considered. The method used in this procedure is as follows. Simulate 10000 observations from generalized Pareto distribution with different values of α and β. Construct frequency polygon of the simulated data using the same class interval used in the observed data plot. Impose this polygon on the histogram of the data. Adjust the values of α and β so that both figures coincide at least approximately. The best estimates of α and β are those values of α and β for which the figures coincide. The following Figure 6.6.4 represents the histogram of the observed data and embedded generalized Pareto frequency polygon.

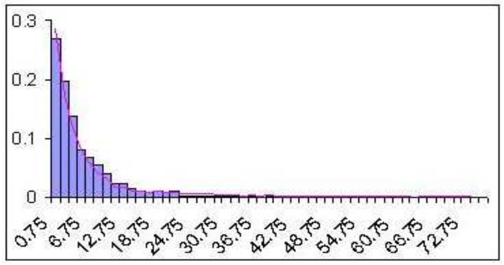

Figure 6.6.4 The histogram and embedded frequency polygon

From the figures we can observe that the data is a good fit for generalized Pareto distribution with $\alpha = 1.63$ and $\beta = 0.3278$.

Exercises 6.6.

6.6.1. Obtain a necessary and sufficient condition that $X_n = a \min(X_{n-1}, \varepsilon_n)$ define a proper $AR(1)$ process.

6.6.2. Develop a minification process with exponential marginal distribution.

6.6.3. Obtain the innovation structure of a minification process with Weibull marginal distribution.

6.6.4. Derive the innovation structure of a semi-Parato minification process.

6.6.5. Examine whether you can develop a minification process with logistic marginal distribution.

6.6.6. What is meant by min geometric infinite divisibility? Give two examples.

6.6.7. Develop the concept of max geometric infinite divisibility. Obtain necessary and sufficient conditions for its validity.

6.6.8. Describe a generalized logistic distribution. Construct an $AR(1)$ process with minification structure and generalized logistic marginal distribution.

6.6.9. Derive the reliability characteristics of a Marshall-Olkin exponential distribution.

6.6.10. Develop a Marshall-Olkin Weibull distribution. Derive the hazard rate and interpret its behavior.

6.6.11. Develop a maximal process with structure $X_n = a \max(X_{n-1}, \varepsilon_n)$ with Burr marginals.

References

Alice Thomas and Jose, K.K. (2003). Marshall-Olkin Pareto Processes, *Far East J. Theo. Stat.*, **9(2)**, 117-132.

Alice Thomas and Jose, K.K. (2004). Bivariate semi-Pareto minification Processes, *Metrika*, **59**, 305-313.

Anderson, T.W. (1971). *The Statistical Analysis of Time Series*, Wiley , New York.

Arnold, B. C. (1983). *Pareto Distributions*, International Cooperative Publishing House, Fairland, Maryland.

Box, G.E.P. and Jenkins, G.M. (1970). *Time Series Analysis: Forcasting and Control*, Holden Day, London.

Erdélyi, A., et al (1955). *Higher Transcendental Functions*, **3**, McGraw Hill, New York.

Fujita, Y. (1993). A generalization of the results of Pillai, *Ann. Inst. Statist. Math.*, **45(2)**, 361-365.

Gaver, D.P. and Lewis, P.A.W. (1980). First-order autoregressive gamma sequences and point processes, *Adv. Appl. Prob.*, **12**, 727-745.

Haubold, H.J. and Mathai, A.M. (2000). The Fractional Kinetic Equation and Thermonuclear Functions, *Astrophysics and Space Science*, **273**, 53-63.

Jayakumar, K. and Pillai, R.N. (1993). The first-order autoregressive Mittag-Leffler process, *J. Appl. Prob*, **30**, 462-466.

Jayakumar, K., Kalyanaraman, K. and Pillai, R.N. (1995). α-Laplace Processes, *Mathl. Comput. Modelling* , **22**(1), 109-116.

Jayakumar, K. and Ajitha, B. K. (2003). On the geometric Mittag-Leffler distributions, *Cal. Statist. Assoc. Bull.*, **54**, Nos. 215-216, 195-208.

Jose, K.K. and Seetha Lekshmi, V. (1999). On geometric exponential distribution and its applications, *J. Ind. Statist. Assoc.*, **37**, 51-58.

Kagan, A.M., Linnik, Yu, V. and Rao, C.R., (1973) *Characterization Problems in Mathematical Statistics*, Wiley, New York.

Klebanov, L.B., Maniya, G.M. and Melamed, I.A., (1984) A problem of Zolotarev and analogs of infinitely divisible and stable distribution in a scheme for summing a random number of random variables, *Theory Probab. Appl.*, **29**, 791-794.

Kozubowski, T.J and Rachev, S.T., (1999) Univariate geometric stable laws, *J. Comp. Anal. Appl.* (Preprint).

Lawrance, A. J. and Lewis, P. A. W. (1982). A mixed time series exponential model, *Management Science*, **28(9)**, 1045-1053.

Lin, G.D. (1998a). A note on the Linnik distributions, *J. Math. Anal. Appl.*, **217**, 701-706.

Lin, G.D. (1998b). On the Mittag-Leffler distributions, *J. Statist. Plann. Inference*, **74**, 1-9.

Mathai, A.M. (1993). On non-central generalized Laplacianness of quadratic forms in normal variables, *J. Multivariate Anal.*, **45(2)**, 239-246.

McIntire, D. (1977). A new approach to ARMA modelling, Ph.D. Thesis, Southern Methodist University, USA.

Mohan, N.R., Vasudeva, R. and Hebbar, H.V., (1993) On geometrically infinitely divisible laws and geometric domains of attraction, *Sankhya*, **55 A, 2**, 171-179.

Pakes, A.G. (1998). Mixture representations for symmetric generalized Linnik laws, *Statist. Prob. Letts.*, **37**, 213-221.

Papoulis, A. (1985). *Probability, Random Variables and Stochastic Processes*, Second Edition, McGraw-Hill, New York.

Parzen, E. (1961). *Time Series Analysis Papers*, Holden Day, London.

Pillai, R.N. (1985). Semi α-Laplace distribution. *Comm.Statist.-Theory and Methods*, **14(4)**, 991-1000.

Pillai, R.N. (1990a). On Mittag-Leffler and related distributions, *Ann. Inst. Statist. Math.*, **42(1)**, 157-161.

Pillai, R.N. (1990b). Harmonic mixtures and geometric infinite divisibility, *J. Ind. Statist. Assoc.* , **28**, 87-98.

Pillai, R. N. (1991) Semi Pareto Processes. *Journal of Applied Probability,* **28**, 461-465.

Pillai, R.N. and Jayakumar, K. (1995). Discrete Mittag-Leffler distributions, *Statist. Prob. Letts.*, **23**, 271-274.

Pillai, R.N. and Sandhya, E. (1990). Distributions with complete monotone derivative and geometric infinite divisibility, *Adv. Appl. Prob.* , **22**, 751-754.

Rachev, S.T. and Resnick, S. (1991). Max-geometric infinite divisibility and stability. *Commun. Statist. Stochastic Models.*, **7(2)**, 191-218.

Saxena, R.K., Mathai, A.M. and Haubold, H.J. (2004a). Unified fractional kinetic equation and a fractional diffusion equation, *Astrophysics and Space Science*, **209**, 299-310.

Saxena, R.K., Mathai, A.M. and Haubold, H.J. (2004b). On generalized fractional kinetic equations, *Physica A*, **344**, 657-664.

Seetha Lekshmi, V. and Jose, K.K. (2002) Geometric Mittag-Leffler tailed autoregressive processes, *Far East J. Theo. Stat.*, **6(2)**, 147-153.

Seetha Lekshmi, V., Joy Jacob and Jose, K.K. (2003(a)). Generalized Laplacian and geometric α-Laplace distributions with applications in time series modeling, *Statistical Methods*, **5(2)**, 140-155.

Seetha Lekshmi, V. and Jose, K.K. (2003(b)). Autoregressive models in geometric exponential tailed marginal distributions, *Journal of Statistical. Studies*, **23**, 33-37.

Seetha Lekshmi, V. and Jose, K.K. (2004(a)). An autoregressive process with geometric α-Laplace marginals, *Statistical Papers*, **45**, 337-350.

Seetha Lekshmi, V. and Jose, K.K., (2004(b)). Geometric Mittag-Leffler distributions and processes, *Journal of Applied Statistical Science*, **13(4)**, 335-342.

Seetha Lekshmi, V., Johny Scaria and Jose, K.K. (2005). On min geometric stable distributions and their applications, *Recent Advances in Statistical Theory and Applications*, **1**, 183-196.

Seetha Lekshmi, V. and Jose, K.K. (2006). Autoregressive processes with Pakes and geometric Pakes generalized Linnik marginals. *Statist. Prob.Letters*. **76**, 318-326.

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

Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics, *J. Statist. Phy.* , **52**, Nos.1-2, 479-487.

Weron, K. and Kotulski, M. (1996). On the Cole-Cole relaxation function and related Mittag-Leffler distribution, *Physica A*, **232**, 180-188.

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