CHAPTER 2

Integer Valued Autoregressive Processes with Generalized Discrete Mittag-Leffler Marginals

2.1 Introduction

The function $E_\alpha(z) = \sum_{k=0}^{\infty} \left[\frac{z^k}{\Gamma(1+k\alpha)} \right]$ was first introduced by Mittag-Leffler in 1903 (Erdelyi,1955). Many properties of the function follow from Mittag-Leffler integral representation

$$E_\alpha(z) = \frac{1}{2\pi i} \int_C \frac{t^{\alpha-1} e^t}{t^\alpha - z} dt$$

Some results included in this chapter form part of the papers Jose and Mariyamma (2012a) and Jose and Mariyamma (2012b).

where the path of integration C is a loop which starts and ends at $-\infty$ and encircles the circular disc $|t| \leq z^{\frac{1}{\alpha}}$. The popularity of Mittag-Leffler function has considerably increased among engineers and scientists due to its vast potential for applications in several areas, such as fluid flow, diffusive transport, electric networks, probability theory, statistical distribution theory etc. Prabhakar (1971) developed a generalization of Mittag-Leffler function. 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. They obtained the solution of a unified form of generalized fractional kinetic equations, which provides a unification and extension of the earlier results. Such behaviors occur frequently in Chemistry, Thermodynamical and Statistical analysis. In all such situations the solutions can be expressed in terms of generalized Mittag-Leffler functions.

Pillai (1990) proved that $F_\alpha(x) = 1 - E_\alpha(-x^\alpha),\ \ 0 < \alpha \leq 1$ are distribution functions, having the Laplace transform $\phi(t) = (1 + t^\alpha)^{-1},\ \ t \geq 0$ 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. Jayakumar and Pillai (1993) developed a first order autoregressive process with Mittag-Leffler marginal distribution. Weron and Kotulski (1996) use Mittag-Leffler distribution in explaining Cole-Cole relaxation. 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. Jose and Abraham (2011) developed new count models with Mittag-Leffler waiting times as generalization of Poisson process. A discrete version of the Mittag-Leffler distribution was introduced by Pillai and Jayakumar (1995).

A first order autoregressive model for count (or integer valued) data is developed through the thinning operator $*$ which is due to Steutel and van Harn (1979). Let X be an

$\mathbb{Z}_+$-valued random variable and $\gamma \in (0,1)$, then the thinning operator '$*$' is defined by

$$\gamma * X = \sum_{i=1}^{X} V_i \tag{2.1.1}$$

where V_i's are *i.i.d.* Bernoulli random variables with $P(V_i = 1) = 1 - P(V_i = 0) = \gamma$, and are independent of X. If $G_X(s) = \sum_{j=0}^{\infty} P[X = j]s^j = E[s^X]$ represents the probability generating function (pgf) of X, then the pgf of $\gamma * X$ is obtained as $G_X(1-\gamma+\gamma s)$.

A sequence $\{X_n, n \in \mathbb{Z}\}$ of $\mathbb{Z}_+$-valued random variables is said to be an integer-valued first order autoregressive (INAR(1)) process if for any $n \in \mathbb{Z}$,

$$X_n = \gamma * X_{n-1} + \epsilon_n \tag{2.1.2}$$

where $\gamma \in (0,1)$ is the first order autocorrelation coefficient of the process and ϵ_n is the innovation process. Under the assumption of strict stationarity, (2.1.2) can be rewritten in terms of *pgf* as

$$G(s) = G(1-\gamma+\gamma s)G_\epsilon(s),\ |s| \leq 1,\ \gamma \in (0,1) \tag{2.1.3}$$

where $G_\epsilon(s)$ is a proper *pgf*.

McKenzie (1986) introduced a class of discrete valued sequences with negative binomial and geometric marginal distributions obtained as discrete analogues of the standard autoregressive time series models of Lawrance and Lewis (1980), replacing the scalar multiplication by the thinning operation. We consider the alternate probability generating function (*apgf*) defined as A(s) = G(1-s) = $E[(1-s)^X]$ instead of the *pgf* in (2.1.2), which yields an expression analogous to the Laplace transform for positive valued continuous random variables, so that (2.1.3) can be rewritten as

$$A(s) = A(\gamma s)A_\epsilon(s) \tag{2.1.4}$$

for every $\gamma \in (0,1)$. The above equation (2.1.4), is analogous to the definition of self-decomposability for continuous random variables.

In this chapter, we consider a generalization of discrete Mittag-Leffler distributions. We introduce and study the properties of a new distribution called geometric generalized discrete Mittag-Leffler distribution. Autoregressive processes with geometric generalized discrete Mittag-Leffler distributions are developed and studied. The distributions are further extended to develop a more general class of geometric generalized discrete semi-Mittag-Leffler distributions. The processes are extended to higher orders also. Autoregressive processes with bivariate discrete semi Mittag-Leffler marginals are developed and their generalization are also studied. An application with respect to an empirical data on customer arrivals in a bank counter is also given. Various areas of potential applications like human resource development, insect growth, epidemic modeling, industrial risk modeling, insurance and actuaries, town planning etc. are also discussed.

2.2 Generalized Discrete Mittag-Leffler distribution

Definition 2.2.1. *A random variable X on $\mathbb{Z}_+$ is said to follow generalized discrete Mittag-Leffler distribution denoted by GDML(α, c, β), if it has the pgf*

$$G(s) = \left\{\frac{1}{1+c(1-s)^{\alpha}}\right\}^{\beta};\ |s| \le 1, 0<\alpha\le 1, c>0, \beta>0. \tag{2.2.1}$$

For $\beta = 1$, it is the DML(α). When $\alpha = 1$, $\beta = 1$ and $c = \frac{q}{p}$ where $q = 1-p$, it reduces to geometric distribution.

Theorem 2.2.1. *The $GDML(\alpha, c, \beta)$ distribution is discrete self-decomposable (or discrete class L).*

Proof: From (2.1.3) the *pgf* of $GDML(\alpha, c, \beta)$ is

$$\begin{aligned} G(s) &= \left[\frac{1}{1+c\gamma^{\alpha}(1-s)^{\alpha}}\right]^{\beta}\left[\gamma^{\alpha}+(1-\gamma^{\alpha})\frac{1}{1+c(1-s)^{\alpha}}\right]^{\beta} \\ &= G(1-\gamma+\gamma s)G_{\gamma}(s). \end{aligned}$$

Theorem 2.2.2. *The $GDML(\alpha, c, \beta)$ distribution is geometrically infinitely divisible and hence infinitely divisible.*

Proof: Consider the *pgf* of $GDML(\alpha, c, \beta)$

$$G(s) = \left[\frac{1}{1 + c(1-s)^{\alpha}}\right]^{\beta}$$

Using the criterion used in Pillai and Sandhya (1990), we see that the $GDML(\alpha, c, \beta)$ is geometrically infinitely divisible.

Theorem 2.2.3. *Let $G(s)$ be the pgf of a GDML distribution with $\gamma \in (0,1), c > 0, \beta > 0, 0 < \alpha \leq 1, |s| \leq 1$. Then there exists a stationary INAR(1) process $\{X_n, n \in \mathbb{Z}\}$, having structure given by (2.1.2) with $G(s)$ as the pgf of its marginal distribution. Also the marginal distribution of the innovation sequences $\{\epsilon_n, n \in \mathbb{Z}\}$ has apgf $A_{\epsilon}(s)$ given by*

$$A_{\epsilon}(s) = \left\{\frac{1 + c\gamma^{\alpha}s^{\alpha}}{1 + cs^{\alpha}}\right\}^{\beta}. \tag{2.2.2}$$

Proof: In terms of *apgf* defined as $A(s) = G(1-s)$, the INAR(1) model defined in (2.1.2) can be rewritten as

$$A_{X_n}(s) = A_{X_{n-1}}(\gamma s)A_{\epsilon_n}(s)$$

Under stationarity it reduces to

$$A_X(s) = A_X(\gamma s)A_{\epsilon}(s)$$

Hence

$$A_{\epsilon}(s) = \frac{A_X(s)}{A_X(\gamma s)}$$

The INAR(1) with GDML marginals is defined, only if there exists an innovation sequence $\{\epsilon_n\}$ such that $A_{\epsilon}(s)$ is an *apgf*.

From (2.2.1)

$$A_X(s) = \left\{ \frac{1}{1 + cs^\alpha} \right\}^\beta$$

Then we have,

$$\begin{aligned} A_\epsilon(s) &= \left\{ \frac{1 + c\gamma^\alpha s^\alpha}{1 + cs^\alpha} \right\}^\beta \\ &= \left\{ \gamma^\alpha + (1 - \gamma^\alpha)\frac{1}{1 + cs^\alpha} \right\}^\beta \end{aligned}$$

Therefore, the innovations ϵ_n are β-fold zero-inflated convolutions of DML(α, c, β).

2.2.1 Joint Distribution of (X_{n-1}, X_n)

The joint *pgf* of (X_{n-1}, X_n) is given by

$$\begin{aligned} G_{X_{n-1},X_n}(s_1, s_2) &= E[s_1^{X_{n-1}} s_2^{\gamma * X_{n-1} + \epsilon_n}] \\ &= G_{\epsilon_n}(s_2) G_{X_{n-1}}[s_1(1 - \gamma + \gamma s_2)] \\ &= \left[\frac{1 + c\gamma^\alpha s_2^\alpha}{1 + cs_2^\alpha} \right]^\beta \left[\frac{1}{1 + c\{s_1(1 - \gamma + \gamma s_2)\}} \right]^\beta \end{aligned}$$

By inverting this expression the joint distribution can be obtained. The above expression is not symmetric in s_1 and s_2 and hence the process is not time reversible.

2.3 Geometric Generalized Discrete Mittag-Leffler Distribution

Jose et al. (2010) introduced and studied the geometric generalized Mittag-Leffler distribution and its properties. Now we shall introduce its discrete analogue as follows.

Definition 2.3.1. *A random variable X on $\mathbb{Z}_+$ is said to follow geometric generalized discrete Mittag-Leffler distribution and write $X \overset{d}{=} GGDML(\alpha, c, \beta)$, if it has the apgf,*

$$A(s) = \frac{1}{1 + \beta \ln[1 + cs^\alpha]}; \ |s| \le 1, 0 < \alpha \le 1, c > 0, \beta > 0 \qquad (2.3.1)$$

Remark 2.3.1. *The GGDML distribution is geometrically infinitely divisible.*

Theorem 2.3.1. *Let $X_1, X_2, \cdots,$ be iid GGDML random variables and $Y = X_1 + X_2 + \cdots + X_{N(p)}$ where $N(p)$ follows geometric distribution with pdf, $P[N(p) = k] = p(1-p)^{k-1},\ k = 1, 2, \cdots, 0 < p < 1$. Then $Y \stackrel{d}{=} GGDML(\alpha, c, \frac{\beta}{p})$.*

Proof: Taking the *apgf* of Y we have,

$$\begin{aligned} A_Y(s) &= \sum_{k=1}^{\infty} \{A_X(s)\}^k p(1-p)^{k-1} \\ &= \frac{1}{1 + \frac{\beta}{p} \ln[1 + cs^{\alpha}]}. \end{aligned}$$

2.3.1 Geometric Generalized Discrete Mittag-Leffler Processes

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

Theorem 2.3.2. *Let $\{X_n, n \geq 1\}$ be defined as*

$$X_n = \begin{cases} \epsilon_n, & \text{with probability} \quad p \\ X_{n-1} + \epsilon_n, & \text{with probability} \quad (1-p) \end{cases} \tag{2.3.2}$$

where $\{\epsilon_n\}$ is a sequence of iid random variables. A necessary and sufficient condition that $\{X_n\}$ is a strictly stationary Markov process with $GGDML(\alpha, c, \beta)$ marginals is that ϵ_n are distributed as geometric Mittag-Leffler provided X_0 is distributed as geometric generalized discrete Mittag-Leffler.

Proof: Rewriting (2.3.2) in terms of *apgf* we have,

$$A_{X_n}(s) = pA_{\epsilon_n}(s) + (1-p)A_{X_{n-1}}(s)A_{\epsilon_n}(s). \tag{2.3.3}$$

Assuming strict stationarity, it becomes,

$$A_X(s) = A_{\epsilon}(s)\{p + (1-p)A_X(s)\}.$$

That is,

$$A_\epsilon(s) = \frac{A_X(s)}{p + (1-p)A_X(s)}$$

where

$$A_X(s) = \frac{1}{1 + \beta \ln[1 + cs^\alpha]}.$$

On simplification we get,

$$A_\epsilon(s) = \frac{1}{1 + p\beta \ln[1 + cs^\alpha]}.$$

and hence $\epsilon_n \stackrel{d}{=} GGDML(\alpha, c, p\beta)$.

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

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

The rest follows similarly.

2.3.2 INAR(p) Process with GGDML Marginal Distribution

Now we consider a p^{th} order integer-valued autoregressive (INAR(p)) process with probability structure,

$$X_n = \begin{cases} \gamma_1 * X_{n-1} + \epsilon_n, & \text{with probability} \quad \delta_1 \\ \gamma_2 * X_{n-2} + \epsilon_n, & \text{with probability} \quad \delta_2 \\ \cdots & \cdots \quad \cdots \\ \cdots & \cdots \quad \cdots \\ \gamma_p * X_{n-p} + \epsilon_n, & \text{with probability} \quad \delta_p \end{cases} \tag{2.3.4}$$

where $0 < \gamma_i, \delta_i \leq 1,\ i = 1, 2, \cdots, p;\ \sum_{i=1}^{p} \delta_i = 1$.

In terms of *apgf* the above equation can be written as

$$A_{X_n}(s) = A_{\epsilon_n}(s) \sum_{i=1}^{p} \delta_i A_{X_{n-i}}(\gamma_i s).$$

Assuming strict stationarity it reduces to

$$A_X(s) = A_\epsilon(s) \sum_{i=1}^{p} \delta_i A_X(\gamma_i s).$$

Hence

$$A_\epsilon(s) = \frac{A_X(s)}{\sum_{i=1}^{p} \delta_i A_X(\gamma_i s)}.$$

For the GGDML marginals, the innovation sequence of the process has *apgf*,

$$A_\epsilon(s) = \frac{[1 + \beta \ln(1 + cs^\alpha)]^{-1}}{\sum_{i=1}^{p} \delta_i \left[1 + \beta \ln(1 + c\gamma_i^\alpha s^\alpha)\right]^{-1}}. \tag{2.3.5}$$

For the particular case of $\gamma_i = \gamma$, for $i = 1, 2, \cdots, p$, (2.3.5) yields the similar pattern of *apgf* defined in (2.2.2). Hence with an error sequence $\{\epsilon_n\}$ distributed as GGDML random variables, the p^{th} order GGDML autoregressive processes are properly defined.

2.4 Further Extensions of GDML and GGDML Distributions

In this section we extend the GDML distribution to obtain a more general class of distributions called generalized discrete semi- Mittag-Leffler (GDSML) distribution and study its properties.

Definition 2.4.1. *A random variable X on $\mathbb{Z}_+$ is said to follow generalized discrete semi Mittag-Leffler distribution and write $X \stackrel{d}{=} GDSML(\alpha, c, \beta)$, if it has the pgf given by*

$$G(s) = \left\{ \frac{1}{1 + \psi(1 - s)} \right\}^\beta \tag{2.4.1}$$

where $\psi(s)$ satisfies the functional equation $a\psi(s) = \psi(a^\alpha s)$ for all $0 < s < 1$.

Remark 2.4.1. *The solution of the functional equation is given by* $\psi(s) = s^{\alpha}h(s)$ *where* $h(s)$ *is a periodic function in* $\ln s$ *with period* $\frac{-2\pi\alpha}{\ln a}$. *This is a special case of the general equation given in pp.310 in Aczel (1966). For more details see Jayakumar (1997) and Kagan et al.(1973).*

In a similar manner we can define a geometric generalized discrete semi Mittag-Leffler distribution and write $X \stackrel{d}{=} GGDSML(\alpha, c, \beta)$, if it has the *pgf*,

$$G(s) = \frac{1}{1+\beta\ln[1+\psi(1-s)]} \tag{2.4.2}$$

where $\psi(\cdot)$ satisfies the above conditions. It can also be verified that the GGDSML distribution is geometrically infinitely divisible.

Theorem 2.4.1. *Let* $X_1, X_2, \cdots$, *are independently and identically distributed geometric generalized discrete semi Mittag-Leffler random variables with parameters* α *and* β *where* $Y = X_1 + X_2 + \cdots + X_{N(p)}$ *such that* $N(p)$ *follows geometric with mean* $\frac{1}{p}$, $P[N(p) = k] = p(1-p)^{k-1},\ k = 1, 2, \cdots, 0 < p < 1$. *Then* $Y \stackrel{d}{=} GGDSML(\alpha, c, \frac{\beta}{p})$.

Proof: Taking the *apgf* of Y we have,

$$\begin{aligned} A_Y(s) &= \sum_{k=1}^{\infty}\{A_X(s)\}^k p(1-p)^{k-1} \\ &= \frac{1}{1+\frac{\beta}{p}\ln[1+\psi(s)]} \end{aligned}$$

Theorem 2.4.2. *Geometric Generalized Discrete Semi Mittag-Leffler distribution is the limit distribution of geometric sum of GDSML* $(\alpha, \frac{\beta}{n})$ *random variables.*

Proof: We have,

$$(1+\psi(s))^{-\beta} = \left\{1 + [1+\psi(s)]^{\frac{\beta}{n}} - 1\right\}^{-n}$$

is the *apgf* of a probability distribution since generalized discrete semi Mittag-Leffler distribution is infinitely divisible. Hence by lemma 3.2 of Pillai (1990)

$$A_n(s) = \left\{1 + n[1 + \psi(s)]^{\frac{\beta}{n}} - 1\right\}^{-n}$$

is the *apgf* of a geometric sum of independently and identically distributed discrete semi Mittag-Leffler random variables. Taking limit as $n \to \infty$

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

2.4.1 Geometric Generalized Discrete Semi Mittag-Leffler Processes

Here we develop a first order new autoregressive process with geometric generalized discrete semi Mittag-Leffler marginals.

Theorem 2.4.3. *Let $\{X_n, n \geq 1\}$ be defined as*

$$X_n = \begin{cases} \epsilon_n, & \text{with probability} \quad p \\ X_{n-1} + \epsilon_n, & \text{with probability} \quad 1-p \end{cases} \tag{2.4.3}$$

where $\{\epsilon_n\}$ is a sequence of i.i.d. random variables. A necessary and sufficient condition that $\{X_n\}$ is a strictly stationary Markov process with GGDSML(α, c, β) marginals is that ϵ_n are distributed as geometric generalized discrete semi Mittag-Leffler.

Proof: Rewriting in terms of the *apgf*, the equation (2.4.3) reduces to

$$\begin{aligned} A_{X_n}(s) &= pA_{\epsilon_n}(s) + (1-p)A_{X_{n-1}}(s)A_{\epsilon_n}(s) \\ &= A_{\epsilon_n}(s)\{p + (1-p)A_{X_{n-1}}(s)\}. \end{aligned} \tag{2.4.4}$$

When X_n is weak stationary, we have

$$A_X(s) = A_\epsilon(s)\{p + (1-p)A_X(s)\}.$$

This gives,

$$A_\epsilon(s) = \frac{A_X(s)}{p + (1-p)A_X(s)}$$

where

$$A_X(s) = \frac{1}{1 + \beta \ln[1 + \psi(s)]}.$$

On simplification we get,

$$A_\epsilon(s) = \frac{1}{1 + p\beta \ln[1 + \psi(s)]}$$

and hence $\epsilon_n \stackrel{d}{=} GGDSML(\alpha, c, p\beta)$.

The converse part can be proved by the method of mathematical induction as follows. Now assume that $X_n \stackrel{d}{=} GGDSML(\alpha, c, \beta)$. Then

$$\begin{aligned} A_{X_{n-1}}(s) &= A_{\epsilon_n}(s)\{p + (1-p)A_{X_{n-2}}(s)\} \\ &= \frac{1}{1 + p\beta \ln[1 + \psi(s)]} \left[p + (1-p)\left\{\frac{1}{1 + \beta \ln[1 + \psi(s)]}\right\}\right] \\ &= \frac{1}{1 + \beta \ln[1 + \psi(s)]}. \end{aligned}$$

The rest follows easily.

2.4.2 kth Order GGDSML Processes

Consider the kth order autoregressive process.

$$X_n = \begin{cases} \epsilon_n, & \text{with probability} \quad p \\ X_{n-1} + \epsilon_n, & \text{with probability} \quad p_1 \\ X_{n-2} + \epsilon_n, & \text{with probability} \quad p_2 \\ \cdots & \cdots \qquad\qquad\qquad \cdots \\ \cdots & \cdots \qquad\qquad\qquad \cdots \\ X_{n-k} + \epsilon_n, & \text{with probability} \quad p_k \end{cases} \tag{2.4.5}$$

where $p + p_1 + p_2 + \cdots + p_k = 1,\ 0 < p, p_i \leq 1,\ i = 1, 2, \cdots, k$ and $\{\epsilon_n\}$ is a sequence of *i.i.d.* random variables independent of $\{X_n, X_{n-1}, \cdots\}$. In terms of *apgf* we have

$$\begin{aligned} A_{X_n}(s) &= pA_{\epsilon_n}(s) + p_1 A_{X_{n-1}}(s) A_{\epsilon_n}(s) + \cdots + p_k A_{X_{n-k}}(s) A_{\epsilon_n}(s) \\ &= A_{\epsilon_n}(s)\{p + p_1 A_{X_{n-1}}(s) + \cdots + p_k A_{X_{n-k}}(s)\}. \end{aligned} \tag{2.4.6}$$

Under stationarity equilibrium, this gives,

$$A_\epsilon(s) = \frac{A_X(s)}{p + (1-p)A_X(s)}.$$

This shows that the results developed in $(2.4.3)$ can be applied in this also. This gives rise to the kth order geometric generalized discrete semi Mittag-Leffler autoregressive processes.

2.5 Bivariate Discrete Semi Mittag-Leffler (BDSML) Distribution

Autoregressive processes whose stationary marginals follow bivariate discrete Mittag-Leffler distribution have been developed by Jayakumar et al. (2010). The $apgf$ of a random variable with bivariate discrete Mittag-Leffler distribution is

$$A(s_1, s_2) = \frac{1}{(1 + c_1 s_1^{\alpha_1})(1 + c_2 s_2^{\alpha_2}) - \delta^2 c_1 c_2 s_1^{\alpha_1} s_2^{\alpha_2}}$$

$0 < \alpha_1, \alpha_2 \leq 1,\ c_1, c_2 > 0,\ \delta^2 > 0$. It is denoted by BDML($\alpha_1, \alpha_2, c_1, c_2, \delta^2$).

In this section, we develop bivariate discrete semi Mittag-Leffler distribution and processes.

Definition 2.5.1. *A bivariate random variable (X, Y) on $\mathbb{Z}_+$ is said to follow bivariate discrete semi Mittag-Leffler distribution denoted by, $(X, Y) \stackrel{d}{=} BDSML(\alpha_1, \alpha_2, c_1, c_2, \delta^2)$ if it has the $apgf$,*

$$A(s_1, s_2) = \frac{1}{(1 + \psi_1(s_1))(1 + \psi_2(s_2)) - \delta^2 \psi_1(s_1) \psi_2(s_2)}$$

where $\psi(s)$ satisfies certain conditions given in definition (2.4.1).

Definition 2.5.2. *A bivariate random variable (X, Y) on $\mathbb{Z}_+$ is said to follow geometric BDSML distribution denoted by, $(X, Y) \stackrel{d}{=} GBDSML(\alpha_1, \alpha_2, c_1, c_2, \delta^2)$ if it has the $apgf$,*

$$A(s_1, s_2) = \frac{1}{1 + \ln\{[1 + \psi_1(s_1)][1 + \psi_2(s_2)] - \delta^2 \psi_1(s_1) \psi_2(s_2)\}}$$

where $\psi(s)$ satisfies certain conditions given in definition (2.4.1).

Remark 2.5.1. *Generalized BDSML distribution can be defined as the distribution with $apgf$,*

$$A(s_1, s_2) = \left[\frac{1}{(1 + \psi_1(s_1))(1 + \psi_2(s_2)) - \delta^2 \psi_1(s_1) \psi_2(s_2)}\right]^{\beta}$$

where $\psi(s)$ satisfies certain conditions given in definition (2.4.1).

2.6 Geometric Generalized Bivariate Discrete Semi Mittag-Leffler Distribution

Definition 2.6.1. *A bivariate random variable (X, Y) on $\mathbb{Z}_+$ is said to follow geometric generalized bivariate discrete Semi Mittag-Leffler distribution denoted by, $(X, Y) \stackrel{d}{=} GGBDSML(\alpha_1, \alpha_2, c_1, c_2, \delta^2, \beta)$ if it has the apgf,*

$$A(s_1, s_2) = \frac{1}{1 + \beta \ln \{[1 + \psi_1(s_1)][1 + \psi_2(s_2)] - \delta^2 \psi_1(s_1)\psi_2(s_2)\}}$$

where $\psi(s)$ satisfies certain conditions given in definition (2.4.1).

Remark 2.6.1. *Geometric generalized bivariate discrete semi Mittag-Leffler distribution is geometrically infinitely divisible.*

Theorem 2.6.1. *Let $U_i = (U_{1i}, U_{2i})$ and $U_1, U_2, \cdots$, are independently and identically distributed geometric generalized bivariate discrete semi Mittag-Leffler random variables with parameters $\alpha_1, \alpha_2, \beta$ and δ^2 where $T = U_1 + U_2 + \cdots + U_{N(p)}$ such that $N(p)$ follows geometric with mean $\frac{1}{p}$,*

$$P[N(p) = k] = p(1 - p)^{k-1}, \ k = 1, 2, \cdots, 0 < p < 1.$$

Then $T \stackrel{d}{=} GGBDSML(\alpha_1, \alpha_2, c_1, c_2, \beta, \delta^2)$.

Proof: Taking the $apgf$ of T we have,

$$\begin{aligned} A_T(s_1, s_2) &= \sum_{k=1}^{\infty} \{A_U(s_1, s_2)\}^k p(1 - p)^{k-1} \\ &= \frac{1}{1 + \frac{\beta}{p} \ln \{[1 + \psi_1(s_1)][1 + \psi_2(s_2)] - \delta^2 \psi_1(s_1)\psi_2(s_2)\}}. \end{aligned}$$

Theorem 2.6.2. *Geometric generalized bivariate discrete Semi Mittag-Leffler distribution is the limit distribution of geometric sum of GBDSML $(\alpha_1, \alpha_2, \frac{\beta}{n})$ random variables.*

Proof: we have, $[1+(1+\psi_1(s_1))(1+\psi_2(s_2))-\delta^2\psi_1(s_1)\psi_2(s_2)]^{-\beta}$

$$=\left\{1+(1+[1+\psi_1(s_1)][1+\psi_2(s_2)]-\delta^2\psi_1(s_1)\psi_2(s_2))^{\frac{\beta}{n}}-1\right\}^{-n}$$

is the $apgf$ of a probability distribution since generalized bivariate discrete semi Mittag-Leffler distribution is infinitely divisible. Hence by lemma 3.2 of Pillai (1990), we have

$$A_n(s_1,s_2)=\left\{1+n[1+(1+\psi_1(s_1))(1+\psi_2(s_2))-\delta^2\psi_1(s_1)\psi_2(s_2)]^{\frac{\beta}{n}}-1\right\}^{-n}$$

is the $apgf$ of a geometric sum of independently and identically distributed generalized bivariate discrete semi Mittag-Leffler random variables. Taking limit as $n\to\infty$, we get

$$\begin{aligned} A(s_1,s_2) &= \lim_{n\to\infty} A_n(s_1,s_2) \\ &= \left\{1+\lim_{n\to\infty}\left(n[1+(1+\psi_1(s_1))(1+\psi_2(s_2))-\delta^2\psi_1(s_1)\psi_2(s_2)]^{\frac{\beta}{n}}-1\right)\right\}^{-1} \\ &= \left\{1+\beta\ln[1+(1+\psi_1(s_1))(1+\psi_2(s_2))-\delta^2\psi_1(s_1)\psi_2(s_2)]\right\}^{-1}. \end{aligned}$$

2.6.1 GGBDSML Processes

Here we develop a first order new autoregressive process with geometric generalized bivariate discrete semi Mittag-Leffler marginals.

Theorem 2.6.3. *Let $\{U_n, n\geq 1\}$ be defined as*

$$U_n=\begin{cases}\epsilon_n, & \textit{with probability} \quad p \\ U_{n-1}+\epsilon_n, & \textit{with probability} \quad 1-p\end{cases} \tag{2.6.1}$$

where $\{\epsilon_n\}$ is a sequence of i.i.d. random variables. A necessary and sufficient condition that $\{U_n\}$ is a strictly stationary Markov process with GGBDSML$(\alpha_1,\alpha_2,c_1,c_2,\beta,\delta^2)$ marginals is that ϵ_n are distributed as geometric generalized discrete semi Mittag-Leffler, where $U_n=(U_{1n},U_{2n})$ and $\epsilon_n=(\epsilon_{1n},\epsilon_{2n})$ are sequences of bivariate random variables.

Proof: Taking the $apgf$ of an equation (2.6.1) we get

$$\begin{aligned} A_{U_n}(s_1,s_2) &= pA_{\epsilon_n}(s_1,s_2)+(1-p)A_{U_{n-1}}(s_1,s_2)A_{\epsilon_n}(s_1,s_2) \\ &= A_{\epsilon_n}(s)\{p+(1-p)A_{U_{n-1}}(s_1,s_2)\}. \end{aligned} \tag{2.6.2}$$

Under stationarity it reduces to

$$A_U(s_1,s_2)=A_\epsilon(s_1,s_2)\{p+(1-p)A_U(s_1,s_2)\}.$$

Hence,

$$A_\epsilon(s_1,s_2)=\frac{A_U(s_1,s_2)}{p+(1-p)A_U(s_1,s_2)}$$

where

$$A_U(s_1,s_2)=\frac{1}{1+\beta\ln[1+(1+\psi_1(s_1))(1+\psi_2(s_2))-\delta^2\psi_1(s_1)\psi_2(s_2)]}.$$

On simplification we get,

$$A_\epsilon(s_1,s_2)=\frac{1}{1+p\beta\ln[1+(1+\psi_1(s_1))(1+\psi_2(s_2))-\delta^2\psi_1(s_1)\psi_2(s_2)]}$$

and hence $\epsilon_n \stackrel{d}{=} GGBDSML(\alpha_1,\alpha_2,c_1,c_2,\delta^2,p\beta)$.

The converse part can be proved by the method of mathematical induction as follows. Now assume that $U_n \stackrel{d}{=} GGBDSML(\alpha_1,\alpha_2,c_1,c_2,\delta^2,\beta)$. Then

$$\begin{aligned} A_{U_{n-1}}(s_1,s_2) &= A_{\epsilon_n}(s_1,s_2)\{p+(1-p)A_{U_{n-2}}(s_1,s_2)\} \\ &= \frac{1}{1+p\beta\ln[1+(1+\psi_1(s_1))(1+\psi_2(s_2))-\delta^2\psi_1(s_1)\psi_2(s_2)]} \\ &\quad \left[p+(1-p)\left\{\frac{1}{1+\beta\ln[1+(1+\psi_1(s_1))(1+\psi_2(s_2))-\delta^2\psi_1(s_1)\psi_2(s_2)]}\right\}\right] \\ &= \frac{1}{1+\beta\ln[1+(1+\psi_1(s_1))(1+\psi_2(s_2))-\delta^2\psi_1(s_1)\psi_2(s_2)]}. \end{aligned}$$

The rest follows easily.

2.7 Application to an Empirical Data

In this section we apply the model to a data on the inter-arrival times of customers in a bank counter measured in terms of number of months from January 1994 to October 2003, which is taken from the file bank.arrivals.xlsx available in the website www.westminstercolle ge.edu. The empirical pdf shows a decreasing trend in the probabilities. Figure 1 gives the empirical pdf and theoretical pdf of GDML$(\alpha,\ c,\ \beta)$.

The mean, variance, coefficient of skewness and kurtosis measure for the data are respectively 1.5435, 3.2314, 0.9845 and 2.6714. The Durbin-Watson test confirms strong autocorrelation in the data with first order autocorrelation coefficient 0.92 so that INAR models are needed to explore the future behaviour of the data. Since the mean is less than variance the geometric distribution is a possible probability model. Since geometric distribution is a special case of GDML$(\alpha,\ c,\ \beta)$, we shall examine whether it is a suitable model to the above data. We obtain the estimates of the parameters as $\alpha = 0.99$, $\beta = 1$ and $c = 0.91$.

Now we apply the Kolmogorov-Smirnov [K.S.] test for testing H_0: GDML distribution with parameters $\alpha = 0.99$, $\beta = 1$ and $c = 0.91$ is a good fit for the given data. Since the computed value of the K.S. test statistic is obtained as 0.1212 and the critical value corresponding to the significance level 0.01 is 0.2403, the GDML assumption for inter-arrival times is justified. Using this we can obtain the probabilities associated with the stationary distribution of the INAR(1) model as well as predict the future values of the process. This will help in developing optimal service policies for ensuring customer satisfaction.

2.8 Conclusions

In this chapter we have considered GDML distributions and introduced a new family of distributions called GGDML distributions and developed integer-valued time series models. We also developed various generalizations such as GGDSML and INAR (p) processes.

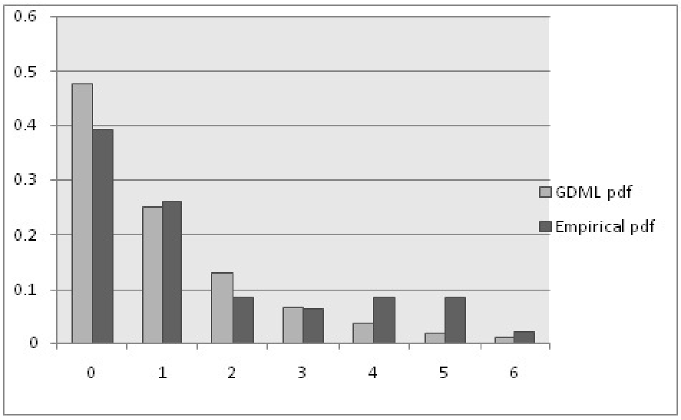

Figure 2.1: The empirical *pdf* and theoretical *pdf* of GDML (0.99, 0.91, 1) distribution.

The use of the model is illustrated by fitting it to an empirical data on customer arrivals in a bank counter and the goodness of fit is established. The processes developed in this paper can be used for modeling time series data on counts of events, objects or individuals at consecutive points in time such as the number of accidents, number of breakdowns in manufacturing plants, number of busy lines in a telephone network, number of patients admitted in a hospital, number of claims in an insurance company, number of persons unemployed in a particular year, number of aero planes waiting for take-off, number of vehicles in a queue, etc. Thus the models have applications in various contexts like studies relating to human resource development, insect growth, epidemic modeling, industrial risk modeling, insurance and actuaries, town planning etc.

References

Aczel, J. (1966). *Lectures on functional equations and their applications*. Academic Press INC, New York.

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

Haubold, H.J., Mathai, A.M. (2000). The fractional kinetic equation and thermonuclear functions. *Astrophysics and Space Science*, **273**, 53-63.

Jayakumar, K., Pillai, R.N. (1993). The first-order autoregressive Mittag-Leffler process. *Journal of Applied Probability*, **30**, 462-466.

Jayakumar, K. (1997). First order autoregressive semi alpha-Laplace processes. *Statistica*, LVII, pp. 455-463.

Jayakumar, K., Ajitha, B.K. (2003). On the geometric Mittag-Leffler distributions. *Calcutta Statistical Association Bulletin*, **54**, Nos. 215-216, 195-208.

Jayakumar, K., Ristic, M.M., Davis, A.M. (2010). A generalization to bivariate Mittag-Leffler and bivariate discrete Mittag-Leffler autoregressive processes. *Communications in Staistics-Theory and Methods*, **39**, 942-955.

Jose, K.K., Abraham, B. (2011). A Count Model based on with Mittag-Leffler interarrival times. *Statistica*, anno LXXI, n.**4**, 501-514.

Jose, K.K., Mariyamma, K.D., (2012a). Integer valued autoregressive processes with generalized discrete Mittag-Leffler marginals. *Statistica*, anno LXXII, n.2, 195-209.

Jose, K.K., Mariyamma, K.D., (2012b). Geometric Generalizations of Discrete Semi Mittag-Leffler Distributions and Processes. *VISTAS*, **1**(1), 71-80.

Jose, K.K., Uma, P., Seetha Lekshmi, V., Haubold, H.J. (2010). Generalized Mittag-Leffler distributions and processes for Applications in Astrophysics and Time Series Modeling. *Astrophysics and Space Science Proceedings*, 79-92.

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

Lawrance, A., Lewis, P. (1980). The exponential autoregressive moving average EARMA (p,q) Process. *Journal of the Royal Statistical Society*, **B 42**, 150-156.

McKenzie, E. (1986). Autoregressive-moving average processes with negative binomial and geometric marginal distributions. *Advances in Applied Probability*, **18**, 679-705.

Pillai, R.N. (1990). Harmonic mixtures and geometric infinite divisibility. *Journal of Indian Statistical Association*, **28**, 87-98.

Pillai, R.N., Jayakumar, K. (1995). Discrete Mittag-Lefler distribution. *Statistics and Probability Letters*, **23**, 271-274.

Pillai, R.N., Sandhya, E. (1990). Distributions with monotone derivative and geometric infinite divisibility. *Advances in Applied Probability*, **22**, 751-754.

Prabhakar, T.R. (1971). A singular integral equation with a generalized Mittag-Leffler function in the kernel. *Yokohama Mathematical Journal*, **19**, 7-15.

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

Seetha Lekshmi, V., Jose, K.K. (2002). Geometric Mittag-Leffler tailed autoregressive processes. *Far East Journal of Theoretical Statistics*, **6**, 147-153.

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

Steutel, F. W., van Harn, K. (1979). Discrete analogues of self- decomposability and stability. *Annals of Probability*, **7**, 893-899.

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