"Science Stays True Here"
Journal of Mathematics and Statistical Science (ISSN 2411-2518, USA), Vol.7, Issue 7, 183-202 | Science Signpost Publishing

Interswitching of Transmuted Gamma Autoregressive Random Processes

Rasaki Olawale Olanrewaju[1], Johnson Funminiyi Ojo[2], Adekola Lanrewaju Olumide[3]

1. Department of Mathematical Sciences, Pan African University Institute for Basic Sciences, Technology and Innovation. P.O. Box 62000-00200, Nairobi, Kenya. Email: rasakiolawale@gmail.com.

2. Department of Statistics, University of Ibadan, Ibadan, P.O. Box 200284, Oyo state, Nigeria.

3. Department of Physical Sciences, Bells University of Technology, Ota, Nigeria.

Abstract

We study the assortment of autoregressive random processes via a transmuted Gamma distributed noise. We consider a transmuted re-parameterization of the Gamma parameters in terms of μ and σ^2, afterwards ascertained that the transmuted Gamma is a proper probability density function, then proceeded to spelt-out the structural form and traits of the Gamma Mixture Autoregressive generalization in its *k*-components. The mean and variance of the Gamma Autoregressive model were ascertained coupled with its first and second-order stationarity. The ingrained *k*-components' autoregressive coefficients, re-parameterization Gamma coefficients, *k*-regime transitional weights were estimated via Expectation-Maximization (EM) algorithm. However, some step ahead predictions were derived as well as the model sub-setting estimation via Levinson-Durbin recursive technique.

***Keywords**:* Expectation-Maximization (EM) algorithm, Gamma Autoregressive model, *k*-components, *k*-regime transitional weights, Levinson-Durbin recursive, Transmuted.

***MSC**:* 33B20, 37A25, 37M10, 62M10

1. Introduction

This novel paper aims at developing, extending and discussing mixture conditional of autoregressive processes to a transmuted Gamma mixture autoregressive processes, denoted by $GAMMAR(k; p_1, p_2, \ldots, p_k)$ with *k*-components states (regimes). The build-up will start from specifying the mixture autoregressive model denoted by $MAR(k; p_1, p_2, \dots, p_k)$. This will be extended to $GAMMAR(k; p_1, p_2, \dots, p_k)$ with k-regimes mixture of k-stationary or non-stationary shifting autoregressive processes per component. Firstly, the conventional two-parameter Gamma distribution will be transmuted via its location and shape parameters of α and β respectively in terms of μ & σ^2. The transmuted two-parameter distribution would be sized-up to ascertain if it is a proper probability density function, that is, if $\int f(x_t) = 1$ before incorporating and substituting it as the white noise for $MAR(k; p_1, p_2, \dots, p_k)$ model. The mean and variance of the multimodal conditional random process and transitional model will be determined coupled with the first and second-order stationary. According to Kalliovirta *et al.* (2016) and Olanrewaju *et al.* (2021), one of the unique property of the MAR model is that a mixture of one or more stationary AR component(s) with non-stationary AR component(s) will surely result in an overall stationary process. This would also be extended to $GAMMAR(k; p_1, p_2, \dots, p_k)$ via a necessary and sufficient condition for the process to be stationary in the mean root of its equation.

In addition, the transmuted Gamma distributed noise would make it possible for the stochastic model to handle and include some steps ahead prediction (forward forecast using the immediate or previous estimation(s) and cycles in a positive series contaminated with excess skewness and kurtosis. The Expectation-Maximization (EM) estimation technique will be adopted via E-step and M-step that would lead to a system of equation and Newton-Raphon iterative technique for estimating AR coefficients; transitional weights, transmuted Gamma coefficients per each regime with their associated standard error. Furthermore, sub-setting of the $GAMMAR(k; p_1, p_2, \dots, p_k)$ model will be carried-out to check the optimality combinations of the mixing model. The sub-setting estimation will be via Levinson-Durbin recursive technique.

2. 0 Preliminary

2.1 Specification of Gamma Mixture Autoregressive Model

Le *et al*. (1996) introduced a Gaussian mixture transition distribution (GMTD) models of conditional Gaussian distribution as:

$$g(x) = \lambda_1 f_1(x) + \lambda_2 f_2(x) + \cdots + \lambda_k f_k(x) \quad (1)$$

where $g(x)$ is whole mixture regime-switching probability density function and $f_i(x)(i = 1, \ldots, n)$ are the probability density functions of the random noise such that the Markov transitional weight is given by $\lambda_1 + \lambda_2 + \cdots + \lambda_k \approx 1 \ni \lambda_i > 0$.

Extending Wong & Li (2000); Boshnakov (2006); and Nastic (2014)'s definition of *k*-component of Mixture Autoregressive (MAR) model to Gamma related random noise via transformed parameters gives:

$$F\left(x_{(t)} / f_{t-1}\right) = \sum_{\mathrm{k}=1}^{K} (\lambda_{k,} \frac{1}{\sigma_k^2}, \mu_k \sigma_k^2) \; \Phi\left(\frac{x_t - \phi_{k,0} - \phi_{k,1} x_{t-1} - \cdots - \phi_{k,pk}\, x_{t-pk}}{\sigma_k}\right) \quad (2)$$

Otherwise,

$$X_t = \begin{cases} \phi_{1,0} + \sum_{i=1}^{p1} \phi_{1,i} x_{t-1} + \varepsilon_t^{(1)} \ni \lambda_1 \ \& \ (\frac{1}{\sigma_1^2}, \mu_1 \sigma_1^2) \\ \phi_{2,0} + \sum_{i=1}^{p2} \phi_{2,i} x_{t-2} + \varepsilon_t^{(2)} \ni \lambda_2 \ \& \ (\frac{1}{\sigma_2^2}, \mu_2 \sigma_2^2) \\ \vdots \qquad \vdots \qquad \vdots \\ \phi_{k,0} + \sum_{i=1}^{pk} \phi_{k,i} x_{t-p} + \varepsilon_t^{(k)} \ni \lambda_k \ \& \ (\frac{1}{\sigma_k^2}, \mu_k \sigma_k^2) \end{cases} \quad (3)$$

$\ni$ the transmuted Gamma location and shape parameters are defined in terms of μ and σ^2 for a transmuted and unconventional Gamma distribution as defined below in equation (4) with α and β of the conventional defined below as.

$$\alpha = \frac{1}{\sigma^2}; \quad \beta = \mu \sigma^2 \quad (4)$$

The model is denoted by $GAMMA - MAR(k; p_1, p_2, \ldots, p_k)$,where $F\left(x_{(t)} / f_{t-1}\right)$ is the conditional cumulative distribution function of X_t given the previous past information evaluated at x_t, $\phi_{pk} \in (0,1)$, $0 < \phi_{pk} < 1 \;\; \forall \; k = 1, \cdots, K$ with Markov transitional probabilities of $\sum_{i=1}^{k} \lambda_i \approx 1$, $\Phi(.)$ is the cumulative distribution function of the standard Gamma**.**

$$f_x(x_t/\mu,\sigma)=\frac{x_t^{\frac{1}{\sigma^2}-1}e^{-\frac{x_t}{(\sigma^2\mu)}}}{(\sigma^2\mu)^{\frac{1}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}\qquad x_t\in(0,\infty) \tag{5}$$

such that setting $\alpha=\frac{1}{\sigma^2}$, $\beta=\mu\sigma^2$. Hence $\mu=\alpha\beta,\ \sigma^2=\frac{1}{\alpha}$ gives back the conventional gamma distribution of

$$f_x(x_t)=\left(\frac{x_t}{\beta}\right)^{\alpha-1}\times\frac{e\left(-\frac{x_t}{\beta}\right)}{\beta\Gamma(\alpha)}\qquad x_t\in(0,\infty) \tag{6}$$

where $$\Gamma(\alpha)=\int_0^\infty e^t\,t^{\alpha-1}\,\partial t$$

with scale parameter $\beta>0$ and shape parameter $\alpha>0$. The scale parameter influences the spread (center of location) of the distribution while the shape parameter controls the skewness parameter of the distribution such that as the shape parameter increases the distribution symmetric expands, for a Gamma distribution for positively skew data.

2.2 Investigation of the Proper P.D.F of the Transmuted Gamma Distribution.

Verifying whether the transmuted Gamma distribution is a proper P.D.F,

Given $$f_x(x_t/\mu,\sigma)=\frac{x_t^{\frac{1}{\sigma^2}-1}e^{-\frac{x_t}{(\sigma^2\mu)}}}{(\sigma^2\mu)^{\frac{1}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}$$

$$\int f(x_t)\ \ \partial x_t\ \ =\int_0^\infty\frac{x^{\frac{1}{\sigma^2}-1}e^{-\frac{x}{(\sigma^2\mu)}}}{(\sigma^2\mu)^{\frac{1}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}\ \partial x_t$$

$$=\frac{1}{(\sigma^2\mu)^{\frac{1}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}\int_0^\infty x_t^{\frac{1}{\sigma^2}-1}e^{-\frac{x_t}{(\sigma^2\mu)}}\ \partial x_t \tag{7}$$

Changing of varaiable technique gives,

$$m = \frac{x_t}{\sigma^2 \mu} \qquad \Rightarrow x_t = m\sigma^2 \mu$$

therefore, $$\frac{\partial x_t}{\partial m} = \sigma^2 \mu \qquad \Rightarrow \partial x_t = \sigma^2 \mu \, \partial m$$

It implies that,

$$\int f(x_t) \ \partial x_t = \frac{1}{(\sigma^2 \mu)^{\frac{1}{\sigma^2}} \Gamma(\frac{1}{\sigma^2})} \int_0^\infty \left(m\sigma^2 \mu\right)^{\frac{1}{\sigma^2}-1} e^{-m} \, \sigma^2 \mu \partial x_t$$

$$\int f(x_t) \ \partial x = \frac{1}{(\sigma^2 \mu)^{\frac{1}{\sigma^2}} \Gamma(\frac{1}{\sigma^2})} \int_0^\infty \left(\sigma^2 \mu\right)^{\frac{1}{\sigma^2}-1} \left(m\right)^{\frac{1}{\sigma^2}-1} e^{-m} \left(\sigma^2 \mu\right) \partial x_t \qquad (8)$$

$$\int f(x_t) \ \partial x_t = \frac{\left(\sigma^2 \mu\right)^{\frac{1}{\sigma^2}-1} \left(\sigma^2 \mu\right)}{(\sigma^2 \mu)^{\frac{1}{\sigma^2}} \Gamma(\frac{1}{\sigma^2})} \int_0^\infty \left(m\right)^{\frac{1}{\sigma^2}-1} e^{-m} \, \partial x_t$$

Recall, $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$

So, $\Gamma(\frac{1}{\sigma^2}) = \int_0^\infty \left(m\right)^{\frac{1}{\sigma^2}-1} e^{-m} \partial m$

$$\int f(x_t) \ \partial x_t = \frac{\left(\sigma^2 \mu\right)^{\frac{1}{\sigma^2}-1+1}}{(\sigma^2 \mu)^{\frac{1}{\sigma^2}} \Gamma(\frac{1}{\sigma^2})} \times \ \Gamma(\frac{1}{\sigma^2})$$

$$\int f(x_t) \ \partial x_t = \frac{\left(\sigma^2 \mu\right)^{\frac{1}{\sigma^2}}}{(\sigma^2 \mu)^{\frac{1}{\sigma^2}} \Gamma(\frac{1}{\sigma^2})} \times \ \Gamma(\frac{1}{\sigma^2})$$

$$\int f(x_t) \ \partial x_t = 1 \qquad (9)$$

This connotes and confirms that the transmuted Gamma distribution is a proper P.D.F

2.3 The r^{th} Moment of Transmuted Gamma distribution

$$E(X_t^r) = \int_0^\infty x_t^r \; f_x(x_t / \mu, \sigma)\, \partial x_t^r \tag{10}$$

$$E(X_t^r) = \int_0^\infty x_t^r \, \frac{x_t^{\frac{1}{\sigma^2}-1} e^{-\frac{x_t}{(\sigma^2\mu)}}}{(\sigma^2\mu)^{\frac{1}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})} \partial x_t^r$$

$$E(X_t^r) = \frac{1}{(\sigma^2\mu)^{\frac{1}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})} \int_0^\infty x_t^r \, x_t^{\frac{1}{\sigma^2}-1} e^{-\frac{x_t}{(\sigma^2\mu)}} \partial x_t^r$$

$$E(X_t^r) = \frac{1}{(\sigma^2\mu)^{\frac{1}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})} \int_0^\infty x_t^{\frac{1}{\sigma^2}+r-1} e^{-\frac{x_t}{(\sigma^2\mu)}} \partial x_t^r$$

Recall, $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} \partial t$

$$\text{so, } \left(\sigma^2\mu\right)^{\frac{1}{\sigma^2}} \mu^r \sigma^{2r} \Gamma(\frac{1}{\sigma^2}+r) = \int_0^\infty x_t^{\frac{1}{\sigma^2}+r-1} e^{-\frac{x_t}{(\sigma^2\mu)}} \partial x_t^r \tag{11}$$

Recall, $\left(ab\right)^r \Gamma(r) = \int_0^\infty y^{r+1} e^{-\frac{y}{(ab)}} \partial y^r$

$$E(X_t^r) = \frac{1}{(\sigma^2\mu)^{\frac{1}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})} \left(\sigma^2\mu\right)^{\frac{1}{\sigma^2}} \mu^r \sigma^{2r} \Gamma(\frac{1}{\sigma^2}+r)$$

$$E(X_t^r) = \frac{\mu^r \sigma^{2r} \Gamma(\frac{1}{\sigma^2}+r)}{\Gamma(\frac{1}{\sigma^2})} \quad \text{for} \quad r > -\frac{1}{\sigma^2} \tag{12}$$

At $r = 1$

$$E(X_t^1) = \mu^1 = \mu = \sigma\beta = \frac{1}{\sigma^2} \times \mu\sigma^2 = \mu \tag{13}$$

$$E(X_t^2) = \alpha\beta^2 + \alpha\beta = \alpha\beta(\beta+1) = \mu^2\sigma^2 + \mu^2 = \mu^2(\sigma^2+1) \quad (14)$$

So, $Var(X_t) = \alpha\beta^2 + \alpha\beta - (\alpha\beta)^2$

$$\alpha\beta^2 = \sigma^2\mu^2 \quad (15)$$

For $x_t > 0$ where $\mu > 0$ and $\sigma > 0$. $E(X_t) = \mu$ & $Var(X_t) = \sigma^2\mu^2$

2.4 The Conditional mean and variance for the Transmuted Gamma Mixture Autoregressive Model.

$$E(x_t / GA_{t-1}) = \sum_{k=1}^{K} (\lambda_k, \frac{1}{\sigma_k^2}, \mu_k\sigma_k^2)\left(\phi_{k0} + \phi_{k1}x_{t-1} + \cdots + \phi_{kp_k}x_{t-p_k}\right) = \sum_{k=1}^{K} (\lambda_k, \frac{1}{\sigma_k^2}, \mu_k\sigma_k^2)\,\mu_{k,t} \quad (16)$$

But, $$\mu_{k,t} = \sum_{k=1}^{K}\left(\phi_{i0}\sum_{p=1}^{p_k}\phi_{i1}x_{t-p_k}\right) = \phi_{k0} + \phi_{k1}x_{t-1} + \cdots + \phi_{kp_k}x_{t-p_k} \quad (17)$$

$$E(x_t^2 / GA_{t-1}) = \sum_{k=1}^{K} (\lambda_k, \frac{1}{\sigma_k^2}, \mu_k\sigma_k^2)\sigma_k^2 + \sum_{k=1}^{K} (\lambda_k, \frac{1}{\sigma_k^2}, \mu_k\sigma_k^2)\mu_{k,t}^2 \quad (18)$$

So, $$Var(x_t / GA_{t-1}) = \sum_{k=1}^{K} (\lambda_k, \frac{1}{\sigma_k^2}, \mu_k\sigma_k^2)\sigma_k^2 + \sum_{k=1}^{K} (\lambda_k, \frac{1}{\sigma_k^2}, \mu_k\sigma_k^2)\mu_{k,t}^2 - \left(\sum_{k=1}^{K} (\lambda_k, \frac{1}{\sigma_k^2}, \mu_k\sigma_k^2)\mu_{k,t}\right)^2$$

(19)

Otherwise,

$$Var(x_t / GA_{t-1}) = \sum_{k=1}^{K} (\lambda_k, \alpha_k, \beta_k)\sigma_k^2 + \sum_{k=1}^{K} (\lambda_k, \alpha_k, \beta_k)\mu_{k,t}^2 - \left(\sum_{k=1}^{K} (\lambda_k, \alpha_k, \beta_k)\mu_{k,t}\right)^2 \quad (20)$$

The minus of the last two components are positive functions (that is, non-negative), which makes the conditional variance strictly greater than the expected mean $(\mu_{k,t})$. This makes it desirable for the Gamma Mixture time series model to possibly exhibit and capture thicker tails (Mesokurtic, plytokurtic and lesokurtic) that might over-whelmed the Gaussian distribution. Where $\sum_{k=1}^{K}(\lambda_k, \sigma_k^2, \frac{1}{\sigma_k^2}, \mu_k\sigma_k^2)$ is the specific value for conditional variance (volatility) for each regime, and means

$\mu_{1,t} = \mu_{2,t}, = \mu_{3,t} = \cdots = \mu_{k,t}$

2.5 Stationary Process for the GAMMAR $(k; p_1, p_2, \ldots, p_k)$

From equation (2)

$$F(x_t / GA_{t-1}) = \sum_{k=1}^{K} (\lambda_k, \frac{1}{\sigma_k^2}, \mu_k \sigma_k^2) \Phi \left\{ \frac{x_t - \phi_{k0} - \phi_{k1} x_{t-1} - \cdots - \phi_{kp_k} x_{t-p_k}}{\sigma_k} \right\} \tag{21}$$

Its conditional expectation (1st order stationary) is

$$\mu_t = E(x_t / x_{t-1}) = E\left[(\lambda_0, \frac{1}{\sigma_0^2}, \mu_0 \sigma_0^2) \Phi \left(\frac{\sum_{j=1}^{P} \phi_{pj}}{\sigma_0} \right) \right] + E\left[\sum_{k=1}^{K} (\lambda_0, \frac{1}{\sigma_0^2}, \mu_0 \sigma_0^2) \Phi \left(\frac{x_t - \sum_{i=1}^{P} \phi_{ki} x_{t-i_k}}{\sigma_0} \right) \right] \tag{22}$$

$$GA_t(x_t) = \sum_{k=1}^{K} \lambda_k \int g_k(x_t) GA_{t-1}(x_{t-1}) \partial x_{t-1} \tag{23}$$

Where, $g_k(x_t) = (\frac{1}{\sigma_k^2}, \mu_k \sigma_k^2) \Phi \left(\frac{x_t - \phi_{k1} x_{t-1}}{\sigma_k} \right)$

So, $E(x_t) = \mu_t = E(x_t / x_{t-1}) = \int x_t GA(x_t / FGA_{t-1}) \partial x_t$ where FGA is CDF of GA

$$= \int x_t \sum_{k=1}^{K} (\lambda_k, \frac{1}{\sigma_k^2}, \mu_k \sigma_k^2) g_k(x_t) \partial x_t$$

$$= \sum_{k=1}^{K} \frac{(\lambda_k, \frac{1}{\sigma_k^2}, \mu_k \sigma_k^2)}{\sigma_k} \int \left(x_t - \phi_{k1} x_{t-1} \right) \partial x_t$$

$$= \sum_{k=1}^{K} \frac{(\lambda_k, \frac{1}{\sigma_k^2}, \mu_k \sigma_k^2)}{\sigma_k} \phi_{k0} + \sum_{i=1}^{P} \left(\sum_{k=1}^{K} (\lambda_k, \frac{1}{\sigma_k^2}, \mu_k \sigma_k^2) \phi_{k1} \right) \mu_{t-1} \tag{24}$$

μ_t satisfied the necessary and sufficient 1st order condition of being finite and independent of "t" at first moment. Hence, $\mu_{0,t} = \mu_{0,t-0}$, $\mu_{1,t} = \mu_{1,t-1}$. μ is that the root of equation of the AR, say $s_1, \cdots\cdots, s_p$ is

$$1-\sum_{i=1}^{P}\left(\sum_{k=1}^{K}(\lambda_k,\alpha_k,\beta_k)\right)s^{-i}=0 \tag{25}$$

That is, all the equations (that is the Autoregressive coefficients at each regime) lies outside the unit circle ($\phi_{ki}<1$) $for\ i>p_k$ expect $\phi_{k0}=0$

$$E(x_t^2 / FGAx_{t-1})=\int x_t^2\, GA(x_t / FGA_{t-1})\,\partial x_t \tag{26}$$

$$=\int x_t^2\sum_{k=1}^{K}(\lambda_k,\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)g_k(x_t)\partial x_t$$

But, $g_k(x_t)=\Phi(\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)\left(\frac{x_t-\phi_{k1}x_{t-1}-\phi_{k2}x_{t-2}}{\sigma_k}\right)$

$$=\sum_{k=1}^{K}\frac{(\lambda_k,\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)}{\sigma_k}\int\left(x_t-\phi_{k1}x_{t-1}-\phi_{k2}x_{t-2}+\phi_{k1}x_{t-1}+\phi_{k2}x_{t-2}\right)^2 g_k(x_t)\partial x_t$$

$$=\sum_{k=1}^{K}\frac{(\lambda_k,\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)}{\sigma_k}\left(\sigma_k^2+\phi_{k1}^2\mu_{t-1}^2+\phi_{k2}^2\mu_{t-2}+2\phi_{k1}\phi_{k2}\mu_{t-1}\mu_{t-2}\right) \tag{27}$$

The second-order stationary via variance of $Var\left(x_t / FGA_{t-1}\right)$

$$Var\left(x_t / FGA_{t-1}\right)=E\left[\left(x_t^2 / x_{t-1}\right)\right]-\left[E\left(x_t / x_{t-1}\right)\right]^2 \tag{28}$$

$$=\sum_{k=1}^{K}\frac{(\lambda_k,\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)}{\sigma_k}\left(\sigma_k^2+\phi_{k1}^2\mu_{t-1}^2+\phi_{k2}^2\mu_{t-2}+2\phi_{k1}\phi_{k2}\mu_{t-1}\mu_{t-2}\right)-$$

$$\left[\sum_{k=1}^{K}\frac{(\lambda_k,\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)}{\sigma_k}\phi_{k0}+\sum_{i=1}^{P}\left(\sum_{k=1}^{K}(\lambda_k,\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)\phi_{k1}\right)\mu_{t-1}\right]^2$$

$$=\sum_{k=1}^{K}\frac{(\lambda_k,\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)}{\sigma_k}\left[\left(\sigma_k^2+\phi_{k1}^2\mu_{t-1}^2+\phi_{k2}^2\mu_{t-2}+2\phi_{k1}\phi_{k2}\mu_{t-1}\mu_{t-2}\right)-\right.$$

$$\sum_{k=1}^{K}\frac{(\lambda_k,\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)}{\sigma_k}\left(\phi_{k0}\right)^2+\left(+\sum_{i=1}^{P}\left(\sum_{k=1}^{K}(\lambda_k,\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)\phi_{k1}\right)\mu_{t-1}\right)^2\Bigg] \tag{29}$$

The covariance of x_t & x_{t-1}

$$\operatorname{cov}\left(x_t, x_{t-1}\right)=\Omega_{1,t}=E\left[x_t, x_{t-1}\right]= E\left[E\left(x_t / FGA_{t-1}\right)x_{t-1}\right] \tag{30}$$

$$=E\left[\frac{\sum_{k=1}^{K}(\lambda_k,\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)\phi_{k1}\,\Omega_{0.t-1}}{\sigma_k}\right]+\sum_{k=1}^{K}(\lambda_k,\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)\phi_{k2}\,\Omega_{0.t-1} \tag{31}$$

The process is second-order stationary $\Omega_{1,t}=\Omega_{1,t-1}$

$$\text{So,}\quad \Omega_{1,t}=\frac{\left[\sum_{k=1}^{K}(\lambda_k,\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)/\sigma_k\right]\Omega_{0,t-1}}{1-\sum_{k=1}^{K}(\lambda_k,\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)} \tag{32}$$

This implies that x_t follows a GAMMAR$\left(k;1,1,\cdots,1\right)$model first-order stationary for a requisite and adequate imposition for the process to be second-order stationary.

$$\left|\lambda_1\phi_{11}^2+\lambda_2\phi_{21}^2+\cdots+\lambda_k\phi_{k1}^2\right|<1$$

So, the m^{th} moment of the GAMMAR$\left(k;p_1,p_2,\cdots,p_k\right)$ for the finite regime an order p_k .

$$\sum_{k=1}^{K}(\lambda_k,\frac{1}{\sigma_k^2},\mu_k\sigma_k^2)\left(\phi_{k1}+\Omega_{pkt}\right)^m<1 \qquad \text{for}\quad p=p_1,\cdots,p_k \tag{33}$$

2.6. Regimes (Transitions) Probabilities λ_k of each Multimodality Estimates via Autocorrelation

Different techniques can be used to carve-out the weights (probabilities) of each regime. Forrest *et al*. (2015) define transition between regimes via proportional odds model, however the technique can be translated to time series via the autocorrelation idea of regime estimates;

$$\lambda_1(x_t) = \left(1 + e^{\Omega_1 + \Omega_0 x_t}\right)^{-1}$$

$$\lambda_2(x_t) = \left(1 + e^{\Omega_2 + \Omega_0 x_t}\right)^{-1} - \left(1 + e^{\Omega_1 + \Omega_0 x_t}\right)^{-1}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$\lambda_k(x_t) = 1 - \left(1 + e^{\Omega_k + \Omega_0 x_t}\right)^{-1}$$

where, $\lambda_k > 0 \ni \sum_{k=1}^{K} \lambda_k(x_t) \approx 1$

$$\Omega_k = \sum_{i=1}^{P}\left(\sum_{k=1}^{K}\left((\lambda_k, \frac{1}{\sigma_k^2}, \mu_k\sigma_k^2)\phi_{ki}\right)E\left(X_{t-k}X_{t-i}\right)\right)$$

$$\Omega_k = \sum_{i=1}^{P}\left(\sum_{k=1}^{K}\left((\lambda_k, \frac{1}{\sigma_k^2}, \mu_k\sigma_k^2)\phi_{ki}\right)\Omega_{|k-i|}\right) \quad (34)$$

2.7. GAMMAR$(k; p_1, p_2, \ldots, p_k)$ Parameter Estimation via EM-Algorithm

The estimation procedure to be adopted is the Expectation-Maximization (EM) algorithm proposed by Dempster *et al*. (1997); Wong (1998) and Christou & Fokianos (2014) for non-linear, complex and complicated for conditional Maximum Likelihood (ML).

Let $X = \left\{X_1, X_2, \cdots, X_n\right\}^T$; $\frac{1}{\sigma_k^2} = \left\{\frac{1}{\sigma_0^2}, \frac{1}{\sigma_1^2}, \cdots, \frac{1}{\sigma_k^2}\right\}^T = \left\{\alpha_0, \alpha_1 \cdots, \alpha_k\right\}^T = \alpha_k$

$\mu_k\sigma_k^2 = \left\{\mu_0\sigma_0^2, \mu_1\sigma_1^2, \cdots, \mu_k\sigma_k^2\right\}^T = \left\{\beta_0, \beta_1, \cdots, \beta_k\right\}^T = \beta_k$

$\phi_k = \left\{\phi_{k0}, \phi_{k1}, \cdots, \phi_{k\,p_k}\right\}^T; \lambda_k = \left\{\lambda_0, \lambda_1, \cdots, \lambda_k\right\}^T$

Let "η" be the latent variable such that

$$\eta_{i,t} = \begin{cases} 1 & \textit{provided } X_t \textit{ emaanted from } i^{th} \textit{ state} \\ 0, & \textit{otherwise} \end{cases}$$

Let $\Psi = \left\{\phi_k, \lambda_k, \frac{1}{\sigma_k^2}, \mu_k\sigma_k^2\right\}^T$ be the universal space.

Given η_t, then the complete Gamma distribution for the latent and random variables of $\left(X_t, \eta_t\right)$ gives,

$$L_t(\Psi)=\prod_{k=1}^{K}\left[\lambda_k\,\frac{x_t^{\frac{1}{\sigma^2}-1}e^{-\frac{x_t}{(\sigma^2\mu)}}}{(\sigma^2\mu)^{\frac{1}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}\right]^{\eta_{kt}} \tag{35}$$

Maximizing the conditional log-likelihood function of $L(\Psi)$ gives

$$L(\Psi)=\sum_{t=L+1}^{n}\left[\sum_{k=1}^{K}\eta_{kt}\log(\lambda_k)+\sum_{k=1}^{K}\eta_{kt}\left(\frac{1}{\sigma_k^2}-1\right)\log\frac{X_t}{\mu_k\sigma_k^2}-\sum_{k=1}^{K}\frac{\eta_{kt}X_t}{\mu_k\sigma_k^2}-\eta_{kt}\sum_{k=1}^{K}\Gamma(\frac{1}{\sigma_k^2})\log(\mu_k\sigma_k^2)\right]$$

$$L(\Psi)=\sum_{t=L+1}^{n}\left[\sum_{k=1}^{K}\eta_{kt}\log(\lambda_k)+\left(\sum_{k=1}^{K}\frac{\eta_{kt}}{\sigma_k^2}-\sum_{k=1}^{K}\eta_{kt}\right)\log\frac{X_t}{\mu_k\sigma_k^2}-X_t\sum_{k=1}^{K}\frac{\eta_{kt}}{\mu_k\sigma_k^2}-\sum_{k=1}^{K}\Gamma(\frac{1}{\sigma_k^2})\log(\mu_k\sigma_k^2)\eta_{kt}\right]$$

$$L(\Psi)=\sum_{t=L+1}^{n}\left[\sum_{k=1}^{K}\eta_{kt}\log(\lambda_k)+\log X_t\left(\sum_{k=1}^{K}\frac{\eta_{kt}}{\sigma_k^2}\right)\log\frac{1}{\mu_k\sigma_k^2}-\log X_t\left(\sum_{k=1}^{K}\eta_{kt}\right)\log\frac{1}{\mu_k\sigma_k^2}-\right.$$

$$\left.X_t\sum_{k=1}^{K}\frac{\eta_{kt}}{\mu_k\sigma_k^2}-\sum_{k=1}^{K}\eta_{kt}\log\mu_k\sigma_k^2\Gamma(\frac{1}{\sigma_k^2})\right] \tag{36}$$

Where $X_t=\phi_{k0}+\phi_{k1}x_{t-1}+\cdots+\phi_k x_{t-p_k}=\phi_{k0}\sum_{i=1}^{p_k}\phi_{ki}x_{t-ki}$

1st derivatives of $L(\Psi)$ w.r.t each of the parameter in the universal space gives,

$$\nabla_{\lambda_k}L(\Psi)=\sum_{t=L+1}^{n}\left(\frac{\eta_{kt}}{\lambda_k}-\frac{\eta_{Kt}}{\lambda_K}\right) \tag{37}$$

$$\nabla_{\phi_{k0}}L(\Psi)=\sum_{t=L+1}^{n}\left(\left(\frac{1}{\sigma_k^2}\eta_{Kt}\right)\log\frac{1}{\mu_k\sigma_k^2}-\frac{1}{X_t}\left(\eta_{Kt}\right)\log\frac{1}{\mu_k\sigma_k^2}-\frac{\eta_{Kt}}{\mu_k\sigma_k^2}\right) \tag{38}$$

$$\nabla_{\phi_{ki}}L(\Psi)=\sum_{t=L+1}^{n}X_{t-i}X_t\left(\frac{1}{\sigma_k^2}\log\frac{1}{\mu_k\sigma_k^2}-\log\frac{1}{\mu_k\sigma_k^2}-\frac{1}{X_t\mu_k\sigma_k^2}\right) \tag{39}$$

$$\nabla_{\alpha_k}L(\Psi)=\nabla_{\frac{1}{\sigma_k^2}}L(\Psi)=\sum_{t=L+1}^{n}\eta_{kt}\left(\log\frac{X_t}{\mu_k\sigma_k^2}-\frac{\log\mu_k\sigma_k^2}{\partial\log\Gamma\left(\frac{1}{\sigma_k^2}\right)}\right) \tag{40}$$

$$\nabla_{\beta_k}L(\Psi)=\nabla_{\mu_k\sigma_k^2}L(\Psi)=\sum_{t=L+1}^{n}\frac{\eta_{kt}}{\mu_k\sigma_k^2}\left(\log X_t-\frac{1}{\sigma_k^2}\log X_t+\frac{X_t}{\mu_k\sigma_k^2}-\log\Gamma(\Psi)\right) \tag{41}$$

$\forall\ k = 1, 2, \cdots, K; \quad i = 0, \cdots, p_k$

Second derivatives of $L(\Psi)$ w.r.t each of the parameter in the universal space shall be calculated via a function of the function of a random variable x_t at time $"t"$ and counter $"j"$

$$h(x_t, j) = \begin{cases} 1 & for\ i = 0 \\ x_{t-i} & j > 0 \end{cases}$$

So,

$$\nabla^2_{\lambda_k} L(\Psi) = -\sum_{t=L+1}^{n} \left(\frac{\lambda_{K,t}}{\lambda_K^2} - \frac{\eta_{k,t}}{\lambda_k^2} \right) \tag{42}$$

$$\nabla_{\lambda_k} L(\Psi) \nabla_{\lambda_j} L(\Psi) = -\sum_{t=L+1}^{n} \left(\frac{\eta_{k,t}}{\lambda_k^2} \right) \quad \forall\ k \neq j \tag{43}$$

$$\nabla^2_{\phi_{k0}} L(\Psi) = -\sum_{t=L+1}^{n} \frac{\eta_{kt}}{h(x_{t-i})^2} \log \frac{1}{\mu_k \sigma_k^2} \tag{44}$$

$$\nabla_{\phi_{koi}} L(\Psi) \nabla_{\phi_{koj}} L(\Psi) = -\sum_{t=L+1}^{n} \frac{\eta_{kt}}{h(x_{t-i}) h(x_{t-j})} \log \frac{1}{\mu_k \sigma_k^2} \ \forall\ i \neq j \tag{45}$$

$$\nabla^2_{\phi_{ki}} L(\Psi) = -\sum_{t=L+1}^{n} h(x_{t-i})^2\, X_t\, \eta_{kt} \left(\frac{1}{h(x_{t-i}) X_t\, \mu_k \sigma_k^2} + \log \frac{1}{\mu_k \sigma_k^2} - \frac{1}{\sigma_k^2} \log \mu_k \sigma_k^2 \right) \tag{46}$$

$$\nabla_{\phi_{ki}} L(\Psi) \nabla_{\phi_{kj}} L(\Psi) = -\sum_{t=L+1}^{n} h(x_{t-i})^2\, h(x_{t-j})^2\, X_t\, \eta_{kt} \times$$

$$\left(\frac{1}{h(x_{t-i}) h(x_{t-j}) X_t\, \mu_k \sigma_k^2} + \log \frac{1}{\mu_k \sigma_k^2} - \frac{1}{\sigma_k^2} \log \frac{1}{\mu_k \sigma_k^2} \right) \forall\ k \neq j \tag{47}$$

$$\nabla_{\alpha_k\, \alpha_k} L(\Psi) = \nabla_{\frac{1}{\sigma_k^2} \frac{1}{\sigma_k^2}} L(\Psi) = -\sum_{t=L+1}^{n} \frac{\log \mu_k \sigma_k^2}{\partial^2 \log \Gamma\left(\frac{1}{\sigma_k^2} \right)} \tag{48}$$

$$\nabla_{\alpha_k} L(\Psi) \nabla_{\alpha_j} L(\Psi) = \nabla_{\frac{1}{\sigma_k^2}} L(\Psi) \nabla_{\frac{1}{\sigma_j^2}} L(\Psi) = -\sum_{t=L+1}^{n} \frac{\log \mu_k \sigma_k^2}{\partial \log \Gamma\left(\frac{1}{\sigma_k^2} \right) \partial \log \Gamma\left(\frac{1}{\sigma_j^2} \right)} \forall\ k \neq j \tag{49}$$

$$\nabla_{\beta_k} L(\Psi)\nabla_{\beta_k} L(\Psi) = \nabla_{\mu_k\sigma_k^2} L(\Psi)\nabla_{\mu_k\sigma_k^2} L(\Psi) = -\sum_{t=L+1}^{n} \frac{1}{\left(\mu_k\sigma_k^2\right)^2}\left(\eta_{kt} + X_t\right) \qquad (50)$$

$$\nabla_{\beta_k} L(\Psi)\nabla_{\beta_j} L(\Psi) = \nabla_{\mu_k\sigma_k^2} L(\Psi)\nabla_{\mu_j\sigma_j^2} L(\Psi) = -\sum_{t=L+1}^{n} \frac{1}{\left(\mu_k\sigma_k^2\right)^2\left(\mu_j\sigma_j^2\right)^2}\left(\eta_{kt} + X_t\right)\forall k \neq j \qquad (51)$$

Letting, $$H_{nm} = E\left[-\nabla_{\lambda_k}^2 L(\Psi)\right]_{H_{nm}\ (k,j)} = \sum_{t=L+1}^{n}\left(\frac{\eta_{K,t}}{\lambda_K^2} - \frac{\eta_{k,t}}{\lambda_k^2}\right) \qquad (52)$$

Where H_{mm} is the m-square Hessian matrix

$$\Rightarrow H_{nm\ (k,j)} = \sum_{k=1}^{K}\frac{\gamma_{kt}}{\lambda_k^2}$$

$$H_{nk} = E\left[-\nabla^2_{\left(\phi_{k0},\phi_{ki},\frac{1}{\sigma_k^2},\mu_k\sigma_k^2,\sigma_k\right)} L(\Psi)/\Psi, X\right] = H_{ak}(k,j) \qquad (53)$$

$$H_{nk}(i+1, j+1) = \sum_{k=1}^{K}\frac{\gamma_{k,t} h(x_t, i) h(x_t, j)}{\mu_k\sigma_k^2} \qquad (54)$$

Employing the EM-algorithm procedure for estimating Ψ, the universal parameter space via the $L(\Psi)$ in equation (35).

The E-step, assuming the universal parameter space Ψ is available, then the neglected values for the latent data $\left(\eta_{L,t}\right)$ is then substituted by the means of the observed values $\left(\overline{X}_t\right)$ their parameters. Then $\gamma_{k,t}$ can be calculated by the procedure of Bayes' theorem as stated below:

$$\gamma_{k,t} = \frac{\lambda_k \dfrac{x_t^{\frac{1}{\sigma^2}-1} e^{-\frac{x_t}{(\sigma^2\mu)}}}{(\sigma^2\mu)^{\frac{1}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}}{\sum_{k=1}^{K}\lambda_k \dfrac{x_t^{\frac{1}{\sigma^2}-1} e^{-\frac{x}{(\sigma^2\mu)}}}{(\sigma^2\mu)^{\frac{1}{\sigma^2}}\Gamma(\frac{1}{\sigma^2})}} \qquad (55)$$

$$\forall\, k = 1,\cdots,K; \quad t = L+1,\cdots,n; \quad \eta_{k,t} = \gamma_{k,t}$$

The M-step can be obtained via equation (55) below:

$$\hat{\lambda}_k = \sum_{t=L+1}^{n} \frac{\gamma_{k,t}}{n-L} \tag{56}$$

∋ that the embedded individual parameters in the universal space can be estimated via a system of equations or via Newton-Raphson iterative procedure.

$$\Psi_k^{r+1} = \Psi_k^r + \left[E\left(-n\nabla_{\Psi_k}^2 L(\Psi) /_{\left(\phi_{k0},\phi_{ki},\frac{1}{\sigma^2},\mu_k\sigma_k^2,\sigma_k\right)} \right) \right]^{-1} \times \nabla_{\Psi_k} L(\Psi) /_{\left(\phi_{k0},\phi_{ki},\frac{1}{\sigma^2},\mu_k\sigma_k^2,\sigma_k\right)} \tag{57}$$

The estimate of Ψ is then iterate by these two steps until convergence is reached.

3.8. Standard Errors for the Associated Parameters

Adopting the method of estimating the dispersion matrix of any parameter space as defined by Louis (1982) and Ghosh *et al.* (2006). Let the Ψ estimate of the inverse observed information be H^{-1}, then,

$$H = H_n - H_m = E\left[-n\nabla_{\Psi}^2 L(\Psi)\right]_{\hat{\Psi}} - \sigma^2\left[n\nabla_{\Psi} L(\Psi)\right]_{\hat{\Psi}} \tag{58}$$

Where, H could be obtain from the information matrix, H_n , and the missing information matrix H_m such the complete information matrix is

$$H_n = \begin{pmatrix} H_{no} & \cdots & \cdots & \\ \cdots & H_{n1} & \cdots & \\ \cdots & & \ddots & \cdots \\ \cdots & & \cdots & H_{nk} \end{pmatrix} \tag{59}$$

Where H_n is a block diagonal matrix , a square matrix of (k-1) dimension. Also,

$$H_m = \begin{pmatrix} H_{maa} & \cdots & \cdots \\ & H_{mbb} & \cdots \\ & \ddots \ddots & \vdots \\ & & H_{mkk} \end{pmatrix} \quad (60)$$

is a symmetric diagonal matrix

where,

$$H_{maa} = \sigma^2 \left(\frac{\partial L(\Psi)}{\partial \lambda_k} = \nabla_{\lambda_k} L(\Psi)\ ; \Psi, X_t \right) \quad (61)$$

$$H_{mkk} = \sigma^2 \left(\frac{\partial L(\Psi)}{\partial \Theta_{\left(\phi_{k0}, \phi_{ki}, \frac{1}{\sigma^2}, \mu_k \sigma_k^2, \sigma_k\right)}} = \nabla_{\left(\phi_{k0}, \phi_{ki}, \frac{1}{\sigma^2}, \mu_k \sigma_k^2, \sigma_k\right)} L(\Psi); \Psi, X_t \right) \quad (62)$$

2.9 Some Steps Ahead Prediction for the GAMMAR $(k; p_1, p_2, \cdots, p_k)$ Model

From equation (17), a step a-head prediction could be obtained. Assuming an horizon of $"t"$ different from the conventional series counter of $"t"$ is needed to generalized starting one-step ahead, two-step ahead etc.. Assuming further also that the highest lag in the combined regimes is three, then,

$$\widehat{x}_{t+1/t} = E\left[\widehat{x}_{t+1} / \widehat{x}_t, \widehat{x}_{t-1}, \widehat{x}_{t-2}, \widehat{x}_{t-3}, \cdots \right] \quad (63)$$

Using equation (1)

$$\widehat{x}_{t+1/t} = \lambda_1 \left[\widehat{\phi}_{11} \widehat{x}_t + \widehat{\phi}_{12} \widehat{x}_{t-1} \right] + \lambda_2 \left[\widehat{\phi}_{21} \widehat{x}_t + \widehat{\phi}_{22} \widehat{x}_{t-1} \right] + \lambda_3 \left[\widehat{\phi}_{31} \widehat{x}_t + \widehat{\phi}_{32} \widehat{x}_{t-1} \right] \quad (64)$$

In a similar vein, the two-step ideal foretelling of the estimated series $\widehat{x}_{t+2/t+1}$

$$\widehat{x}_{t+2/t} = E\left[E\left(\widehat{x}_{t+2/t} / \widehat{x}_{t+1}, \widehat{x}_t, \widehat{x}_{t-1}, \widehat{x}_{t-2}, \cdots \right) / \widehat{x}_t, \widehat{x}_{t-1}, \widehat{x}_{t-2}, \cdots \right] \quad (65)$$

Using evaluation from equation (62) to simplify

$$\widehat{x}_{t+2/t} = \left[\lambda_1 \widehat{\phi}_{11} + \lambda_2 \widehat{\phi}_{21} + \lambda_3 \widehat{\phi}_{31} \right] \widehat{x}_{t+1/t} + \left[\lambda_1 \widehat{\phi}_{12} + \lambda_2 \widehat{\phi}_{12} + \lambda_2 \widehat{\phi}_{22} + \lambda_3 \widehat{\phi}_{32} \right] \widehat{x}_t \quad (66)$$

For three-step ahead

$$\widehat{x}_{t+3/t} = E\left[E\left(\widehat{x}_{t+3} / \widehat{x}_{t+2}, \widehat{x}_{t+1}, \widehat{x}_t, \widehat{x}_{t-1}, \cdots \right) / \widehat{x}_t, \widehat{x}_{t-1}, \widehat{x}_{t-2}, \cdots \right] \quad (67)$$

This is gives,

$$\hat{x}_{t+3/t} = \left[\lambda_1\hat{\phi}_{11} + \lambda_2\hat{\phi}_{21} + \lambda_3\hat{\phi}_{31}\right]\hat{x}_{t+2/t} + \left[\lambda_1\hat{\phi}_{12} + \lambda_2\hat{\phi}_{12} + \lambda_2\hat{\phi}_{22} + \lambda_3\hat{\phi}_{32}\right]\hat{x}_{t+1/t} \quad (68)$$

In a general form,

$$\hat{x}_{t+i/t} = E\left[E\left(\hat{x}_{t+i} / \hat{x}_{t+i-1}, \hat{x}_{t+i-2}, \hat{x}_{t+i-3}, \hat{x}_{t+i-4}, \cdots\right) / \hat{x}_t, \hat{x}_{t-1}, \hat{x}_{t-2}, \cdots\right] \quad (69)$$

$$\hat{x}_{t+i/t} = \left[\lambda_1\hat{\phi}_{11} + \lambda_2\hat{\phi}_{21} + \lambda_3\hat{\phi}_{31} + \cdots + \lambda_i\hat{\phi}_{i1}\right]\hat{x}_{t+i-1/t} +$$

$$\left[\lambda_1\hat{\phi}_{12} + \lambda_2\hat{\phi}_{12} + \lambda_2\hat{\phi}_{22} + \lambda_3\hat{\phi}_{32} + \cdots + \lambda_i\hat{\phi}_{i2}\right]\hat{x}_{t+i-2/t} + \cdots +$$

$$\left[\lambda_1\hat{\phi}_{1i} + \lambda_2\hat{\phi}_{1i} + \lambda_2\hat{\phi}_{2i} + \lambda_3\hat{\phi}_{3i} + \cdots + \lambda_i\hat{\phi}_{ii}\right]\hat{x}_{t/t} \quad (70)$$

Where

$$e_t = x_t - \hat{x}_t$$

$$e_{t-1} = x_{t-1} - \hat{x}_{t-1}$$

$$\vdots \qquad \vdots$$

$$e_{t-i} = x_{t-i} - \hat{x}_{t-i}$$

2.10 Levinson-Durbin Method of Sub-setting of Coefficients of GAMMAR $(k : p_1, \ldots, p_k)$ Model

The Recursion method of the Levinson-Durbin will be adopted in carrying out the sub-setting of the $GAMMAR(k : p_1, \ldots, p_k)$ if peradventure there exist problem of parsimony or the need arises to verify the selected order in each regime are the optimal to the fitted components of the *k*-regimes. According to McLeod & Zhang (2008), Hannan & Quinn (1979), the recursion of the Levinson-Durbin is an ordered recursion of M-steps, which can either be a forward or a backward step of combining the coefficients via a linear complexity. Incorporating into the $GAMMAR(k : p_1, \ldots, p_k)$, since each of the regime is a linear of individual AR, it implies each of the regime has mean s $\mu_k^{'}$ written as

$$\phi_k(B_k) = \left(x_{tk}^{'} - \mu_k^{'}\right) = \left(x_{tk}^{'} - (\alpha\beta)_k^{'}\right) = b_{tk} \quad (71)$$

$$\ni \quad \phi(B_k) = 1 - \phi_{1k}B - \phi_{2k}B -, \cdots, \phi_{pk}B$$

For $"t"$ which ranges from $1,\cdots,n_k$

$$\ni \quad x_{tk}^{'} \sim \left((\alpha\beta)_k^{'},(\alpha\beta^2)_k^{'}\right)=\left((\mu)_k^{'},(\mu^2\sigma^2)_k^{'}\right) \tag{72}$$

Such that the sub-setting via the Levinson-Durbin recursion is defined as

$$\phi_{mi,\rho m+1}=\phi_{mi,\rho m}-\phi_{mi+1,\rho m+1}\phi_{mi+1-i,\rho m} \qquad for \;\; i=1,\cdots,m$$

With a changeover of 1-1 mapping of

$$B:\left(\lambda_1,\lambda_2,\cdots,\lambda_{pm}\right)\rightarrow\left(\phi_{k1},\phi_{k2},\cdots,\phi_{kp}\right) \tag{73}$$

With log-likelihood from the Gamma distribution a defined in equation (36)

$$L(\Theta)=\sum_{t=L+1}^{n}\left[\sum_{k=1}^{K}\eta_{kt}\log(\lambda_k)+\sum_{k=1}^{K}\lambda_{kt}\left(\frac{1}{b_{tk}}-1\right)\log\frac{b_{tk}X_t}{\mu_k^{'}}-\sum_{k=1}^{K}\frac{\eta_{kt}b_{tk}X_t}{\mu_k^{'}}-\eta_{kt}\sum_{k=1}^{K}\Gamma(\frac{1}{b_{tk}})\log(\frac{\mu_k^{'}}{b_{tk}})\right] \tag{74}$$

According to McLeod & Zhang (2016) and Hossain (2015). the technique of selection will solely lie on the modified BIC gotten from $\left[-2L(\Theta)+(\rho m\times m)\log(n)\right]$ where $"n"$ is the length of the whole regime series, $(\rho m\times m)$ is the number of the AR parameters of the combined $GAMMAR(k:p_1,\ldots,p_k)$ with $AR_\lambda\left(m_1,m_2,\cdots,m_p\right)$ and $\hat{\sigma}_b^2$ approximately equals, $b_0\left(1-\hat{\phi}_{m_1,m_1}\right)\cdots\left(1-\hat{\phi}_{\rho m,\rho m}\right)$ for $"b_0"$ the sample variation. So, the modified BIC equals

$$BIC_\lambda\left(m_1,\ldots,\rho m\right)=n\log L(\Theta)+(\rho m\times m)\log(n) \tag{75}$$

Such that the below algorithm could use in obtaining minimum BIC_λ model:

1. Choose the ρm, the highest order for each of the AR in each of the regime via autocorrelation. Count the number of parameters estimated initially.

2. Distinguished the absolute values of AR coefficients that are less than one, which are stationary coefficients of AR in a descending order of magnitude.

3. Evaluate the $BIC_\lambda\left(m_1,\ldots,\rho m\right)$ for $\rho m=1,\cdots,\rho m_k$ and choose the minimum BIC_λ model for the betterment of forecasting evaluation and performance.

3. Conclusion

We have seen above that the transmuted Gamma distribution was a proper probability density function for its to substantially drive the Gamma mixture autoregressive model with k-components. The first and second-order stationarity process of the $GAMMAR(k : p_1, \ldots, p_k)$ fall within the unit circle to confirm some of it k-components stationarity. Mixture of any stationary k-component(s) with any non-stationary component(s) makes the whole process stationary. The Levinson-Durbin M-steps recursion via a forward or a backward step was incorporated into the $GAMMAR(k : p_1, \ldots, p_k)$ model to enable carve-out some derivations for some steps ahead prediction.

4. References

[1]. G.N. Boshnakov (2006). *Prediction with Mixture Autoregressive Models*. Research Report No. 6, 2006, Probability and Statistics Group School of Mathematics, The University of Manchester, London.

[2]. V. Christou & K. Fokianos (2014). Quasi-likelihood inference for negative binomial time series models. *Journal of Time Series Analysis*, 35(1), 55– 78. doi:10.1111/jtsa.12050.

[3]. N. Dempster, A.P.Laird, & D.Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39, 1-8.

[4]. W. C. Forrest, A.S. Marc, & , E.W. Robert (2015). Sex, lies and self reported counts: Bayesian mixture models for heaping in longitudinal count data via birth–death processes. *The Annals of Applied Statistics*, 9(2), 572–596. doi:10.1214/15-AOAS809.

[5]. H. Ghosh, M.A.Iquebal, & , P. Sankhya (2006). On mixture non-linear time series modeling and forecasting for ARCH Effects. *The Indian Journal of statistics*, 68(1), 111-129.

[6]. E.J. Hannan & B.G. Quinn (1979) The determination of the order of an autoregression. *Journal of the Royal Statistical Society*, B 41,190–195.

[7]. S. Hossain (2015). Complete Bayesian analysis of some mixture time series models. A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Engineering and Physical Sciences School of Mathematics.

[8]. L. Kalliovirta, M. Meitz, & P. Saikkonen (2016). A Guassian Mixture Vector Autoregressive. *Journal of econometrics*, Vol. 192, 485-498.

[9]. N.D. Le, R.D. Martin, & A.E. Raftery (1996): Modeling fiat stretches, bursts and outliers in time series using mixture transition distribution models. *Journal of American Statistical Association*, 9 (1), 1504-1514.

[10]. T.A. Louis (1982). Finding the observed information matrix when using the EM algorithm. *Journal of Royal Statistics Society*, B44, 226-233.

[11]. A.I. McLeod & Y. Zhang (2016). Subset autoregression: a new approach. The University of Western Ontario and Acadia University. http//:www.arXiv:1611.01370v1.

[12]. A. I. McLeod & Y. Zhang (2008). Improved Subset Autoregression: with R Package. *Journal of Statistical Software*, 28(2) 1- 28. http://www.jstatsoft.org/

[13]. A.S. Nastic (2014). On suitability of negative binomial marginal and geometric counting sequence in some applications of combined INAR(*p*) model. *Facta Universitatis (NIS) Ser. Math. Inform*, 29(1), 25–42.

[14]. R.O. Olanrewaju, A.G. Waititu, & L.A. Nafiu (2021). Frechet Random Noise for k-Regime-Switching Mixture Autoregressive Model. *American Journal of Mathematics and Statistics*, 11(1), 1-10. Doi:10.5923/j.ajms.20211101.01.

[15]. C.S. Wong & W.K. Li (2000): On mixture autoregressive model. *Journal of the Royal Statistical Society*, B 62, Part 1,95-115.

[16]. C.S. Wong (1998): Statistical inference for some nonlinear time series models. The HKU Scholars Hub. The University of Hong Kong. http:hdl.handle.net/10722/35431.