Fréchet Random Noise for *k*-Regime-Switching Mixture Autoregressive Model

Rasaki Olawale Olanrewaju[1,*], Anthony Gichuhi Waititu[2], Nafiu Lukman Abiodun[3]

[1]Pan African University, Institute for Basic Sciences, Technology and Innovation (PAUSTI), Kenya
[2]Department of Mathematical Sciences, Jomo Kenyatta University of Agriculture and Technology, Kenya
[3]Department of Mathematics and Statistics, Kabale University, Kabale, Uganda

Abstract   This paper describes Fréchet distribution as a random noise for capturing multimodalities, regime-switching and change-points attributed to uniformly time-varying series via causality of fluctuations, extreme values and heavy-tailed time series. Fréchet Mixture Autoregressive (FMAR) model of *k*-regime-switching, denoted by $FMAR(k; p_1, p_2, \ldots, p_k)$ was developed and Expectation-Maximization (EM) algorithm was used as a method of parameter estimation for the embedded coefficients of AR of *k*-mixing weights and lag p_k. The limiting distribution of the $FMAR(k; p_1, p_2, \ldots, p_k)$ model via Gnedenko-Fisher Tippet limiting property was derived to asymptotically approach an exponential function.

Keywords   Fréchet distribution, Expectation-Maximization, Gnedenko-Fisher Tippet, *k*-regime-switching, Mixture Autoregressive, Multimodalities

1. Introduction

Some of the problems of time series are cyclical fluctuation and regime-switching, full range of shape changing predictive distributions (multimodalities), change like behavior and inability to handle cycles traits that were characterized by the data series [2]. Some non-linear time series models such as Threshold Auto-Regressive (TAR), and Self Exciting Threshold Autoregressive (SETAR) were proposed contain the made mentioned stylized properties, but were limited to two regime-shifting [8]. Mixture Autoregressive (MAR) model $MAR(k; p_1, p_2, \ldots, p_k)$ proposed by [14] captured and modeled the aforementioned stylized traits with strictly Gaussian or normal distributed error term.

The problem of regime-switching model via mixture models are yet to be fully and completely addressed by $MAR(k; p_1, p_2, \ldots, p_k)$ model with Gaussian random noise [1].

According to [4], structural $MAR(k; p_1, p_2, \ldots, p_k)$ time-series model has been statistically deficiency due to inadequate to take into consideration multimodalities, large fluctuations, heavy-tail and non-Gaussian distributional random noise that are causal traits of the multimodalities and tenure-changing. However, the MAR model with Gaussian marginal distribution is noted to inadequately capture

extreme events that are usually the causal of distortion that leads to multimodalities and tenure-changing [7]. Consequently, the stated problem of MAR model with Gaussian error term would be addressed by Fréchet random noise, being one of the three Extreme Value Distributions (EVDs). In other words, Fréchet Mixture Autoregressive $FMAR(k; p_1, p_2, \ldots, p_k)$ model with Fréchet error term will be developed and derived as a substitute for MAR to capture extreme events (outliers) that are causes of large fluctuations in arising due to abnormal kurtosis and skewness. Hence, this article develops, ascertains and estimates embedded parameters (AR coefficients, mixing weights and Fréchet coefficients) via Expectation-Maximization algorithm of the $FMAR(k; p_1, p_2, \ldots, p_k)$ process with *k*-regime-switching. In addition, the limiting distribution (asymptotic property) of the $FMAR(k; p_1, p_2, \ldots, p_k)$ model will be ascertained via Gnedenko-Fisher Tippet limiting behavior as well as the ergodic of the FMA process. The mean and variance of the $FMAR(k; p_1, p_2, \ldots, p_k)$ process will be ascertained as well.

2. Literature Review

The history of mixture time series model was originated and propounded by [15] in the late 1900s, when he used normal mixture to model outliers and employ the method of moments to decompose normal mixtures. He deduced that Autoregressive Integrated Moving Average (ARIMA), Self-Exciting Threshold Autoregressive (SETAR) model, and other nonlinear models can only depict a unimodal

predictive time series random noise distribution. He affirmed that mixture model would be needed for stylist, cyclical, switching tenure properties to be revealed. Afterwards, [3] predicted with the MAR model and deduced that MAR model possessed the attractive property in such a way that spatial attributes of the conditional density of its prediction contingents upon on the immediate past events of the process. Additionally, he maintained that the multiple steps' distribution would be a collection (mixture) of normal distributions if and only if the source MAR model is a mixture of normal densities.

However, despite propounded of all these models, no model studied the stationarity and ergodicity traits of MAR related models. It was in line of this deficiency that [8] propounded and affirmed the theoretical advantage of Gaussian mixture autoregressive (GMAR) conditions for stationarity and ergodicity to be met. They claimed that GMAR properties were straightforward to establish. In addition, they maintained that p[th] order model can be explicitly expressed in a (Optimal order (p) +1)– lengthy stationary distribution with ascertained constant mixing weights to connote mixture of normal probability distributions.

In advancement, [7] proposed a novel nonlinear Vector Autoregressive (VAR) otherwise called Gaussian Mixture Vector autoregressive (GMVAR) model. They developed and explained its asymptotic theory of maximum likelihood whose usefulness is vital when dealing with bivariate time series settings.

Furthermore, [6] proposed a new nonlinear time series via a vector incorporation into the MAR model called Vector-MAR (VAR) model otherwise known as Gaussian Mixture Vector Autoregressive (GMVAR) model. They affirmed and claimed that the GMVAR model can be categorized as member of the mixture VAR models because of its normally distributed error term, constructed and developed for studying nonlinear models with time-varying series that exhibited tenure changing traits. They maintained that primarily the GMVAR model differs from the variety of VAR models in the meaning of the mingling proportions that regularize the tenure likelihoods.

[11] provided an up-to-date theoretical and methodological developments on finite mixture models due to its flexibility increasingly exploitation and applications. They adopted the semiparametric technique way of parameter estimation via the EM algorithm to carve out the distributional shape changing, group-structure and cluster analysis in the data captured to the model. Proposed models reviewed were formulated to capture, and correct the stylized properties; entire outrange of pattern switching distributions (multimodalities), change points like behavior, regime switching (capable of handling recurring periodical sequence), and time-varying volatilities (conditional means-variances), but no method/model yet to cater for heavy-tailed, long-memory and non-Gaussian MAR model. Hence, a full sketch $FMAR(k; p_1, p_2, \ldots, p_k)$ model will be developed such that its mean and variance will as ascertained as well as its parameter estimation, limiting distribution and ergodic process.

3. Specification of the Fréchet Mixture Autoregressive Model

[10] and [13] defined a *k*-component of Mixture Autoregressive (MAR) model $MAR(K : p_1, \ldots, p_k)$ to be

$$F\left(x_{(t)} / f_{t-1}\right) = P\left(X_{(t)} \leq x / Ft-1\right) = \sum_{i=1}^{k} w_k \; \Phi\left(\frac{x_t - \phi_{k,0} - \phi_{k,1}x_{t-1} - \cdots - \phi_{k,pk} \; x_{t-pk}}{\sigma_k}\right) \quad (1)$$

Rewriting the *k*-component of MAR model in (1) in Fréchet Mixture Autoregressive $FMAR(k; p_1, p_2, \ldots, p_k)$ gives

$$X_t = \sum_{i=1}^{k} \left(w_k, \alpha_k, \beta_k\right) \; \Phi\left(\frac{x_t - \phi_{k,0} - \phi_{k,1}x_{t-1} - \cdots - \phi_{k,pk} \; x_{t-pk}}{\sigma_k}\right) \quad (2)$$

where; $X_t = \phi_{k,0} + \phi_{k,1}x_{t-1} + \cdots + \phi_{k,p}x_{t-pk} + \varepsilon_t^{(k)}$ $\varepsilon_t^{(k)} = x_t - \phi_{k,0} - \phi_{k,1}x_{t-1} - \cdots - \phi_{k,p}x_{t-pk}$

$\mu_{k,t} = \phi_{k,0} + \phi_{k,1}x_{t-1} + \cdots + \phi_{k,p}x_{t-pk}$

Otherwise,

$$X_t = \begin{cases} \phi_{1,0} + \sum_{i=1}^{p1} \phi_{1,i} x_{t-1} + \varepsilon_t^{(1)} & \text{with probability } \omega_1 \; \& \, (\alpha_1, \beta_1) \\ \phi_{2,0} + \sum_{i=1}^{p2} \phi_{2,i} x_{t-2} + \varepsilon_t^{(2)} & \text{with probability } \omega_2 \; \& \, (\alpha_2, \beta_2) \\ \vdots \qquad\qquad \vdots \qquad\qquad\quad \vdots \\ \phi_{k,0} + \sum_{i=1}^{pk} \phi_{k,i} x_{t-p} + \varepsilon_t^{(k)} & \text{with probability } \omega_k \; \& \, (\alpha_k, \beta_k) \end{cases} \quad (3)$$

Where, $\phi_{pk} \in (0,1)$, for $k = 1,\ldots\ldots, K$, $p \geq 1$. For mixing weights $\omega_1 + \ldots + \omega_k \approx 1$, $\omega_i > 0$, for $k = 1 \cdots K$ $\Phi(.)$ is the

Cumulative Distribution Function of the standard distributions, where $\varepsilon_t^{(k)}$ & $X_t \approx g(x_t; \alpha, \beta) = \dfrac{\alpha}{\beta}\left(\dfrac{\beta}{x_t}\right)^{\alpha+1} \exp\left(-\left(\dfrac{\beta}{x_t}\right)^{\alpha}\right)$,

with mean $E(X) = \beta\Gamma\left(1 - \dfrac{1}{\alpha}\right)$ and variance $\sigma^2 = \beta^2\left[\Gamma\left(1 - \dfrac{2}{\alpha}\right) - \left(\Gamma\left(1 - \dfrac{1}{\alpha}\right)\right)\right]^2$ *for* $\alpha > 2$.

3.1. The Conditional Mean and Variance for FMA Model

The conditional mean and variance for FMA model of x_t given the previous preceding values is as follow:

$$E\left(x_t \,/\, g(\alpha,\beta)_{t-1}\right) = \sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\left(\phi_{k0} + \phi_{k1}x_{t-1} + \cdots + \phi_{kpk}x_{t-pk}\right) = \sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\,\mu_{k,t} \tag{4}$$

since, $\mu_{k,t} = \sum_{k=1}^{K}\left(\phi_{i0} + \sum_{p=1}^{pk}\phi_{i1}x_{t-pk}\right) = \phi_{k0} + \phi_{k1}x_{t-1} + \cdots + \phi_{kpk}x_{t-pk}$

The mean of Fréchet is $E(X) = \beta\Gamma\left(1 - \dfrac{1}{\alpha}\right)$, therefore

$$E\left(x_t \,/\, g(\alpha,\beta)_{t-1}\right) = \sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\,\mu_{k,t} = \sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\,\beta_k\Gamma\left(1 - \dfrac{1}{\alpha_k}\right) \tag{5}$$

which relies on previous preceding values of the time series & $g(\alpha,\beta)$ is the PDF of Fréchet.

$$E(x_t^2 \,/\, g(\alpha,\beta)x_{t-1}) = \sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\sigma_k^2 + \sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\mu_{k,t}^2$$

So,

$$Var(x_t \,/\, g(\alpha,\beta)x_{t-1}) = \sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\sigma_k^2 + \sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\mu_{k,t}^2 - \left(\sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\mu_{k,t}\right)^2$$

But, $\left(\sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\mu_{k,t}\right)^2 = \sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\left(\mu_{k,t} - \sum_{s=1}^{K}(\omega_s,\alpha_s,\beta_s)\mu_{s,t}\right)^2$ for $s = t-1$ (previous preceding values of

the time series that relies on "t" the current values).

$$\sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\left(\mu_{k,t} - \sum_{s=1}^{K}(\omega_s,\alpha_s,\beta_s)\mu_{s,t}\right)^2 =$$

$$= \sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\mu_{k,t}^2 + \left(\sum_{s=1}^{K}(\omega_s,\alpha_s,\beta_s)\mu_{s,t}\right)^2 - 2\left(\sum_{s=1}^{K}(\omega_s,\alpha_s,\beta_s)\mu_{s,t}\right)^2$$

$$= \sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\mu_{k,t}^2 + \left(\sum_{s=1}^{K}(\omega_k,\alpha_k,\beta_k)\mu_{k,t}\right)^2$$

Therefore,

$$Var(x_t \,/\, g(\alpha,\beta)x_{t-1}) = \sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\sigma_k^2 - \left(\sum_{s=1}^{K}(\omega_k,\alpha_k,\beta_k)\mu_{k,t}\right)^2 \tag{6}$$

The expression $\sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\sigma_k^2-\left(\sum_{s=1}^{K}(\omega_k,\alpha_k,\beta_k)\mu_{k,t}\right)^2$ is a positive quantity. **Hence, this meets the non-negative property of variance.**

Since the mean and variance of Fréchet PDF are; $\beta\Gamma\left(1-\dfrac{1}{\alpha}\right)$ and $\sigma^2=\beta^2\left[\Gamma\left(1-\dfrac{2}{\alpha}\right)-\left(\Gamma\left(1-\dfrac{1}{\alpha}\right)\right)\right]^2$ *for* $\alpha>2$ respectively

Alternatively, $Var(x_t\,/\,g(\alpha,\beta)x_{t-1})$ of equation (6) can be written as

$$Var(x_t\,/\,g(\alpha,\beta)x_{t-1})=\sum_{k=1}^{K}(\omega_k,\alpha_k,\beta_k)\beta_k^2\left[\Gamma\left(1-\frac{2}{\alpha_k}\right)-\left(\Gamma\left(1-\frac{1}{\alpha_k}\right)\right)\right]^2-\left(\sum_{s=1}^{K}(\omega_k,\alpha_k,\beta_k)\beta_k\Gamma\left(1-\frac{1}{\alpha_k}\right)\right)^2$$

$$-\left(\sum_{s=1}^{K}(\omega_k,\alpha_k,\beta_k)\beta_k\Gamma\left(1-\frac{1}{\alpha_k}\right)\right)^2 \tag{7}$$

3.2. Parameter Estimation for FMA via EM Algorithm

Adopting the Expectation-Maximization (EM) algorithm propounded by [10] for non-linear and regime-switching. Let $X=\{X_1,X_2,\cdots,X_n\}$; $\alpha_k=\{\alpha_0,\alpha_1,\cdots,\alpha_k\}^T$ $\beta_k=\{\beta_0,\beta_1,\cdots,\beta_k\}^T$; $\phi_k=\{\phi_{k0},\phi_{k1},\cdots,\phi_{k\,pk}\}^T$ $\omega_k=\{\omega_0,\omega_1,\cdots,\omega_k\}^T$ for $k=1,\cdots,K$

Additionally, supposing that "S" is the unobserved random variable where S_t is a $k-$dimensional vector such that $S=\{S_1,S_2,\cdots,S_n\}$

$S_t=\{S_1,S_2,\cdots,S_t\}^T$ whose component is

$$S_{i,t}=\begin{cases}1 & if\ X_t\ originated\ from\ the\ j^{th}\ weight\,(state)\\0, & otherwise\end{cases}$$

For $1\le j\le K$, that is,

Let $\Theta=\{\phi_k,\omega_k,\alpha_k,\beta_k\}^T$ be the universal space of parameters.

Given S_t, the Fréchet distribution of the complete data (X_t,S_t) is given by

$$L_t(\Theta)=\prod_{k=1}^{K}\left[\omega_k\frac{\alpha_k}{\beta_k}\left(\frac{\beta_k}{X_t}\right)^{\alpha_k+1}\exp\left(-\left(\frac{\beta_k}{X_t}\right)^{\alpha_k}\right)\right]^{S_{kt}} \tag{8}$$

If $L(\Theta)=\sum_{t=1}^{n}L_t$ is the log-likelihood. Letting $L(\Theta)=\sum_{t=L+1}^{n}L_t$ be the joint conditional log-likelihood function for large sample size (n) for maximizing the function gives

$$L_t(\Theta)=\log\prod_{k=1}^{K}\left[\omega_k\frac{\alpha_k}{\beta_k}\left(\frac{\beta_k}{X_t}\right)^{\alpha_k+1}\exp\left(-\left(\frac{\beta_k}{X_t}\right)^{\alpha_k}\right)\right]^{S_{kt}} \tag{9}$$

$$L(\Theta)=\sum_{t=L+1}^{n}\left[\sum_{k=1}^{K}S_{kt}\log(\omega_k)+\sum_{k=1}^{K}S_{kt}\log\frac{\alpha_k}{\beta_k}+\sum_{k=1}^{K}S_{kt}\alpha_k\log\beta_k+\sum_{k=1}^{K}S_{kt}\alpha_k\log X_t+\sum_{k=1}^{K}S_{kt}\log\frac{\beta_k}{X_t}-\sum_{k=1}^{K}S_{kt}\beta_k^{\alpha_k}X_t^{-\alpha_k}\right] \tag{10}$$

Where $X_t=\phi_{k0}+\phi_{k1}x_{t-1}+\cdots\cdots+\phi_kx_{t-pk}=\phi_{k0}+\sum_{i=1}^{Pk}\phi_{ki}x_{t-ki}$ First derivatives of $L(\Theta)$ w. r. t each of the parameter gives,

$$\nabla_{w_k} L(\Theta) = \sum_{t=L+1}^{n} \left(\frac{S_{kt}}{w_k} - \frac{S_{Kt}}{w_K} \right) \text{ for } k = 1, \cdots, K-1 \tag{11}$$

$$\nabla_{\phi_{k0}} L(\Theta) = \sum_{t=L+1}^{n} S_{Kt} \left(\frac{\alpha_k - 1}{X_t} + \alpha_k \left(\beta_k^{\alpha_k} \frac{1}{X_t^{\alpha_k+1}} \right) \right) \tag{12}$$

$$\nabla_{\phi_{ki}} L(\Theta) = \sum_{t=L+1}^{n} \left(\frac{\alpha_k S_{Kt} - S_{Kt}}{X_t X_{t-i}} + \alpha_k \left(S_{Kt} \beta_k^{\alpha_k} X_t^{-\alpha_k-1} X_{t-i}^{-\alpha_k-1} \right) \right)$$

$$= \sum_{t=L+1}^{n} S_{Kt} \left(\frac{\alpha_k - 1}{X_t X_{t-i}} + \alpha_k \beta_k^{\alpha_k} \left(X_t^{-\alpha_k-1} X_{t-i}^{-\alpha_k-1} \right) \right) \tag{13}$$

$$\nabla_{\alpha_k} L(\Theta) = \sum_{t=L+1}^{n} \left(\frac{S_{kt}}{\alpha_k} + S_{kt} \log \beta_k + S_{kt} \log X_t - S_{kt} \beta_k X_t^{-\alpha_k} e^{\alpha_k} + S_{kt} e^{-\alpha_k} X_t^{-\alpha_k} \beta_k \right)$$

$$= \sum_{t=L+1}^{n} S_{kt} \left(\frac{1}{\alpha_k} + \log \beta_k + \log X_t - X_t^{-\alpha_k} \beta_k \left(e^{\alpha_k} - e^{-\alpha_k} \right) \right) \tag{14}$$

$$\nabla_{\beta_k} L(\Theta) = \sum_{t=L+1}^{n} \frac{S_{kt} \alpha_k}{\beta_k} \left(1 - \beta_k^{\alpha_k} X_t^{-\alpha_k} \right) \tag{15}$$

For $k = 1, 2, \cdots, K; \quad i = 0, \cdots, pk$

Then, the second derivatives of individual coefficient estimated from (11) - (15) depends of setting a function x_t to be a random variable at time "t" and counter "j"

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

$$\nabla^2_{\omega_k} L(\Theta) = -\sum_{t=L+1}^{n} \left(\frac{S_{K,t}}{\omega_K^2} - \frac{S_{k,t}}{\omega_k^2} \right) \tag{16}$$

$$\nabla_{\omega_k} L(\Theta) \nabla_{\omega_j} L(\Theta) = -\sum_{t=L+1}^{n} \left(\frac{S_{k,t}}{\omega_k^2} \right) \text{ for } k \neq j \tag{17}$$

$$\nabla^2_{\phi_{k0}} L(\Theta) = \sum_{t=L+1}^{n} S_{Kt} \left(\frac{\alpha_k}{h(x_{t-i})^2} - \frac{1}{h(x_{t-i})^2} + \alpha_k \left(\beta_k^{\alpha_k} \frac{1}{h(x_{t-i})^{2(\alpha_k+1)}} \right) \right) \tag{18}$$

$$\nabla_{\phi_{koi}} L(\Theta) \nabla_{\phi_{koj}} L(\Theta) = \sum_{t=L+1}^{n} \frac{S_{Kt}}{h(x_{t-i}) h(x_{t-j})} ((\alpha_k - 1) +$$

$$\alpha_k \left(\beta_k^{\alpha_k} \frac{1}{h(x_{t-i})^{\alpha_k} h(x_{t-j})^{\alpha_k}} \right) \right) \text{ for } i \neq j \tag{19}$$

$$\nabla^2_{\phi_{ki}} L(\Theta) = \sum_{t=L+1}^{n} S_{Kt} \left(-\left(\frac{\alpha_k - 1}{X_t h(x_{t-i})^2} + \alpha_k \beta_k^{\alpha_k} \left(X_t^{-\alpha_k-1} h(x_{t-i})^{2(\alpha_k-1)} \right) \right) \right) \tag{20}$$

$$\nabla_{\phi_{ki}} L(\Theta) \nabla_{\phi_{kj}} L(\Theta) = \sum_{t=L+1}^{n} S_{Kt} \left(\left(\frac{\alpha_k - 1}{X_t \, h\left(x_{t-i}\right)^2 h\left(x_{t-j}\right)^2} + \right. \right.$$

$$\left. \left. \alpha_k \beta_k^{\alpha_k} \left(X_t^{-\alpha_k - 1} h\left(x_{t-i}\right)^{2(\alpha_k - 1)} h\left(x_{t-j}\right)^{2(\alpha_k - 1)} \right) \right) \right) \; for \; i \neq j \tag{21}$$

$$\nabla_{\alpha_k \, \alpha_k} L(\Theta) = \sum_{t=L+1}^{n} S_{kt} \left(\frac{1}{h\left(x_{t-i}\right)} - h\left(x_{t-i}\right)^{-2\alpha} \beta_k \left(e^{2\alpha_k} - e^{-2\alpha_k} \right) \right) \tag{22}$$

$$\nabla_{\alpha_k} L(\Theta) \nabla_{\alpha_j} L(\Theta) = \sum_{t=L+1}^{n} S_{kt} \left(\frac{1}{h\left(x_{t-i}\right) h\left(x_{t-j}\right)} - \right.$$

$$\left. h\left(x_{t-i}\right)^{-2\alpha_k} h\left(x_{t-j}\right)^{-2\alpha_j} \beta_k \left(e^{2\alpha_k} e^{2\alpha_j} - e^{2\alpha_k} e^{-2\alpha_j} \right) \right) \; for \; k \neq j \tag{23}$$

$$\nabla_{\beta_k} L(\Theta) \nabla_{\beta_k} L(\Theta) = \sum_{t=L+1}^{n} -\frac{S_{kt} \alpha_k}{\beta_k^2} \left(-\alpha_k \beta_k^{\alpha_k - 1} X_t^{-\alpha_k} \right) \tag{24}$$

$$\nabla_{\beta_k} L(\Theta) \nabla_{\beta_j} L(\Theta) = \sum_{t=L+1}^{n} -\frac{S_{kt} \alpha_k}{\beta_k^2 \beta_j^2} \left(-\alpha_k \beta_k^{\alpha_k - 1} \beta_j^{\alpha_j - 1} X_t^{-\alpha_k} X_t^{-\alpha_j} \right) \; for \; k \neq j \tag{25}$$

Adopting the EM algorithm for estimating the parameter space Θ in $L(\Theta)$ of (10).

(a) The first step of the Expectation-Maximization algorithm which stands for E-step, If each of the coefficient (parameter) in the parameter space Θ is known, such that the unknown values of the unobserved data $\left(S_{L,t}\right)$ is then substituted by enforcing their means on each coefficient on the observed data X_t. Subjecting $\Psi_{k,t}$ has the enforced mean on $S_{k,t}$, then the individual probability of $\Psi_{k,t}$ is the total probability of the enforced means over individual mean, as given in equation below:

$$\Psi_{k,t} = \frac{\omega_k \dfrac{\alpha_k}{\beta_k} \left(\dfrac{\beta_k}{X_t} \right)^{\alpha_k + 1} \exp\left(-\left(\dfrac{\beta_k}{X_t} \right)^{\alpha_k} \right)}{\sum_{k=1}^{K} \omega_k \dfrac{\alpha_k}{\beta_k} \left(\dfrac{\beta_k}{X_t} \right)^{\alpha_k + 1} \exp\left(-\left(\dfrac{\beta_k}{X_t} \right)^{\alpha_k} \right)} \tag{26}$$

For $k = 1, \cdots, K; \quad t = L+1, \cdots, n; \quad S_{k,t} = \Psi_{k,t}$

(b) The M-step where the unknown values of S is presumed via guessing and their enforced means are assumed to be the coefficients. The mixing weight per each regime is then estimate via

$$\hat{\omega}_k = \sum_{t=L+1}^{n} \frac{\Psi_{k,t}}{n - L} \tag{27}$$

For $\hat{\phi}_{k \, pk}$, where $k = 1, \cdots, p_k$ such that the estimates of parameters of these two steps are iterated until convergence is reached. The criterion for checking convergence was obtained by [12] as,

$$\max\left(\left| \frac{\Theta_j^{s+1} - \Theta_j^s}{\Theta_j^s} \right|; \; s, j \geq 1 \right) \leq 10^{-5} \tag{28}$$

for Θ_j^s being the jth component of Θ^s.

3.3. Limiting Distribution (Asymptotic Property) for FMA Model

The Limit distribution (Asymptotic property) for FMA model would be based on limiting laws for Extreme Value Distributions stated by Gnedenko-Fisher Tippet theorem [9].

The theorem stated that: Let $\{X_i\}_{i=1}^{n}$ be a sequence of IID random variables of sample size (n) with common Cumulative Density Function $F(x_t)$ with its maximum denoted as $M_n = (X_1, X_2, \cdots, X_n)$ such that distribution (Limiting behavior) of M_n as the sample size (n) approaches infinity is

$$P(M_n \leq x_t) = P(X_1 \leq x_t, X_2 \leq x_t, \cdots, \leq X_n \leq x_t) = \prod_{i=1}^{n} P(X_i \leq x_t) = F^n(x_t)$$

Investigating the wavering of M_n in a similar vein to that of Central Limit Theorem (CTL) for sum of the random variables $S_n = \sum_{i=1}^{n} X_i$, then the distribution of $F(x)$ belongs to the maximum domain of non-degenerate distribution $G_\xi(x_t)$ if there exist sequences of constants $\{a_n\}$ and $\{b_n\} > 0$, such there exist a real value x_t for

$$\lim_{n \to \infty} P\left(\frac{M_n - b_n}{a_n} \leq x_t\right) = \lim_{n \to \infty} F^n(a_n x_t + b_n) \to G_\xi(x_t) \qquad x_t \varepsilon \mathbb{R}.$$

Such that $G_\xi(x_t)$ is non-degenerate distribution function, and $G(x_t)$ belongs to any of the $G_\xi(x_t)$

Starting with,

$$Pr\left(\frac{M_n - b_n}{a_n} \leq x_t\right) = F^n(a_n x_t + b_n)$$

the Cumulative Distribution Function for Fréchet is

$$F(x_t) = \Pr(X_t \leq x_t) = \exp(-x_t)^{-\alpha} \qquad x_t \geq 0, \, \alpha > 0 \tag{29}$$

So,

$$Pr\left(\frac{M_n - b_n}{a_n} \leq x_t\right) = \left\{\exp\left\{-(a_n X_t + b_n)\right\}\right\}^{-\alpha n} \tag{30}$$

Where, $X_t = \phi_{k,0} + \phi_{k,1} x_{t-1} + \cdots + \phi_{k,p} x_{t-pk} = \sum_{k=1}^{K}\left(\phi_{i0} + \sum_{p=1}^{pk} \phi_{i1} x_{t-pk}\right)$

$$\approx \exp\left\{-\left(a_n\left(\sum_{k=1}^{K}\left(\phi_{i0} + \sum_{p=1}^{pk}\phi_{i1}x_{t-pk}\right)\right) + b_n\right)\right\}^{\alpha} + \exp\left\{-\left(a_n\left(\sum_{k=1}^{K}\left(\phi_{i0} + \sum_{p=1}^{pk}\phi_{i1}x_{t-pk}\right)\right) + b_n\right)\right\}^{2\alpha}$$

$$+ \exp\left\{-\left(a_n\left(\sum_{k=1}^{K}\left(\phi_{i0} + \sum_{p=1}^{pk}\phi_{i1}x_{t-pk}\right)\right) + b_n\right)\right\}^{3\alpha} + \cdots \tag{31}$$

According to [5], the normalizing constants of a_n and b_n for Fréchet distribution are approximately $\dfrac{n^{\frac{1}{\alpha}}}{\alpha}$ and $n^{\frac{1}{\alpha}}$ respectively, where "n" is the set of natural numbers.

$$= \exp\left\{n^{\frac{1}{\alpha}}\left(\left(\sum_{k=1}^{K}\left(\phi_{i0} + \sum_{p=1}^{pk}\phi_{i1}x_{t-pk}\right)\right) + \alpha\right)\right\} + \exp\left\{2n^{\frac{1}{\alpha}}\left(\left(\sum_{k=1}^{K}\left(\phi_{i0} + \sum_{p=1}^{pk}\phi_{i1}x_{t-pk}\right)\right) + \alpha\right)\right\}$$

$$+ \exp\left\{ 3n^{\frac{1}{\alpha}} \left(\left(\sum_{k=1}^{K} \left(\phi_{i0} \sum_{p=1}^{pk} \phi_{i1} x_{t-pk} \right) \right) + \alpha \right) \right\} + \cdots \tag{32}$$

Logarithm of both sides

$$\log_e \Pr\left(\frac{M_n - b_n}{a_n} \le x_t \right) = \log_e F^n \left(a_n x_t + b_n \right)$$

$$\lim_{n\to\infty} \log_e \Pr\left(\frac{M_n - b_n}{a_n} \le x_t \right) = \lim_{n\to\infty} \log_e F^n \left(a_n x_t + b_n \right)$$

$$\log_e \lim_{\infty\to n} \Pr\left(\frac{M_n - b_n}{a_n} \le x_t \right) = \log_e \lim_{\infty\to n} F^n \left(a_n x_t + b_n \right)$$

$$= \sum_{i=1}^{n} \left(\left(\sum_{k=1}^{K} \left(\phi_{i0} + \sum_{p=1}^{pk} \phi_{i1} x_{t-pk} \right) \right) + \alpha \right)$$

$$= n\alpha + \sum_{i=1}^{n} \left(\left(\sum_{k=1}^{K} \left(\phi_{i0} + \sum_{p=1}^{pk} \phi_{i1} x_{t-pk} \right) \right) \right) \tag{33}$$

Taking exponent of both sides so as to naturalize the logarithm

$$\exp\left(\log_e \lim_{\infty\to n} \Pr\left(\frac{M_n - b_n}{a_n} \le x_t \right) \right) = \exp\left(\log_e \lim_{\infty\to n} F^n \left(a_n x_t + b_n \right) \right)$$

$$\lim_{\infty\to n} \Pr\left(\frac{M_n - b_n}{a_n} \le x_t \right) = \lim_{\infty\to n} F^n \left(a_n x_t + b_n \right)$$

$$= \exp\left(n\alpha + \sum_{i=1}^{n} \left(\left(\sum_{k=1}^{K} \left(\phi_{i0} + \sum_{p=1}^{pk} \phi_{i1} x_{t-pk} \right) \right) \right) \right) \tag{34}$$

Which means that as $n \to \infty$, the limiting distribution of the FMA asymptotically approaches

$$\exp\left(n\alpha + \sum_{i=1}^{n} \left(\left(\sum_{k=1}^{K} \left(\phi_{i0} + \sum_{p=1}^{pk} \phi_{i1} x_{t-pk} \right) \right) \right) \right)$$

3.4. Ergodic Process for FMA Model

Strictly stationary and ergodic processes for FMA model would be ascertained via

$$\lim_{l\to\infty} P\left(\max\left| \frac{1}{n} \sum_{i=1}^{n} \left[l_t(\theta) - E\left(l_t(\theta)\right) \right] \right| \ge \varepsilon \right) = 0$$

Where, $l_t(\theta) \,\&\, E\left(l_t(\theta)\right)$ are likelihoods and expectations of likelihoods respectively.

$$l_t(\Theta) = \sum_{t=L+1}^{n} \left[\sum_{k=1}^{K} S_{kt} \log(\omega_k) + \sum_{k=1}^{K} S_{kt} \log\frac{\alpha_k}{\beta_k} + \sum_{k=1}^{K} S_{kt}\alpha_k \log\frac{\beta_k}{X_t} + \sum_{k=1}^{K} S_{kt} \log\frac{\beta_k}{X_t} - \sum_{k=1}^{K} S_{kt}\beta_k^{\alpha_k} X_t^{-\alpha_k} \right] \tag{35}$$

$$E\left(l_t(\Theta)\right) = E\left[\sum_{t=L+1}^{n} \left[\sum_{k=1}^{K} S_{kt} \log(\omega_k) + \sum_{k=1}^{K} S_{kt} \log\frac{\alpha_k}{\beta_k} + \sum_{k=1}^{K} S_{kt}\alpha_k \log\frac{\beta_k}{X_t} + \sum_{k=1}^{K} S_{kt} \log\frac{\beta_k}{X_t} - \sum_{k=1}^{K} S_{kt}\beta_k^{\alpha_k} X_t^{-\alpha_k} \right] \right] \tag{36}$$

But,

$$\left[l_t(\theta) - E\left(l_t(\theta)\right) \right] = \left[\sum_{t=L+1}^{n} \left[\sum_{k=1}^{K} S_{kt}\alpha_k \log X_t - \sum_{k=1}^{K} S_{kt} \log X_t - \sum_{k=1}^{K} S_{kt}\beta_k^{\alpha_k} X_t^{-\alpha_k} \right] - \right.$$
$$\left. \left[\sum_{t=L+1}^{n} \left[\sum_{k=1}^{K} S_{kt}\alpha_k E(\log X_t) - \sum_{k=1}^{K} S_{kt} E(\log X_t) - \sum_{k=1}^{K} S_{kt}\beta_k^{\alpha_k} E\left(X_t^{-\alpha_k}\right) \right] \right] \right] \quad (37)$$

From Jensen's inequality, $E(\ln(X)) \le \ln E(X)$ for every positively integrable X.

$$\text{So, } \left[l_t(\theta) - E\left(l_t(\theta)\right) \right] \le \left[\sum_{t=L+1}^{n} \left[\sum_{k=1}^{K} S_{kt}\alpha_k \log X_t - \sum_{k=1}^{K} S_{kt} \log X_t - \sum_{k=1}^{K} S_{kt}\beta_k^{\alpha_k} X_t^{-\alpha_k} \right] - \right.$$
$$\left. \left[\sum_{t=L+1}^{n} \left[\sum_{k=1}^{K} S_{kt}\alpha_k \log E(X_t) - \sum_{k=1}^{K} S_{kt} \log E(X_t) - \sum_{k=1}^{K} S_{kt}\beta_k^{\alpha_k} E\left(X_t^{-\alpha_k}\right) \right] \right] \right] \quad (38)$$

Recall, $E(X_t) = \beta_k \Gamma\left(1 - \dfrac{1}{\alpha_k}\right)$

$$\approx \sum_{t=L+1}^{n} S_{kt}\alpha_k \left(\log \frac{X_t}{\beta_k \Gamma\left(1 - \frac{1}{\alpha_k}\right)} \right) - \sum_{t=L+1}^{n} S_{kt}\left(\log X_t \beta_k \Gamma\left(1 - \frac{1}{\alpha_k}\right) \right) - \sum_{t=L+1}^{n} S_{kt}\beta_k^{\alpha_k}\left(X_t^{-\alpha_k} + \beta_k \Gamma\left(1 - \frac{1}{\alpha_k}\right)^{-\alpha_k} \right) \quad (39)$$

$$\left| \frac{1}{n}\sum_{i=1}^{n}\left[l_t(\theta) - E\left(l_t(\theta)\right) \right] \right| = \approx \sum_{t=L+1}^{n} S_{kt}\alpha_k \left(\log \frac{X_t}{\beta_k \Gamma\left(1 - \frac{1}{\alpha_k}\right)} \right) - \sum_{t=L+1}^{n} S_{kt}\left(\left(\log X_t \beta_k \Gamma\left(1 - \frac{1}{\alpha_k}\right) \right) + \right.$$
$$\left. \beta_k^{\alpha_k}\left(X_t^{-\alpha_k} + \beta_k \Gamma\left(1 - \frac{1}{\alpha_k}\right)^{-\alpha_k} \right) \right) \quad (40)$$

$$\max \left| \frac{1}{n}\sum_{i=1}^{n}\left[l_t(\theta) - E\left(l_t(\theta)\right) \right] \right| = \sum_{t=L+1}^{n} S_{kt}\alpha_k \left(\log \frac{X_t}{\beta_k} \right) - \sum_{t=L+1}^{n} S_{kt}\left(\left(\log X_t \beta_k \right) + \beta_k^{\alpha_k}\left(X_t^{-\alpha_k} + \beta_k \right) \right) \quad (41)$$

since, $\Gamma(1) = 1! = 1$

$$\max \left| \frac{1}{n}\sum_{i=1}^{n}\left[l_t(\theta) - E\left(l_t(\theta)\right) \right] \right| \ge \varepsilon = \sum_{t=L+1}^{n} S_{kt}\alpha_k \left(\log \frac{X_t}{\beta_k} \right) - \sum_{t=L+1}^{n} S_{kt}\left(\left(\log X_t \beta_k \right) + \beta_k^{\alpha_k}\left(X_t^{-\alpha_k} + \beta_k \right) \right) \ge \varepsilon$$
$$\Rightarrow \sum_{t=L+1}^{n} S_{kt}\left(\left(\alpha_k \log \frac{X_t}{\beta_k} \right) - \left(\left(\log X_t \beta_k \right) + \beta_k^{\alpha_k}\left(X_t^{-\alpha_k} + \beta_k \right) \right) \right) \ge \varepsilon \quad (42)$$
$$\Rightarrow \lim_{l \to \infty} P\left(\sum_{t=L+1}^{n} S_{kt}\left(\left(\alpha_k \log \frac{X_t}{\beta_k} \right) - \left(\left(\log X_t \beta_k \right) + \beta_k^{\alpha_k}\left(X_t^{-\alpha_k} + \beta_k \right) \right) \right) \ge \varepsilon \right) < \infty \quad (43)$$

4. Conclusions

The proposed model of $FMAR(k; p_1, p_2, \ldots, p_k)$ being one of the prominent models for extreme valued mixture autoregressive was devised in order to sieze and objurgate stylized traits; entire greater range of multimodalities, regime-switching, (ability to handle cycles), time-varying volatilities (conditional variances), time-varying mixing weights with Fréchet Probability Density Function (PDF). Long-memory, heavy-tailed distributed observations are the causal that arise to outliers, change points like behavior, abnormal kurtosis, and skewness. The formulated and proposed $FMAR(k; p_1, p_2, \ldots, p_k)$ will be an ideal and substitute mixture autoregressive model whenever heavy-tailed distributed observations are the causal that arise to outliers,

abnormal kurtosis, and skewness brings about the regime-switching.

REFERENCES

[1]   A. Ang, and A. Timmermann, "Regime Changes and Financial Markets" Annual Review of Financial Economics, Vol. 4(1), pp. 313-337, 2011.

[2]   G. Behrooz., S.M.N. Abu Hassan, and S. Tamat, "Application of the Threshold Model for Modelling and Forecasting of Exchange Rate in Selected ASEAN Countries". Sains Malaysiana, Vol.43(10), pp. 1609-1622, 2014.

[3]   G.N. Boshankov, "Prediction with Mixture Autoregressive Models" Research Report No.6, 2006, Probability and Statistics Group School of Mathematics. The University of Manchester, 2006.

[4]   T.A. Chawsheen, and M. Broom, "Seasonal time-series modeling and forecasting of monthly mean temperature for decision making in the Kurdistan Region of Iraq". Journal of Statistical Theory and Practice. Vol. 11(4), pp.604– 633, 2017.

[5]   J. Y. Fortin, and M. Clusel, "Applications of extreme value statistics in physics". Journal of Physics, A Mathematical and Theoretical, pp. 1-39, 2015.

[6]   L. Kalliovirta, M. Meitz, and P. Saikkonen., "Identification and estimation of non-Gaussian Structural Vector Autoregressions" Journal of Econometric, Vol. 196, pp. 288-304, 2017.

[7]   L. Kalliovirta, M. Meitz, and P. Saikkonen, A Guassian Mixture Vector Autoregressive. Journal of econometrics, Vol.192, pp. 485-498, 2016.

[8]   L. Kalliovirta, M. Meitz, P. and Saikkonen, "A Guassian Mixture autoregresive model for univariate series" Journal of Time Series Analysis, Vol. 36, pp. 247-266, 2012.

[9]   O. Karpa, "Development of Bivariate Extreme Value Distributions for Applications in Marine Technology". (Doctoral thesis, Norwegian University of Science and Technology (NTNU), Norway), 2015.

[10] M. Li, and M. Raftery, "Modeling flat stretches, burst and outliers in time series using Mixture Transition Distribution Models". Journal of American Statistics Association, Vol. 91, pp.1504-1514, 1966.

[11] G.J. McLachlan, S.X. Lee, and I. R. Suren. Finite Mixture Models. Annual Review of Statistics and Its Application Finite Mixture Models, Vol. 6, pp. 355–378, 2019.

[12] V. Melnykova, and I. Melnykovb "Initializing the EM algorithm in gaussian mixture models with an unknown number of components" Computational Statistics and Data Analysis, Vol. 56, pp. 1381–1395, 2012.

[13] N. Nguyen, "Hidden Markov Model for Stock Trading". International Journal Financial studies, Vol. 6 (36), 2018.

[14] C.S. Wong, W.S, Chan, and P.L. Kam, "A student t-Mixture Autoregressive Model with Application to heavy-tailed financial data" Singapore Economic Review Conference, pp. 1-10, 2009.

[15] C. S. Wong, "Statistical inference for some nonlinear time series models" Ph.D thesis, University of Hong Kong, Hong Kong, 1998.