VALUE-AT-RISK PREDICTION USING GARCH MODEL AND BAYESIAN EXTREME VALUE FOR MIXTURE DISTRIBUTIONS

[(1)]*Frihi, Redhouane.,* [(2)]*Rassoul, Abdelaziz.,* [(3)] *Hamid, Ouldrouis*

[(1)]Blida 1 University, BP 270 road soumaa, Blida, Algeria.
[(2)] National Higher School of Hydraulic, Guerouaou, BP 31, Blida, 09000, Algeria
[(3)]Blida 1 University, BP 270 road soumaa, Blida, Algeria.
[(1)] rfrihi@yahoo.fr, [(2)]a.rassoul@ensh.dz,[(3)]horouis@hotmail.comr

ABSTRACT

In the current paper, we proposed a method to estimate value-at-risk (VaR) for the model with GARCH effect when the distribution of independent innovations of the residuals of the GARCH model has a mixture of generalized hyperbolic secant distribution (*GHSD*), with two tailed generalized Pareto distributions (*GPD*), and estimating the parameters of the *GHSD* and the *GPD* distributions by the Bayesian inference approch.
This approach has two steps : The first step is the estimation of the variability (volatility) followed by the GARCH model using the maximum likelihood method. The second step is used to take into account the uncertainties in the parameters, including the a priori information to estimate the parameters of the hyperbolic secant distribution (*GHSD*) and (*GPD*) distributions to obtain the a posteriori distribution of the parameters. We apply the mixture model to the innovations obtained from the residuals to derive the value-at-risk (VaR) estimates.

Key words : Generalized Pareto distribution (GPD), Generalized Hyperbolic secant distribution (GHSD) , GARCH model, Bayesian inference, Value-at-Risk.

1. INTRODUCTION

The purpose of this paper is to quantify risk. Value-at-Risk (VaR) is one such measure of risk, which quantifies the largest possible profit of a portfolio over a fixed time period for a given small probability. To estimate VaR using one of the statistical models, many assumptions are necessary. One of these is that daily returns are identical, independent and normally distributed. However, in the real world, financial data are not normally distributed and contain skewness or kurtosis properties. Therefore, modeling VaR with the assumption of normality, without accounting for large and unexpected losses that occur in the tail of the distribution, results in underestimated or overestimated VaR forecasts. Due to the inability of the normal distribution to model the tail of financial return series, many researchers have used skewed and heavy-tailed distributions to forecast VaR to overcome this problem. Many models have been proposed to capture the clustering effect of volatility, the most widely used being the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model.

Extreme value models have been widely used in financial applications such as risk analysis, forecasting and pricing models. Extreme value models describe the stochastic dynamics of a process for states that have a small probability of realizing and generally outside the range of observed data (Beirlant et al. 2004).

This paper develops a new model based on extreme value theory (EVT) to estimate VaR, using the extreme quantiles of the return series after taking into account the dependence structure induced by the volatility clustering modeled by the GARCH model. A two-step approach commonly used in the financial literature (due to McNeil and Frey 2000). In particular, a mixture of three distributions noted (GHG) are used to capture the full distribution of innovations. The generalized secant distribution (GHSD) (unimodal and symmetric) is used to describe the main mode of the innovation distribution and generalized Pareto (GPD) is used to simultaneously extrapolate the gains (on the right) and losses (on the left) beyond certain thresholds, which represent the upper and lower tails of the innovation distribution. We will use this two-step methodology as the basis for the approach taken in this paper.

Bayesian inference is used for fitting the mixture model as it can take advantage of any expert prior information, which can be important in tail estimation due to the inherent sparsity of extreme data. The estimation method for the proposed mixture model is firstly evaluated, followed by application of the two stage GARCH-GHSD-GPD mixture model to forecasting VaR.

The rest of the paper is organized as follows.
Section 2, restricted to the one sided tailed GPD distribution. In the section 3, we presente the mixture GPD-GHSD-GPD model for the independent residuals. In Section 4, we give the VaR extreme in presence of the GARCH effect. Section 5 is devoted to estimate the parameters of the GPD-GHSD-GPD model using the bayesian inference.

2. ONE SIDED GPD DISTRIBUTION

The generalized Pareto distribution (GPD) is a model with asymptotic justification when applied to excesses which occur over a sufficiently high threshold. The GPD can equivalently be defined for excesses below a suitably low threshold for capturing the lower tail of a distribution. Let X be an i.i.d random variable with $X \geq u$ following a $G(x \mid \xi, \beta, u)$ with scale parameter β (dependent on threshold u) β and shape parameter ξ, which has a distribution function given by :

$$G(x \mid \xi,\beta,u) = P(X < x \mid X > u) = \begin{cases} 1-\left[1+\xi\left(\frac{x-u}{\beta}\right)\right]_+^{-1/\xi} & \xi \neq 0 \\ 1-\exp\left[-\left(\frac{x-u}{\beta}\right)\right] & \xi = 0 \end{cases}$$

where $x \geq u, \beta > 0$ and $y_+ = \max(y,0)$. There are three types of tail behaviour detertmined by the shape parameter : $\xi = 0$ gives an exponential tail, $\xi < 0$ gives a short tail with an upper bound given by $u - \beta/\xi$ and a heavier tail than an exponential is indicated if $\xi > 0$. The task is to find the lowest threshold such that the GPD fits the sample of exceedances over this threshold adequately.
The problem of choosing the threshold u is still of high theoretical and practical interest. It is desirable to have an intuitive automated threshold selection procedure to use with POT analysis. The simple method is an a priori, or fixed threshold selection based on expertise on the subject matter at hand. Various rules have been suggested, for example, selecting the top 10% of the data, see e.g., DuMouchel (1983), or the top 5%, see, Kelly and Jiang (2014), or the top square root of the sample size see, e.g., Bader et al. (2018), and Silva Lomba et al. (2020).
Drees et al. (2000) suggested the Hill plot, which plots the Hill estimator of the shape parameter based on the top k order statistics against the threshold u.
Many variants of the Hill plot have been proposed (Scarrott and MacDonald (2012).

3. THE GENERAL HYPERBOLIC SECANT DISTRIBUTION

Palmitesta and Provasi (2004) showed that the generalized secant hyperbolic distribution (GSHD) can be an alternative to the Student-t distribution in the interpretation of financial data with heavy tails. During the last few years, several generalizations of the hyperbolic secant distribution have become popular in the context of financial return data because of its excellent fit.

The standard hyperbolic secant distribution (HSD) has its origin in Fisher (2010). Additional properties are developed by Talacko (???? Année). It is symmetric and bell-shaped like the Gaussian distribution but has slightly heavier tails. However, in contrast, both probability density function, cumulative density function and quantile function, admit simple and closed-form expressions, which makes it appealing from a practical and a theoretical point of view. In particular, HSD can be used as starting distribution to obtain generalized distribution systems which exhibit skewness and heavier (or lighter) tails.
. The distribution function H of the standard hyperbolic secant distribution is given by

$$H(z) = \frac{2}{\pi} \arctan \left[\exp \left(\frac{\pi}{2} z \right) \right], \quad z \in \mathbb{R}$$

The quantile function H^{-1} of the standard hyperbolic secant distribution is given by

$$H^{-1}(p) = \frac{2}{\pi} \ln \left[\tan \left(\frac{\pi}{2} p \right) \right], \quad p \in (0,1)$$

The standard hyperbolic secant distribution is generalized by adding location and scale parameters. Suppose that Z has the standard hyperbolic secant distribution and that $\mu \in \mathbb{R}$ and $\sigma \in (0,\infty)$. Then $X = \mu + \sigma Z$ has the hyperbolic secant distribution with location parameter μ and scale parameter σ parameter.

The distribution function H of X is given by

$$H(x) = \frac{2}{\pi} \arctan \left\{ \exp \left[\frac{\pi}{2} \left(\frac{x-\mu}{\sigma} \right) \right] \right\}, \quad x \in \mathbb{R}$$

The density function is given by :

$$h(x) = \frac{1}{2\sigma} \frac{1}{\cosh \pi (\frac{x-\mu}{2\sigma})}$$

And the quantile function H^{-1} of X is given by

$$H^{-1}(p) = \mu + \sigma \frac{2}{\pi} \ln \left[\tan \left(\frac{\pi}{2} p \right) \right], \quad p \in (0,1)$$

Suppose that $\mu \in \mathbb{R}$ and $\sigma \in (0,\infty)$. If X has the hyperbolic secant distribution with location parameter μ and scale parameter σ then

$$U = H(X) = \frac{2}{\pi} \arctan \left\{ \exp \left[\frac{\pi}{2} \left(\frac{X-\mu}{\sigma} \right) \right] \right\}$$

has the standard uniform distribution. If U has the standard uniform distribution then

$$X = H^{-1}(U) = \mu + \sigma \frac{2}{\pi} \ln \left[\tan \left(\frac{\pi}{2} U \right) \right]$$

has the hyperbolic secant distribution with location parameter μ and scale parameter σ.
The skewness and kurtosis of X are $\mathrm{skew}(X) = 0$ and $\mathrm{kurt}(X) = 5$.

4. TWO TAIL GPD-GHSD-GPD MIXTURE MODEL

The two tail GPD mixture model has separate GPD's for the upper and lower tails beyond each threshold, with a suitable distribution between the two thresholds. The thresholds are explicitly specified by model parameters to be estimated. We will denote the two tail distribution GPD innovation with the mixture distributiion of the GHSD and GPD. The distribution function of the mixture model, $P(X \leq x) = F(x)$ where :

$$\begin{aligned} F(x \mid \theta) = &\{H(u_l \mid \theta_1)[1 - G(-x \mid \xi_l, \beta_l, -u_l)]\} I_{(-\infty, u_l}(x) \\ &+ H(x \mid \theta_1) I_{(u, u_r)}(x) + \\ &\{H(u_r \mid \theta_1) + [1 - H(u_r \mid \mu, \sigma)] G(x \mid \xi_r, \beta_r, u_r)\} I_{[u_r, \infty)} \end{aligned}$$

and $H(x)$ is the GHSD cumulative distribution function with location parameter μ and the scale parameter σ and $G(x \mid \xi, \beta, u)$ is the distribution function of GPD defined by equation 1. The subscript on the GPD parameters l denotes the lower (left) tail and r denotes the upper (right) tail. The parameter vector of the model is $\theta = (\mu, \sigma, u_r, \xi_r, \beta_r, u_l, \xi_l, \beta_l)$, $\theta_1 = (\mu, \sigma)$, $\theta_3 = (\xi_r, \beta_r, \xi_l, \beta_l)$ and $\theta_2 = (u_r, u_l)$

For a sample of size $n, x = (x_1, \ldots, x_n)$ from F, parameter vector $\theta = (\mu, \sigma, u_r, \xi_r, \beta_r, u_l, \xi_l, \beta_l)$, $A = \{i : x_i < u_l\}$, $B = \{i : u_l \leq x_i \leq u_r\}$, and $C = \{i : x_i > u_r\}$, the likelihood function is

$$\begin{aligned} L(\theta|\mathrm{x}) = \prod_A h(x \mid \theta_1) \prod_B (1 - H(u_l \mid \theta_1)) &\left(\frac{1}{\beta_l} \left[1 + \frac{\xi_l (x_i - u)}{\beta_l} \right]_+^{-(1+\xi_l)/\xi_l} \right) \\ \prod_C (1 - H(u_| \theta_1)) &\left(\frac{1}{\beta_r} \left[1 + \frac{\xi_r (x_i - u)}{\beta_r} \right]_+^{-(1+\xi_r)/\xi_r} \right) \end{aligned}$$

for $\xi \neq 0$,
and

$$\begin{aligned} \mathrm{L}(\theta \mid \mathrm{x}) = &\textstyle\prod_A h(x \mid \theta_1) \prod_B (1 - H(u_l \mid \theta_1)) ((1/\beta_l) \exp\{(x_i - u_l)/\beta_l\}) \\ &\prod_C (1 - H(u_r \mid \theta_1)) ((1/\beta_r) \exp\{(x_i - u_r)/\beta_r\}) \end{aligned}$$

, for $\xi = 0$.
Wehere $\theta_1 = (\mu, \sigma)$.

The GHG model is also able to extrapolate two sided tail distributions simultaneously, which is highly relevant in many finance/economics applications. The proposed mixture model has the flexibility in dealing with a variety of distributions, with or without the symmetry, by allowing both tails to follow separate GPD distributions.

5. GARCH MODEL

Let $\{R_t\} = \ln X_t - \ln X_{t-1}$ be a strictly stationary daily log return series on a financial asset at time t.

GARCH$(1,1)$, which is the most commonly used process of all GARCH models. It is specified as follows :

$$\begin{aligned} R_t &= \mu_t + \varepsilon_t \\ \sigma_t^2 &= \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{aligned}$$

where $\omega > 0, \alpha_1 > 0, \beta_1 > 0$, and $\alpha_1 + \beta_1 < 1$ to ensure stationarity, and $\varepsilon_t = \sigma_t z_t$ is a random variable denoting the mean corrected return/random shock. z_t is a sequence of i.i.d. r.v. with parameter θ and μ_t is the expected return at time t and σ_t is the volatility estimator from a GARCH model. The distribution of $\varepsilon_t \mid I_{t-1}$ is conditional on all information available up to time $t-1$. The dynamic behavior of the conditional variance is accounted by ε_t. This implies that σ_t^2, the conditional variance of today, is dependent on past squared disturbances, ε_t^2 Therefore, the distribution function of the observation z_t can be written as :

$$z_t \sim F(x, |\theta)$$

$$\begin{aligned} F(x|\theta) = & \{H(u_l \mid \mu,\sigma)[1-G(-x \mid \xi_l,\beta_l,-u_l)]\} I_{(-\infty,u_l}(x) \\ & + H(x \mid \mu,\sigma) I_{(u,u_r)}(x) + \\ & \{H(u_r \mid \mu,\sigma) + [1-H(u_r \mid \mu,\sigma)] G(x \mid \xi_r,\beta_r,u_r)\} I_{[u_r,\infty)} \end{aligned}$$

6. TWO STAGE APPROACH

Suppose the log-returns, $R_t = \ln X_t - \ln X_{t-1}$, follow the model

$$\varepsilon_t = \sigma_t z_t$$

Let VaR_t^α denote the corresponding value at risk. Suppose z_t are independent and identical with the mixture distribution GHG. The relationship between the VaR and the standard deviation at time t can be expressed as follows :

$$VaR_t^\alpha = \mu_t + q_F \sigma_t$$

where $q_F = F^{-1}(\alpha)$ is the α-quantile of the mixture distribution GHG.

The two stage approach to estimate the VaR is as follows :

1. Fit a GARCH volatility model to $\{R_t\}$ and obtain the standardized innovation term z_t as $R_t = \mu_t + \sigma_t z_t$.

Here, the μ_t is the expected return at time t and σ_t is the volatility estimator from a GARCH model. The form of GARCH can be selected according to the particular application.

2. Fit the proposed GHG mixture model to $\{z_t\}$ (the standardized innovation sequence) as described above. The upper tail of the mixture model represent gains and the lower tail represents the losses.

The first-stage GARCH model is fitted using a standard maximum likelihood method, as this stage is less critical for VaR estimation. However, the GHG mixture model is estimated using Bayesian inference because the complexity of the likelihood for this model means that it would be difficult to maximize it directly, and Bayesian inference also allows for the use of prior information that can greatly facilitate the estimation of tail quantities (such as VaR).

7. BAYESIAN INFERENCE FOR MIXTURE MODEL

Bayesian inference is used to estimate the mixture model parameters to combine the a priori information from the experts potentially with the sample data. Markov chain Monte Carlo (MCMC) method was used to obtain the a posteriori distribution. In this section, we principally apply Bayesian inference. In the absence of expert knowledge, it is convenient to make use of so called objective priors, such as the Jeffreys prior (Jeffreys, 1961) and the maximal data information (MDI) prior (first mentioned in Zellner, 1971, and given explicitly in Zellner, 1996).

7.1. Prior Distribution

The parameter vector $\theta = (\mu,\sigma,u_r,\xi_r,\beta_r,u_l,\xi_l,\beta_l)$ can be decomposed into three components $\theta_1 = (\mu,\sigma)$, $\theta_2 = (\xi_l,\beta_l)$ $\theta_3 = (\xi_r,\beta_r)$ and $\theta_4 = (u_r,u_l)$, associated with the GHSD, GPD parameters, and the thresholds respectively. In this study we explicitly specify priors with little information. In specific applications, however, expert information could be included to give more informative priors which could reduce the uncertainty associated with parameter estimates.

7.2. Prior for the *GHSD* parameters

If the chosen distribution $H(x \mid \theta_1)$ is gamma (due to Behrens et al (2004)). Moreover, it is logical to assume the prior independence between the shape parameter and the location parameter. We then define $\mu \sim \mathrm{Ga}(a,b)$ and $\sigma \sim \mathrm{Ga}(c,d)$, where a,b,c and d are known hyperparameters. The joint precedence of $\theta_1 = (\mu,\sigma)$ is then the following

$$\pi(\theta_1) = \frac{b^a}{\Gamma(a)}\mu^{a-1}\mathrm{e}^{-b\mu}\,\frac{d^c}{\Gamma(c)}\left(\frac{\mu}{\sigma}\right)^{c-1}\mathrm{e}^{-d\mu/\sigma}\left(\frac{\mu}{\sigma^2}\right)$$

7.3. Prior for the *GPD* parameters

The idea we use here is from Coles and Tawn (1996) refers to the elicitation of information within a parameterization on which experts are familiar. More precisely, by the inversion of Equation (1.1), we obtain the $1-p$ quantile of the distribution,

$$q = u + \frac{\beta}{\xi}\left(p^{-\xi} - 1\right)$$

The formulation of the prior elicited on the quantile differences also permits consideration of the known negative dependence between the shape ξ and scale β parameters of the *GPD*. A gamma prior distribution is used to describe the quantile differences. We assume the quantile differences follow a gamma distribution, so that : $d_1 = q_1 \sim \mathrm{Ga}(a_1,b_1)$ and $d_2 = q_2 - q_1 \sim \mathrm{Ga}(a_2,b_2)$ The marginal prior distribution for β and ξ for the excesses above (or below if lower tail) the threshold, the prior for upper tail is defined as

$$\begin{aligned}\pi(\beta,\xi) \propto \exp&\left\{-b_1\left[u+\frac{\beta}{\xi}\left(p_1^{-\xi}-1\right)\right]\right\}\left[u+\frac{\beta}{\xi}\left(p_1^{-\xi}-1\right)\right]^{a_1-1}\\ &\times\exp\left\{-b_2\left[\frac{\beta}{\xi}\left(p_2^{-\xi}-p_1^{-\xi}\right)\right]\right\}\left[\frac{\beta}{\xi}\left(p_2^{-\xi}-p_1^{-\xi}\right)\right]^{a_2-1}\\ &\times\left|\frac{\beta}{\xi^2}\left[(p_1p_2)^{-\xi}(\log p_1-\log p_2)+p_2^{-\xi}\log p_2-p_1^{-\xi}\log p_1\right]\right|\end{aligned}$$

The prior for the lower tail GPD parameters is similarly defined. In this paper we have used the quantile differences for the conditional tail probabilities ($p_1 = 0.1$ and $p_2 = 0.01$) following Coles and Tawn (1996). The tail probabilities considered for p_1 and p_2 can be altered according to the application and the available expert information.

7.4. Prior for the thresholds

There are many ways to set up a prior distribution for u. We can assume that u follows a truncated normal distribution with parameters $\left(\mu_u,\sigma_u^2\right)$, which are truncated at the minimum and maximum of the sample data respectively (and thresholds), due to Behrens et al. (2004) :

$$\pi\left(u \mid \mu_u,\sigma_u^2\right) \propto \exp\left[-\frac{1}{2}\left(\frac{u-\mu_u}{\sigma_u^2}\right)^2\right]$$

for the lower threshold $u = u_l$ and upper threshold $u = u_r$.

A continuous uniform prior is another alternative. A discrete distribution can also be assumed.

for the lower threshold $u = u_l$ and upper threshold $u = u_r$. The priors for the GHSD and GPD components are assumed independent giving the logarithm of the posterior distribution

$L(\theta \mid x) \propto \pi(\theta)L(x \mid \theta)$ for $\xi \neq 0$:

$$\begin{aligned}
&L(\theta \mid x) = \pi(\theta_1)\pi(\theta_2)\pi(\theta_3)\pi(\theta_4)L(x \mid \theta)\\
&= \frac{b^a}{\Gamma(a)}\mu^{a-1}\mathrm{e}^{-b\mu}\frac{d^c}{\Gamma(c)}\left(\frac{\mu}{\sigma}\right)^{c-1}\mathrm{e}^{-d\mu/\sigma}\left(\frac{\mu}{\sigma^2}\right)\\
&\times\left\{-b_1^l\left[u+\frac{\beta_l}{\xi}\left(p_1^{-\xi_l}-1\right)\right]\right\}\left[u+\frac{\beta_l}{\xi}\left(p_1^{-\xi_l}-1\right)\right]^{a_1-1}\\
&\times\exp\left\{-b_2^l\left[\frac{\beta_l}{\xi_l}\left(p_2^{-\xi_l}-p_1^{-\xi_l}\right)\right]\right\}\left[\frac{\beta_l}{\xi_l}\left(p_2^{-\xi_l}-p_1^{-\xi_l}\right)\right]^{a_2^l-1}\\
&\times\left|\frac{\beta}{\xi_l^2}\left[\left(p_1^l p_2^l\right)^{-\xi}\left(\log p_1^l-\log p_2^l\right)+p_2^{-\xi_l}\log p_2^l-p_1^{-\xi_l}\log p_1^l\right]\right|\\
&\times\left\{-b_1^r\left[u+\frac{\beta_r}{\xi_r}\left(p_1^{-\xi_r}-1\right)\right]\right\}\left[u_r+\frac{\beta_r}{\xi_r}\left(p_1^{-\xi_r}-1\right)\right]^{a_1^r-1}\\
&\times\exp\left\{-b_2^r\left[\frac{\beta_r}{\xi_r}\left(p_2^{-\xi_r}-p_1^{-\xi_r}\right)\right]\right\}\left[\frac{\beta_r}{\xi_r}\left(p_2^{-\xi_r}-p_1^{-\xi_r}\right)\right]^{a_2^r-1}\\
&\times\left|\frac{\beta_r}{\xi_r^2}\left[(p_1^r p_2^r)^{-\xi_r}(\log p_1^r-\log p_2^r)+p_2^{-\xi_r}\log p_2^r-p_1^{-\xi_r}\log p_1^r\right]\right|\\
&\exp\left[-\frac{1}{2}\left(\frac{u-\mu_u}{\sigma_u^2}\right)^2\right]\times L(\theta \mid x)
\end{aligned}$$

In the case $\xi = 0$, the a posteriori distribution is deduced by replacing ξ by 0 on the left and/or on the right.

The computation is done through the MCMC methods, via Metropolis-Hastings (M-H) algorithm. In the algorithm implementation, each subset of parameters is updated at each iteration step in terms of the importance order of the parameters as $(\xi_r, \beta_r, u_r), (\xi_l, \beta_l, u_l)$ and finally (μ, σ).

8. CONCLUSIONS

In this paper, we have tried to present the Bayesian approach to estimate the value-at-risk in the case of a dynamic model with GARCH effect by assuming that the error of the residuals is of mixed distribution whose central law is followed by the GHSD distribution and the two tails are followed by the GPD distribution.

9. REFERENCES

[1] Bader, B., J. Yan, and X. Zhang (2018). Automated threshold selection for extreme value analysis via ordered goodness-of-fit tests with adjustment for false discovery rate. Ann. Appl. Stat. 12, 310–329.

[2] Behrens, C., Lopes, H., Gamerman, H., 2004. Bayesian analysis of extreme events with threshold estimation. Statistical Modelling 4, 227–244.

[3] Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. Statistics of Extremes : Theory and Applications. Wiley Series in Probability and Statistics.

[4] Chan, N., Deng, S., Peng, L., Xia, Z., 2007. Interval estimation of Value-at-Risk based on GARCH models with heavy-tailed innovations. Journal of Econometrics 137, 556–576.

[5] COLES,S., TAWN,J. A Bayesian Analysis of Extreme Rainfall Data. Appl. Statist. (1996) 45, No. 4, pp. 463-478.

[6] DuMouchel, W. H. (1983), "Estimating the Stable Index in Order to Measure Tail Thickness : A Critique," The Annals of Statistics, 11, 1019–1031.

[7] Drees, H., de Haan, L., and Resnick, S. (2000), How to Make a Hill Plot,"The Annals of Statistics, 28, 254-274.

[8] Fischer,.M, Generalized hyperbolic secant distributions. With applications to finance, Book Series : SpringerBriefs in Statistics, 2014/01/01.

[9] Fischer, M. : Generalized Tukey-type distribution with applications to financial return data.Stat. Pap. 1, 41–56 (2010).

[10] McNeil, A. J., Frey, R., 2000. Estimation of tail-related risk measures for heteroscedastic financial time series an extreme value approach. Journal of Empirical Finance 7, 271–300.

[11] Kelly, B., H. Jiang, 2014. Tail risk and asset prices. Review of Financial Studies 27(10), 2841–2871.

[12] Silva Lomba, J., Fraga Alves, M.I. L-moments for automatic threshold selection in extreme value analysis. Stoch Environ Res Risk Assess 34, 465–491 (2020).

[13] Talacko, J. : A note about a family of Perks distribution, Sankhya. Indian J. Stat. 20(3,4), 323–328 (1958).

[14] Zellner, A., An Introduction to Bayesian Inference in Econometrics. Willey classics library Edition Pulished 1996.