

DISCRIMINATING BETWEEN THE NORMAL INVERSE GAUSSIAN AND GENERALIZED HYPERBOLIC SKEW-T DISTRIBUTIONS WITH A FOLLOW-UP THE STOCK EXCHANGE DATA

Hanieh PANAHI
*Department of Mathematics and Statistics
Lahijan branch, Islamic Azad University, Lahijan, Iran,
panahi@liau.ac.ir*

Received: August 2017 / Accepted: March 2018

Abstract: The statistical methods for the financial returns play a key role in measuring the goodness-of-fit of a given distribution to real data. As is well known, the normal inverse Gaussian (NIG) and generalized hyperbolic skew-t (GHST) distributions have been found to successfully describe the data of the returns from financial market. In this paper, we mainly consider the discrimination between these distributions. It is observed that the maximum likelihood estimators (MLEs) cannot be obtained in closed form. We propose to use the EM algorithm to compute the maximum likelihood estimators. The approximate confidence intervals of the unknown parameters have been constructed. We then perform a number of goodness-of-fit tests to compare the NIG and GHST distributions for the stock exchange data. Moreover, the Vuong type test, based on the Kullback-Leibler information criteria, has been considered to select the most appropriate candidate model. An important implication of the present study is that the GHST distribution function, in contrast to NIG distribution, may describe more appropriate for the proposed data.

Keywords: Generalized hyperbolic skew-t distribution, EM algorithm, Goodness-of-fit, Normal inverse Gaussian distribution, Stock exchange, Vuong type test.

MSC: 91G70, 62F10, 62F03.

1. INTRODUCTION

For portfolio risk modeling and basket derivative pricing, it is essential to determine the correct statistical distribution. However, it is widely acknowledged

that returns and other financial variables are not normally distributed. Therefore, selecting the statistical model that represents the leptokurtic behavior of the financial returns is an important research topic. Thus, we must consider the distributions that are heavier or fatter than the normal distribution. The class of generalized hyperbolic (GH) distributions that has heavy-tail property is one of the important distributions in financial data analysis. The NIG and GHST distributions, special cases of the GH distribution, have been used with success in many areas of their application such as turbulence (Barndorff-Nielsen, 1997), biology (Bilsild, 1981), and finance (Goncu and Yang (2016), Lee and McLachlan (2016), Mabitsela et. al., (2015), Nakajima (2014), Necula (2009), Azzalini and Capitanio (2003), Lillestøl, (1998; 2000)). Although, the two distribution functions have adequately fit for the financial returns, still it is desirable to select the correct or more nearly correct model for this data. In this paper, we want to discriminate between the NIG and GHST distributions for the stock exchange data. For this purpose, we first estimate the unknown parameters of the NIG and GHST distributions. The maximum likelihood procedure is one of the important methods for estimating the unknown parameters. In many situations where the maximum likelihood estimators of the parameters cannot be expressed in a closed form, we can obtain the MLEs using different numerical algorithms. The standard Newton-Raphson algorithm is one of the algorithms to estimate the parameters, but it does not converge always. So, we can use some very powerful algorithm, say the EM algorithm (Dempster, 1977) to compute the MLEs. Due to its efficiency, the EM algorithm has been studied by several authors such as Ng et. al., (2002), Rastogi and Tripathi (2012), Panahi (2016; 2017), and Panahi and Sayyareh (2016). We then apply a number of goodness-of-fit tests to discriminate between the NIG and GHST distributions. The goodness-of-fit (GOF) tests are used for verifying whether the experimental data come from the postulated model. These tests are a hypothesis testing problem, the problem concerned with the choice of one of the two alternative hypotheses

H_0 : Data come from the specified distribution.

H_1 : Data don't come from the specified distribution.

The goodness-of-fit tests have been focus of investigation for many authors, see for example, Pewsey and Kato (2016), Kreer et. al., (2015), Pakyari and Balakrishnan(2013), Cao et. al., (2010), Lim and Park (2007). We also, considered the Vuong (Vuong;1989) test as a model selection test. Based on this test, we can compare the rival models and then select the best one. The rest of the paper is organized as follows. In Section 2, we first discuss some properties of the NIG distribution. Then, we obtain the MLEs of the four unknown parameters using the EM algorithm. Section 3 provides the EM algorithm for the GHST distributions. The observed Fisher information matrix is presented in Section 4. Different goodness-of-fit and model selection tests are provided in Section 5, which enable us to discriminate the NIG and GHST distributions. Analysis of the daily closed Tehran stocks exchanges is presented in Section 6, and finally in Section 7, we

conclude the paper.

2. PARAMETERS ESTIMATION OF NIG USING THE EM ALGORITHM

A random variable X is said to has NIG distribution if the probability density function (PDF) is of the following form;

$$f_{NIG}(x; \alpha, \beta, \gamma, \delta) = \frac{\alpha\gamma}{\pi} \exp\left(\gamma \sqrt{\alpha^2 - \beta^2} + \beta(x - \delta)\right) \times \mathfrak{I}_1^*(x; \alpha, \gamma, \delta). \quad (1)$$

where, $\mathfrak{I}_1^*(x; \alpha, \gamma, \delta) = B_1(\alpha \sqrt{\gamma^2 + (x - \delta)^2}) / \sqrt{\gamma^2 + (x - \delta)^2}$ and $\delta \in R$ is a location parameter, $\gamma > 0$ serves for scaling, $\alpha > 0$ determines the shape, and $\beta (0 < |\beta| < \alpha)$ determines the skewness. Also, $B_1(\cdot)$ denotes the modified Bessel function of the third kind of index 1. The tail heaviness and skewness of the NIG distribution are measured by α and β , respectively. Now, we want to estimate the unknown parameters of the NIG distribution. Suppose that $\underline{X} = (X_1, \dots, X_n)$ are random variables from NIG distribution. The likelihood function becomes

$$l(\underline{x}, \alpha, \beta, \gamma, \delta) = \left(\frac{\alpha\gamma}{\pi}\right)^n \exp\left(n\gamma \sqrt{\alpha^2 - \beta^2} + \beta \sum_{i=1}^n (x_i - \delta)\right) \times \prod_{i=1}^n \mathfrak{I}_1^*(x_i; \alpha, \gamma, \delta). \quad (2)$$

Hence, the log-likelihood function can be written

$$L(\underline{x}, \alpha, \beta, \gamma, \delta) = n \log\left(\frac{\alpha\gamma}{\pi}\right) + n\gamma \sqrt{\alpha^2 - \beta^2} + \beta \sum_{i=1}^n (x_i - \delta) + \sum_{i=1}^n \log \mathfrak{I}_1^*(x_i; \alpha, \gamma, \delta).$$

Differentiating the log-likelihood function $L(\underline{x}, \alpha, \beta, \gamma, \delta)$ partially with respect to the unknown parameters and then equating to zero, we have

$$\frac{\partial L}{\partial \alpha} = \frac{n}{\alpha} + \frac{n\alpha\gamma}{\sqrt{\alpha^2 - \beta^2}} + \sum_{i=1}^n \left(\frac{\partial B_1 / \partial \alpha}{B_1}\right) = 0, \quad (3)$$

$$\frac{\partial L}{\partial \beta} = -\frac{n\beta\gamma}{\sqrt{\alpha^2 - \beta^2}} + \sum_{i=1}^n (x_i - \beta) = 0, \quad (4)$$

$$\frac{\partial L}{\partial \gamma} = \frac{n}{\gamma} + n \sqrt{\alpha^2 - \beta^2} + \sum_{i=1}^n \left(\frac{\partial B_1 / \partial \gamma}{B_1} - \frac{\gamma}{\gamma^2 + (x_i - \beta)^2}\right) = 0, \quad (5)$$

$$\frac{\partial L}{\partial \delta} = -n\beta + \sum_{i=1}^n \left(\frac{\partial B / \partial \delta}{B} - \frac{x_i - \delta}{\gamma^2 + (x_i - \beta)^2}\right) = 0, \quad (6)$$

where,

$$\frac{\partial B_1}{\partial \alpha} = -\left[\mathfrak{I}_2^*(x; \alpha, \gamma, \delta) + B_0(\alpha\iota)\right] \times \iota, \quad \frac{\partial B_1}{\partial \beta} = 0,$$

$$\frac{\partial B_1}{\partial \gamma} = - \left[\mathfrak{J}_2^*(x; \alpha, \gamma, \delta) + B_0(\alpha \iota) \right] \times \frac{\alpha \gamma}{\iota},$$

$$\frac{\partial B_1}{\partial \delta} = - \left[\mathfrak{J}_2^*(x; \alpha, \gamma, \delta) + B_0(\alpha \iota) \right] \times \frac{-\alpha(x - \delta)}{\iota}.$$

Where, $\mathfrak{J}_2^*(x; \alpha, \gamma, \delta) = B_1(\alpha \sqrt{\gamma^2 + (x - \delta)^2}) / \alpha \sqrt{\gamma^2 + (x - \delta)^2}$ and $\iota = \sqrt{\gamma^2 + (x - \delta)^2}$. Unfortunately, analytic solutions for the unknown parameters are not available. We propose to use the EM algorithm for estimating the unknown parameters. The EM algorithm is an iterative algorithm that converges under rather weak conditions to a local maximum of the likelihood function. The idea behind the algorithm is to augment the observed data with latent data, which can be either missing data or parameter values, so that the likelihood conditioned on the data and the latent data has a form that is easy to analyze. The EM algorithm is also suitable for mixture distributions, since the mixing operation in a sense produces missing data, the mixing variables. The algorithm can be broken down into two steps: the expectation (E) step, and the maximization (M) step. Now, we describe the steps of the EM algorithm for NIG distribution.

EM algorithm

Suppose that n observations from NIG distribution are available. First, we consider

$$X = \delta + \beta Z + \sqrt{Z} Y, \quad (7)$$

where, $Y \sim N(0, 1)$, $Z \sim IG(\gamma, \mathfrak{J})$ and Y and Z are independent. Also, $\mathfrak{J} = \sqrt{\alpha^2 + \beta^2}$. It is clear that

$$Z|x \sim GIG(-1, \nu(x), \alpha); \nu(x) = \sqrt{\gamma^2 + (x - \delta)^2}.$$

Note that the IG and GIG denote the inverse Gaussian and generalized inverse Gaussian, respectively. So, based on the equation (7), the log-likelihood function can be rewritten as

$$L(x, \alpha, \beta, \gamma, \delta) = \sum_{i=1}^n \log f_{NIG}(x; \alpha, \beta, \gamma, \delta)$$

$$= \sum_{i=1}^n \log f_{NIG}(x_i | z_i, \beta, \delta) + \sum_{i=1}^n \log f_{NIG}(z_i, \gamma, \mathfrak{J}) = C_1 + C_2,$$

where,

$$C_1 = -\frac{n}{2} \log(2\pi) - 2^{-1} \sum_{i=1}^n \log z_i - 2^{-1} \sum_{i=1}^n z_i^{-1} (x_i - \delta - \beta z_i)^2,$$

$$C_2 = n \log(\gamma) - \frac{\gamma^2}{2} \sum_{i=1}^n \log z_i^{-1} - \frac{\mathfrak{Y}^2}{2} \sum_{i=1}^n z_i + n\gamma\mathfrak{Y},$$

and $\mathfrak{Y} = \sqrt{\alpha^2 + \beta^2}$. For the E-step of the EM algorithm, one needs to compute $E(Z_i | X_i = x_i)$ and $E(Z_i^{-1} | X_i = x_i)$. Also, the M-step involves the maximization of C_1 and C_2 with respect to unknown parameters (for more detail see, Pradhan and Kundu(2009), and Karlis (2002)).

3. PARAMETERS ESTIMATION OF GHST USING THE EM ALGORITHM

A random variable X has the GHST distribution with parameters $\theta = (\eta, \beta, \delta, \mu)$, denoted $X \sim GHST(\eta, \beta, \delta, \mu)$, if it has the probability density function (Aas and Haff, 2006),

$$f_{GHST}(x; \eta, \beta, \delta, \mu) = \frac{2^{\frac{1-\eta}{2}} \delta^\eta |\beta|^{\frac{1+\eta}{2}} B_{\frac{1+\eta}{2}}(\sqrt{\beta^2(\delta^2 + (x - \mu)^2)}) \exp(\beta(x - \mu))}{\Gamma(\eta/2) \sqrt{\pi} (\sqrt{\delta^2 + (x - \mu)^2})^{\frac{1+\eta}{2}}}. \tag{8}$$

Here, $\beta \neq 0$ and using the following relation,

$$B_{\frac{1+\eta}{2}}\left(\sqrt{\beta^2(\delta^2 + (x - \mu)^2)}\right) \xrightarrow{\beta \rightarrow 0} \Gamma\left(\frac{1 + \eta}{2}\right) 2^{\frac{\eta-1}{2}} \left(\sqrt{\beta^2(\delta^2 + (x - \mu)^2)}\right)^{-\left(\frac{1+\eta}{2}\right)},$$

we have

$$f_{GHST}(x; \eta, \beta, \delta, \mu) = \frac{\Gamma\left(\frac{1+\eta}{2}\right) \delta^\eta}{\Gamma(\eta/2) \sqrt{\pi} (\sqrt{\delta^2 + (x - \mu)^2})^{1+\eta}}; \beta = 0.$$

Now, we estimate the unknown parameters using the maximum likelihood method. For applying the EM algorithm to estimate the unknown parameters of GHST distribution, we similarly write, $Z \sim GIG(-\eta/2, \delta)$. So, the joint density of X and Z is given by $f(x, z) = f(x|z)f(z)$. Here, $X|Z = z \sim G(\mu + \beta z, z)$ that G denotes the Gaussian distribution. So, based on equation (8), the log-likelihood function can be written as

$$L(\underline{x}, \eta, \beta, \delta, \mu) = \log l(\underline{x}, \eta, \beta, \delta, \mu) = \sum_{i=1}^n \log f_{x|z}(x|z_i, \beta, \mu) + \sum_{i=1}^n \log f_z(z_i; \delta, v),$$

where, $v = -\eta/2$. For the E-step of the EM algorithm, one needs to compute $o_1 = E(Z_i | X_i = x_i)$, $o_2 = E(Z_i^{-1} | X_i = x_i)$ and $o_3 = E(\log Z_i | X_i = x_i)$. We know that

$$Z \Big| X = x \sim GIG\left(-\frac{(\eta + 1)}{2}, \sqrt{\delta^2 + (x - \mu)^2}, |\beta|\right).$$

Thus the o_1 , o_2 and o_3 can be obtained easily. Now, in the M-step, one computes the parameter estimates resulting from maximizing the likelihood of $f(x, z) = f(x|z)f(z)$ using the pseudo values o_1 , o_2 and o_3 (see Appendix A) from the M-step.

4. THE OBSERVED FISHER INFORMATION MATRIX

In this section, we want to derive the observed Fisher information for the likelihood using (3), (4), (5), and (6), which will enable us to construct confidence intervals for the unknown parameters. We have,

$$I(\hat{\theta}) = I(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta}) = \begin{bmatrix} -\frac{\partial^2 L}{\partial \alpha^2} & -\frac{\partial^2 L}{\partial \alpha \partial \beta} & -\frac{\partial^2 L}{\partial \alpha \partial \gamma} & -\frac{\partial^2 L}{\partial \alpha \partial \delta} \\ -\frac{\partial^2 L}{\partial \beta \partial \alpha} & -\frac{\partial^2 L}{\partial \beta^2} & -\frac{\partial^2 L}{\partial \beta \partial \gamma} & -\frac{\partial^2 L}{\partial \beta \partial \delta} \\ -\frac{\partial^2 L}{\partial \gamma \partial \alpha} & -\frac{\partial^2 L}{\partial \gamma \partial \beta} & -\frac{\partial^2 L}{\partial \gamma^2} & -\frac{\partial^2 L}{\partial \gamma \partial \delta} \\ -\frac{\partial^2 L}{\partial \delta \partial \alpha} & -\frac{\partial^2 L}{\partial \delta \partial \beta} & -\frac{\partial^2 L}{\partial \delta \partial \gamma} & -\frac{\partial^2 L}{\partial \delta^2} \end{bmatrix}_{(\alpha, \beta, \gamma, \delta) = (\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta})}$$

where,

$$\frac{\partial^2 L}{\partial \alpha^2} = -\frac{n}{\alpha^2} - \frac{n\beta^2\gamma}{(\sqrt{\alpha^2 - \beta^2})^3} + \sum_{i=1}^n \left(\frac{(\partial^2 B_1 / \partial \alpha^2) B_1 - (\partial B_1 / \partial \alpha)^2}{B_1^2} \right),$$

$$\frac{\partial^2 L}{\partial \gamma^2} = -\frac{n}{\gamma^2} + \sum_{i=1}^n \left(\frac{(\partial^2 B_1 / \partial \gamma^2) B_1 - (\partial B_1 / \partial \gamma)^2}{B_1^2} + \frac{\gamma^2 - (x_i - \delta)^2}{(\gamma^2 + (x_i - \delta)^2)^2} \right),$$

$$\frac{\partial^2 L}{\partial \delta^2} = \sum_{i=1}^n \left(\frac{(\partial^2 B_1 / \partial \delta^2) B_1 - (\partial B_1 / \partial \delta)^2}{B_1^2} + \frac{(x_i - \delta)^2 - \gamma^2}{(\gamma^2 + (x_i - \delta)^2)^2} \right),$$

$$\frac{\partial^2 L}{\partial \alpha \partial \beta} = -\frac{n\alpha\beta\gamma}{(\sqrt{\alpha^2 - \beta^2})^3}, \quad \frac{\partial^2 L}{\partial \beta^2} = -\frac{n\alpha^2\gamma}{(\sqrt{\alpha^2 - \beta^2})^3},$$

$$\frac{\partial^2 L}{\partial \alpha \partial \gamma} = \frac{n\alpha}{\sqrt{\alpha^2 - \beta^2}} + \sum_{i=1}^n \left(\frac{(\partial^2 B_1 / \partial \alpha \partial \gamma) B_1 - (\partial B_1 / \partial \alpha) (\partial B_1 / \partial \gamma)}{B_1^2} \right),$$

$$\frac{\partial^2 L}{\partial \alpha \partial \delta} = \sum_{i=1}^n \left(\frac{(\partial^2 B_1 / \partial \alpha \partial \delta) B_1 - (\partial B_1 / \partial \alpha) (\partial B_1 / \partial \delta)}{B_1^2} \right),$$

$$\frac{\partial^2 L}{\partial \beta \partial \gamma} = -\frac{n\beta}{\sqrt{\alpha^2 - \beta^2}}, \quad \frac{\partial^2 L}{\partial \beta \partial \delta} = -n,$$

$$\frac{\partial^2 L}{\partial \gamma \partial \delta} = \sum_{i=1}^n \left(\frac{(\partial^2 B_1 / \partial \gamma \partial \delta) B_1 - (\partial B_1 / \partial \gamma) (\partial B_1 / \partial \delta)}{B_1^2} - \frac{2\gamma(x_i - \delta)}{(\gamma^2 + (x_i - \delta)^2)^2} \right),$$

and also,

$$\begin{aligned} \frac{\partial^2 B_1}{\partial \alpha^2} &= \left[\frac{(2 + (\alpha A)^2)^2 B_1(\alpha A)}{(\alpha A)^2} + \frac{B_0(\alpha A)}{\alpha A} \right] \times A^2, \\ \frac{\partial^2 B_1}{\partial \gamma^2} &= \left[\frac{(2 + (\alpha A)^2)^2 B_1(\alpha A)}{(\alpha A)^2} + \frac{B_0(\alpha A)}{\alpha A} \right] \times \left(\frac{\alpha \gamma}{A} \right)^2 \\ &\quad - \left[\frac{B_1(\alpha A)}{\alpha A} + B_0(\alpha A) \right] \times \frac{\alpha(x - \delta)^2}{A^3}, \\ \frac{\partial^2 B_1}{\partial \delta^2} &= \left[\frac{(2 + (\alpha A)^2)^2 B_1(\alpha A)}{(\alpha A)^2} + \frac{B_0(\alpha A)}{\alpha A} \right] \times \left(\frac{-\alpha(x - \delta)}{A} \right)^2 \\ &\quad - \left[\frac{B_1(\alpha A)}{\alpha A} + B_0(\alpha A) \right] \times \frac{\alpha \gamma^2}{(A)^3}, \\ \frac{\partial^2 B_1}{\partial \alpha \partial \beta} &= 0, \quad \frac{\partial^2 B_1}{\partial \beta^2} = 0, \quad \frac{\partial^2 B_1}{\partial \beta \partial \gamma} = 0, \quad \frac{\partial^2 B_1}{\partial \beta \partial \delta} = 0, \\ \frac{\partial^2 B_1}{\partial \alpha \partial \gamma} &= \left[\frac{(2 + (\alpha A)^2)^2 B_1(\alpha A)}{(\alpha A)^2} + \frac{B_0(\alpha A)}{\alpha A} \right] \times (\alpha \gamma) \\ &\quad - \left[\frac{B_1(\alpha A)}{\alpha A} + B_0(\alpha A) \right] \times \frac{\gamma}{A}, \\ \frac{\partial^2 B_1}{\partial \alpha \partial \delta} &= \left[\frac{(2 + (\alpha A)^2)^2 B_1(\alpha A)}{(\alpha A)^2} + \frac{B_0(\alpha A)}{\alpha A} \right] \times (-\alpha(x - \delta)) \\ &\quad - \left[\frac{B_1(\alpha A)}{\alpha A} + B_0(\alpha A) \right] \times \frac{-(x - \delta)}{A}, \\ \frac{\partial^2 B_1}{\partial \gamma \partial \delta} &= \left[\frac{(2 + (\alpha A)^2)^2 B_1(\alpha A)}{(\alpha A)^2} + \frac{B_0(\alpha A)}{\alpha A} \right] \times \left(\frac{-\alpha^2 \gamma (x - \delta)}{A^2} \right) \\ &\quad - \left[\frac{B_1(\alpha A)}{\alpha A} + B_0(\alpha A) \right] \times \frac{\alpha \gamma (x - \delta)}{A^3}, \end{aligned}$$

where, $A = \sqrt{\gamma^2 + (x - \delta)^2}$. The observed Fisher information matrix can be inverted to obtain a local estimate of the asymptotic variance-covariance matrix of the MLE. Also, the A , two-sided normal approximation confidence interval for $\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta})$ can be obtained by using the asymptotic variance-covariance matrix. Note that the observed Fisher information matrix for the GHST distribution can be obtained similarly.

5. GOODNESS-OF-FIT (GOF) AND MODEL SELECTION TESTS

In this section, we describe different available methods for choosing the best fitted model to a given dataset. Suppose that there are two families, the problem is to choose the correct family for a given dataset $\{x_1, \dots, x_n\}$.

5.1. The Kolmogorov-Smirnov Test (K-S Test)

The K-S test is one of the important tests for GOF. It is used to decide if a sample comes from a population with a specific distribution. The K-S test is based on the empirical distribution function as $F_n(x) = \frac{\text{Number of } x_i \leq x}{n}$. Based on the two competitive models F and G , the K-S distances are calculated as

$$D_1 = \sup_{-\infty < x < \infty} |F_n(x) - F(x)|, \quad (9)$$

$$D_2 = \sup_{-\infty < x < \infty} |G_n(x) - G(x)|. \quad (10)$$

To implement this procedure, a candidate from each parametric family that has the smallest Kolmogorov distance should be found, and then, different best fitted distributions should be compared.

5.2. The Anderson Darling Test (A-D Test)

The A-D test is the GOF test of whether a given sample of data is drawn from a population with a specific distribution.. It is a modification of the K-S test and gives more weight to the tails than the K-S test does.

$$A_n^2 = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{F(x)(1 - F(x))} dF(x). \quad (11)$$

The critical value for the significance level of 95% is given by 2.49. Also, the computational form of (11) can be written as

$$A_n^2 = -n \left[1 + n^{-2} \sum_{i=1}^n (2i - 1) \log(F(x_{(i)})(1 - F(x_{(n-i+1)}))) \right]. \quad (12)$$

where, $x_{(i)}$ is the ordered data.

5.3. The Vuong Test

The Vuong (1989) test is designed to compare the two rival models. This test is based on the likelihood ratio (LR) statistic. The LR statistic for the model $F(.; \theta_1) = \{f(.; \theta_1), \theta_1 \in \Theta_1 \subset R^p\}$ against the model $G(.; \theta_2) = \{g(.; \theta_2), \theta_2 \in \Theta_2 \subset R^q\}$ is:

$$LR_n(\hat{\theta}_1, \hat{\theta}_2) = \sum_{i=1}^n \log \frac{f(x; \hat{\theta}_1)}{g(x; \hat{\theta}_2)},$$

where, $\hat{\theta}_1$, and $\hat{\theta}_2$ are the maximum likelihood estimators. The null hypothesis of the Vuong test is

$$H_0 : D_{KL}(h \| f) = D_{KL}(h \| g),$$

where, $D_{KL}(h \| \cdot)$ is the Kullback-Leibler divergence (Kullback and Leibler, 1951) from the true model as

$$D_{KL}(h \| f) = E_h \left(\log \frac{h}{f(X; \theta_2)} \right). \quad (13)$$

Here, h is the true model. In other words, the hypotheses of this test can be written as

$$H_0 : E_h \left(\log \frac{f(X)}{g(X)} \right) = 0, H_f : E_h \left(\log \frac{f(X)}{g(X)} \right) > 0, H_g : E_h \left(\log \frac{f(X)}{g(X)} \right) < 0.$$

For discriminating the two models using the Vuong test, we consider the following Steps:

Step 1: Estimate the unknown parameters of the two models using the maximum likelihood procedure.

Step 2: Calculate $LR_n(\hat{\theta}_1, \hat{\theta}_2) = \sum_{i=1}^n \log \frac{f(x; \hat{\theta}_1)}{g(x; \hat{\theta}_2)}$.

Step 3: Obtain the $\mathfrak{V} = \frac{LR_n(\hat{\theta}_1, \hat{\theta}_2)}{\sqrt{n\hat{\psi}_n^2}}$, where $\hat{\psi}_n^2$ is the empirical variance of the $LR_n(\hat{\theta}_1, \hat{\theta}_2)$ as

$$\hat{\psi}_n^2 = \frac{1}{n} \sum_{i=1}^n \left(\log \frac{f(x_i; \hat{\theta}_1)}{g(x_i; \hat{\theta}_2)} \right)^2 - \left(\frac{1}{n} \sum_{i=1}^n \left(\log \frac{f(x_i; \hat{\theta}_1)}{g(x_i; \hat{\theta}_2)} \right) \right)^2. \quad (14)$$

Step 4: Discriminate about the two models as:

- If the value of the statistic \mathfrak{V} is higher than $Z_{1-\alpha}$ (the $(1 - \alpha)^{th}$ quantile of standard normal distribution), then one rejects the null hypothesis that the models are equivalent in favor of $F(.; \theta_1)$ being better than $G(.; \theta_2)$.

- If \mathfrak{V} is smaller than $-Z_{1-\alpha}$ then, one rejects the null hypothesis in favor of $G(.; \theta_2)$ being better than $F(.; \theta_1)$.

- If $|\mathfrak{V}| < Z_{1-\alpha}$ then, one cannot discriminate between the two rival models based on the given data (two models equal).

6. THE STOCK EXCHANGE DATA ANALYSIS

The data used in this research is the daily closed Tehran stock exchanges. We considered the four shares as prices and cash returns index (A), second Market Index (B), first Market Index (C), and industry Index (D). We used the daily log returns data ($r_t = \log(\mathfrak{J}_t/\mathfrak{J}_{t-1})$) which are obtained totaling a series of N -1 observations for Tehran listed shares. Note that N is the number of the closing prices, and $[\mathfrak{J}_t]_{t \geq 0}$ is the closing price at day t . All the results are obtained using the R program. We want to test the following hypothesis for the proposed data sets.

H_0 : the data are NIG distributed with the parameters $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta})$.

H_0 : the data are GHST distributed with the parameters $(\hat{\eta}, \hat{\beta}, \hat{\delta}, \hat{\rho})$.

We first considered the Normality tests such as Kolmogorov-Smirnov (K-S norm), Shapiro-Wilk (S-W), Jarque- Bera (J-B), and Lillie (L) tests. For each test, the null hypothesis ($H_{0Normal}$) is that the daily log returns are normally distributed. It is observed that all p -values ($< 2.2e^{-16}$) are less than the significant level (0.05). So, the proposed null hypothesis ($H_{0Normal}$) will be rejected. Now, we want to compare the NIG and GHST distributions for the proposed data. For this purpose, we considered the following Steps:

Step 1: Computing the moment estimates (MME) of the parameters (see Appendix B).

Step 2: Econometrically estimating the unknown parameters of the NIG and GHST distributions using the EM algorithm. To calculate the EM algorithm for the real data set, we need to specify the initial values of the parameters. These were taken to be their MMEs.

Step 3: Computing different criteria for comparing the NIG and GHST distributions. We used the Akaike information criterion ($AIC = (parameters \times 2) - 2 \ln L$), the Bayesian information criterion ($BIC = (parameters \times \ln n) - 2 \ln L$), and the maximum log-likelihood value ($LL = - \ln L$).

Step 4: Comparing the NIG and GHST distributions using the p -values of Kolmogorov-Smirnov and the Anderson-Darling tests.

Step 5: Obtaining the Vuong statistic for selecting the best model.

Step 6: Constructing the 95% approximation normal confidence intervals (ACs) for NIG and GHST distributions using the asymptotic variance-covariance matrices.

Tables 1 and 2 present the results from the EM algorithm for the NIG and GHST

distributions, respectively. The AIC, BIC, and LL values for the distributions are presented in Table 3. From the results of Table 3, it is observed that the AIC and BIC criteria of the GHST are smaller than the NIG. Also, the LL of the NIG is larger than GHST distribution. So, the GHST is fitter than the NIG for these data. The results of the Kolmogorov-Smirnov and the Anderson-Darling tests are reported in Table 4. Based on the limited set of stock exchange data, and using several statistical criteria, minimum AIC, minimum BIC, maximum LL value, and the high p -values of the K-S and A-D tests, the GHST distribution function appears to be more appropriate statistical distribution function. Note that we avoid claiming that the NIG is not the appropriate distribution for these data.

It is observed that the steepness, $\zeta = [1 + \gamma(\alpha^2 - \beta^2)^{1/2}]^{-1/2}$, and the asymmetry, $\chi = \beta\zeta/\alpha$, are cited in the domain of variation $\{(\chi, \zeta) : -1 < \chi < 1, 0 < \zeta < 1\}$. For more comparison, we consider the Young test to confirm the results of the previous different tests and criteria. For this purpose, we considered the GHST ($F(\cdot; \theta_1)$), and NIG ($G(\cdot; \theta_2)$) as the two rival models. For different data sets (A, B, C and D), the Young statistic (\mathfrak{Y}) are (3.456), (4.220), (3.874), and (4.628), respectively. It is observed that all the values are higher than $Z_{0.95}$. Thus, we can conclude that the $F(data; \eta, \beta, \delta, \mu) \approx GHST$ is better than $G(data; \alpha, \beta, \delta, \gamma) \approx NIG$ to estimate the true model for a given data. Finally, we calculated the ACs using the asymptotic variance-covariance matrix, which was considered in Section 4. The results for NIG and GHST distributions are reported in Tables 5 and 6, respectively.

Table 1: Parameters estimation for the NIG distribution using the EM algorithm.

Source	α	β	γ	δ
A	7.98728×10^{-1}	-1.89975×10^{-2}	2.89654×10^{-4}	6.3512×10^{-5}
B	3.82744×10^1	-1.17558×10^1	3.29206×10^{-4}	2.3528×10^{-5}
C	42.5695×10^{-1}	-10.1423×10^{-1}	3.42852×10^{-4}	1.4104×10^{-5}
D	1.07788×10^1	-27.8912×10^{-1}	3.33931×10^{-4}	1.4663×10^{-6}

Table 2: Parameters estimation for the GHST distribution using the EM algorithm.

Source	η	β	δ	μ
A	1.04300	-10.811×10^{-2}	3.024×10^{-4}	-7.038×10^{-6}
B	2.06200	-62.301×10^{-1}	6.251×10^{-4}	-4.282×10^{-5}
C	1.74900	11.860×10^{-1}	5.807×10^{-4}	2.109×10^{-6}
D	1.55600	-1.1142×10^{-1}	5.083×10^{-4}	-1.065×10^{-5}

7. CONCLUSION

In this paper, we evaluated the performance of the NIG and GHST distributions in characterizing the Tehran index returns. It is observed that when all pa-

Table 3: The AIC, BIC and LL statistic for the NIG and GHST distributions.

Distribution	Criteria	A	B	C	D
GHST	AIC	-12991.67	-13846.24	-13850.36	-13883.45
	BIC	-12991.81	-13846.37	-13850.49	-13883.55
	LL	6498.836	6926.121	6828.181	6894.723
NIG	AIC	-12850.59	-13737.14	-13712.63	-13815.07
	BIC	-12850.64	-13737.26	-13712.72	-13815.19
	LL	6428.296	6891.569	6809.317	6810.534

Table 4: The goodness-of-fit test for the NIG and GHST distribution.

Distribution	<i>p</i> -values	A	B	C	D
GHST	K-S	0.7832	0.8322	0.7759	0.8027
	A-D	0.8219	0.8601	0.8097	0.8412
NIG	K-S	0.5973	0.6347	0.5889	0.6099
	A-D	0.6275	0.6690	0.6220	0.6638

Table 5: The lengths of the 95% ACs for the NIGs parameters.

Parameters	A	B	C	D
α	2.917	2.176	2.265	3.004
β	1.365	1.701	1.438	1.623
γ	1.895	2.004	1.785	2.038
δ	2.643	2.758	2.446	2.226

Table 6: The lengths of the 95% ACs for the NIGs parameters.

Parameters	A	B	C	D
η	3.124	2.643	2.648	3.633
β	1.784	1.989	1.873	1.978
δ	2.112	2.492	1.899	2.665
μ	2.875	3.016	2.780	2.839

parameters are unknown, the maximum likelihood estimates of these distributions cannot be obtained in the explicit form. Thus, we have used the EM algorithm to estimate the parameters. The approximate confidence intervals for the unknown parameters have been constructed using the asymptotic variance-covariance matrix. Furthermore, we have considered the NIG and GHST distributions to critically analyze the data of Tehran stock exchange. Using several criteria and model selection tests, the GHST distribution is shown to be more appropriately than the NIG distribution for these data sets.

REFERENCES

- [1] Azzalini, A., and Capitanio, A., "Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t-Distribution", *Journal of the Royal Statistical Society, Series B, Statistical Methodology*, 65 (2) (2003) 367-389.
- [2] Aas, K., and Haff, I.H., "The Generalized Hyperbolic Skew Students t-Distribution", *Journal of Financial Econometrics*, 4 (2) (2006) 275309.
- [3] Barndorff-Nielsen, O.E., "Normal inverse Gaussian processes and stochastic volatility modeling", *Scandinavian Journal of Statistics*, 24 (1) (1997) 1-13.
- [4] Blasild, B., "The two-dimensional hyperbolic distribution and related distributions, with an application to Johanssen's bean data", *Biometrika*, 68 (1) (1981) 251-263.
- [5] Cao, J., Moosman, A., and Johnson, V.E., "A Bayesian chi-squared goodness-of-fit test for censored data models", *Biometrics*, 66 (2) (2010) 426-34.
- [6] Goncu, A., and Yang, H., "Variance-Gamma and Normal-Inverse Gaussian models: Goodness-of-fit to Chinese high-frequency index returns", *The North American Journal of Economics and Finance*, 36 (2016) 279-292.
- [7] Dempster, A.P., Laird, N.M., and Rubin, D.B., "Maximum likelihood from Incomplete Data via the EM Algorithm", *Journal of the Royal Statistical Society, Series B*, 39 (1) (1997) 1-38.
- [8] Karlis, D., "An EM type algorithm for maximum likelihood estimation of the normalinverse Gaussian distribution", *Statistics and Probability Letters*, 57 (1) (2002) 43-52.
- [9] Kreer, M., Kzlers, A., Thomas, A.W., Egdio dos Reis, and Alfredo, D., "Goodness-of-fit tests and applications for left-truncated Weibull distributions to non-life insurance", *European Actuarial Journal*, 5 (1) (2015) 139163.
- [10] Kullback, S., and Leibler, R., "On Information and Sufficiency", *Annals of Mathematical Statistics*, 22 (1) (1951) 7986.
- [11] Lee, S.X., and McLachlan, G.J., "Finite mixtures of canonical fundamental skew t-distributions", *Statistics and Computing*, 26 (3) (2016) 573589.
- [12] Lillestøl, J., *Fat and skew : can NIG cure? : on the prospects of using the normal inverse Gaussian distribution in finance*, Department of Finance and Management Science, The Norwegian School of Economics and Business Administration, Bergen, Norway, 1998.
- [13] Lillestøl, J., "Risk analysis and the NIG distribution", *Journal of Risk*, 2 (4) (2000) 41-56.
- [14] Lim, J., and Park, S., "Censored Kullback-Leibler information and Goodness-of-Fit test with type II censored data", *Journal of Applied Statistics*, 34 (9) (2007) 1051-1064.
- [15] Mabitsela, L., Mar, E., and Kufakunesu, R., "Quantification of VaR: A Note on VaR Valuation in the South African Equity Market", *Journal of Risk Financial Management*, 8 (1) (2015) 103-126.
- [16] Nakajima, J., "Bayesian analysis of multivariate stochastic volatility with skew distribution", *Econometric Reviews*, 43 (2014) 1-23.
- [17] Necula, C., "Modeling Heavy-Tailed stock index returns using the generalized hyperbolic distribution", *Romanian Journal of Economic Forecasting* 6 (2) (2009) 118-131.
- [18] Ng, T., Chan, C.S., and Balakrishnan, N., "Estimation of parameters from progressively censored data using EM algorithm", *Computational Statistics and Data Analysis*, 39 (4) (2002) 371- 386.

- [19] Pakyari, R., and Balakrishnan, N., "Goodness-of-fit tests for progressively Type-II censored data from location-scale distributions", *Journal of Statistical Computation and Simulation*, 83 (1) (2013) 167-178.
- [20] Panahi, H., "Estimation of the Burr type III distribution with application in unified hybrid censored sample of fracture toughness", *Journal of applied Statistics*, (2017) in press.
- [21] Panahi, H., and Sayyareh, A., "Estimation and prediction for a unified hybrid-censored Burr Type XII distribution", *Journal of Statistical Computation and Simulation*, 86 (1) (2016) 55-73.
- [22] Panahi, H., "Model selection test for the heavy-tailed distributions under censored samples with application in financial data", *International Journal of Financial Studies*, 4 (4) (2016) 1-14.
- [23] Pewsey, R., and Kato, S., "Parametric bootstrap goodness-of-fit testing for WehrlyJohnson bivariate circular distributions", *Statistics and Computing*, 26 (6) (2016) 1307-1317.
- [24] Pradhan, B., and Kundu, D., "On Progressively Censored Generalized Exponential Distribution", *Test*, 43 (2009) 1-17.
- [25] Rastogi, M.K., and Tripathi, Y.M., "Estimating the parameters of a Burr Distribution under progressive type II censoring", *Statistical Methodology*, 9 (3) (2012) 381-391.
- [26] Vuong, Q.H., "Likelihood ratio tests for model selection and Non-nested hypothesis", *Econometrica*, 57 (2) (1989) 307333.

APPENDIX A:

We know that $W = Z|X$ has the generalized inverse Gaussian (GIG) distribution with the following distribution function:

$$f_W\left(-\frac{(\eta+1)}{2}, \sqrt{\delta^2 + (x-\mu)^2}, |\beta|\right) = \left(\frac{|\beta|}{\sqrt{\delta^2 + (x-\mu)^2}}\right)^{-\frac{(\eta+1)}{2}} \frac{w^{-\frac{(\eta+1)}{2}-1}}{2B_{-\frac{(\eta+1)}{2}}(|\beta| \sqrt{\delta^2 + (x-\mu)^2})} e^{-\frac{1}{2}\left\{\frac{(\delta^2+(x-\mu)^2)}{w} + \beta^2 w\right\}}.$$

Note: if $X \sim GIG(\alpha, \beta, \gamma) \Rightarrow X^{-1} \sim GIG(-\alpha, \gamma, \beta)$.

So, based on the properties of the GIG distribution, the values of $E(Z_i|X_i = x_i)$, $E(Z_i^{-1}|X_i = x_i)$ and $E(\log Z_i|X_i = x_i)$ can be written as

$$E(Z_i|X_i = x_i) = \frac{\sqrt{\delta^2 + (x_i - \mu)^2} B_{-\frac{(\eta+1)}{2}+1}(|\beta| \sqrt{\delta^2 + (x_i - \mu)^2})}{|\beta| B_{\frac{1+\eta}{2}}(|\beta| \sqrt{\delta^2 + (x_i - \mu)^2})},$$

$$E(Z_i^{-1}|X_i = x_i) = |\beta| B_{\frac{3+\eta}{2}}(|\beta| \sqrt{\delta^2 + (x_i - \mu)^2}) \times \left\{ \sqrt{\delta^2 + (x_i - \mu)^2} B_{\frac{1+\eta}{2}}(|\beta| \sqrt{\delta^2 + (x_i - \mu)^2}) \right\}^{-1},$$

and

$$E(\log Z_i|X_i = x_i) = \log\left(\frac{\sqrt{\delta^2 + (x_i - \mu)^2}}{|\beta|}\right)$$

$$+ \frac{\partial}{\partial(\frac{\eta+1}{2})} \left[\log \left(B_{\frac{1+\eta}{2}}(|\beta| \sqrt{\delta^2 + (x_i - \mu)^2}) \right) \right],$$

APPENDIX B:

The method of moment (MM) is probably the oldest method for estimating the unknown parameters. Suppose that $X = (X_1, \dots, X_n)$ is the random variable with the probability density function $f(x, \theta_1, \dots, \theta_k)$. The method of moments is a technique for constructing estimators of the parameters that is based on matching the sample moments ($m_j = \frac{1}{n} \sum_{i=1}^n X_i^j$) with the corresponding distribution moments ($\mu_j = E(X^j)$) as $E(X^j) = \frac{1}{n} \sum_{i=1}^n X_i^j$; $j = 1, 2, \dots, k$. Also, we can consider the following equations:

$$\mu_1 = E(X) = \frac{1}{n} \sum_{i=1}^n X_i, \quad E(X - \mu_1)^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2,$$

$$E(X - \mu_1)^3 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^3, \quad E(X - \mu_1)^4 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^4.$$