Comparison of parameter estimation methods for normal inverse Gaussian distribution

Jeongyoen Yoon[a], Jiyeon Kim[b], Seongjoo Song[1,a]

[a]Department of Statistics, Korea University, Korea;
[b]Actuarial Team, KYOBO Life Insurance Co.,Ltd., Korea

Abstract

This paper compares several methods for estimating parameters of normal inverse Gaussian distribution. Ordinary maximum likelihood estimation and the method of moment estimation often do not work properly due to restrictions on parameters. We examine the performance of adjusted estimation methods along with the ordinary maximum likelihood estimation and the method of moment estimation by simulation and real data application. We also see the effect of the initial value in estimation methods. The simulation results show that the ordinary maximum likelihood estimator is significantly affected by the initial value; in addition, the adjusted estimators have smaller root mean square error than ordinary estimators as well as less impact on the initial value. With real datasets, we obtain similar results to what we see in simulation studies. Based on the results of simulation and real data application, we suggest using adjusted maximum likelihood estimates with adjusted method of moment estimates as initial values to estimate the parameters of normal inverse Gaussian distribution.

Keywords: normal inverse Gaussian distribution, feasible domain, maximum likelihood estimation, method of moments estimation, adjusted estimation

1. Introduction

The generalized hyperbolic distribution introduced in Barndorff-Nielsen (1977) has often been used to model return distributions in financial markets since its application to DAX stock returns by Eberlein and Keller (1995). Variance gamma distribution by Madan and Seneta (1990), the skewed Student's *t* distribution used by Hansen (1994), and the normal inverse Gaussian distribution introduced by Barndorff-Nielsen (1997) are special cases of the generalized hyperbolic distribution. The generalized hyperbolic distribution has advantages in capturing stylized empirical facts of financial data such as high peaks and fat tails. Prause (1999) applied the generalized hyperbolic distribution to German stocks and American indices; in addition, Fajardo and Farias (2004) applied it to Brazilian market data and made a review of its empirical features. Behr and Pötter (2009) conducted the goodness-of-fit test of the generalized hyperbolic distribution to the return data of S&P500.

Normal inverse Gaussian (NIG) distribution which is a subclass of the generalized hyperbolic class of distributions has been successfully used in financial literature. Rydberg (1997) examined the performance of the NIG distribution using German and Danish stocks and Eriksson *et al.* (2009) studied the option pricing problem using NIG distribution and conducted empirical analysis using S&P500. They linked the pricing of derivatives to the moments of the risk-neutral density and showed

---

[1] Corresponding author: Department of Statistics, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, Korea.
  E-mail: sjsong@korea.ac.kr

that the option price can be approximated by a relatively simple function of the moments of the NIG distribution.

Some research has been done on the parameter estimation of NIG distribution; for instance, Figueroa-López *et al.* (2011) and Ghysels and Wang (2014). Figueroa-López *et al.* (2011) proposed an approximation method based on moments of NIG distribution and variance gamma distribution for high-frequency financial data. Ghysels and Wang (2014) provided a simple expression for moment estimators in NIG, variance gamma, and the skewed *t* distribution as well as introduced a feasible domain of the moments. Feasible domain is the set of skewness and excess kurtosis in which the moment estimators of NIG, variance gamma, and the skewed *t* distribution must reside. Even when true moment values are in the feasible domain, it is possible that the estimates can be outside of the domain, which makes the parameter estimation difficult. Yoon and Song (2016) examined this problem and proposed the epsilon estimation for NIG parameters. In this paper, we compare the performance of parameter estimation methods for NIG parameters, considering the problem related to the feasible domain. Ordinary maximum likelihood estimator (MLE), ordinary method of moments estimator (MME), epsilon MLE and epsilon MME by Yoon and Song (2016) as well as exponential MLE such as used by Kim (2019) are compared through simulation and real data application. Using the comparison, we suggest a stable and reliable method that can be used to estimate the parameters of NIG distribution.

The remainder of the paper is organized as follows. In Section 2, we define the NIG distribution and its feasible domain. Section 3 explains the adjusted estimation methods to be compared in later sections. In Section 4, simulation and real data application results using S&P 500 return data, U.S. indemnity loss data, and Danish fire loss data are provided. Section 5 finally concludes the paper.

2. The normal inverse Gaussian distribution

The probability density function of the NIG distribution is as follows. For $-\alpha < \beta < \alpha$, $\alpha > 0$, and $\delta > 0$,

$$f_{\text{NIG}}(x; \alpha, \beta, \delta, \mu) = \alpha\delta \frac{e^{\delta\gamma + \beta(x-\mu)}}{\pi \sqrt{\delta^2 + (x-\mu)^2}} K_1\left(\alpha \sqrt{\delta^2 + (x-\mu)^2}\right), \tag{2.1}$$

where K_1 is the modified Bessel function of the second kind with index 1 and γ is $\sqrt{\alpha^2 - \beta^2}$. α controls the steepness of the density and reflects the tail behavior, in the sense that larger values imply lighter tails and smaller values imply heavier tails. The density with $\beta < 0$ is skewed to the left whereas the density with $\beta > 0$ is skewed to the right and the density with $\beta = 0$ is symmetric. μ and δ are the location and the scale parameters, respectively. The parameters of the NIG distribution are generally estimated by the MLE since the density of NIG has the closed form expression.

With the NIG distribution, the population moments are expressed in terms of α, β, δ, and μ as follows (Eriksson *et al.*, 2009).

$$M = \mu + \delta \frac{\beta}{(\alpha^2 - \beta^2)^{\frac{1}{2}}}, \quad V = \frac{\delta\alpha^2}{(\alpha^2 - \beta^2)^{\frac{3}{2}}}, \quad S = \frac{3\beta}{\alpha\delta^{\frac{1}{2}}(\alpha^2 - \beta^2)^{\frac{1}{4}}}, \quad K = \frac{3\left(1 + 4\beta^2/\alpha^2\right)}{\delta(\alpha^2 - \beta^2)^{\frac{1}{2}}}, \tag{2.2}$$

where M, V, S, and K denote the mean, the variance, the skewness, and the excess kurtosis. The excess kurtosis is the kurtosis minus 3 so that a positive excess kurtosis means a fatter tail than normal distribution. Throughout this paper, we denote the sample mean, the sample variance, the sample skewness, and the sample excess kurtosis as $\hat{M}$, $\hat{V}$, $\hat{S}$, and $\hat{K}$, respectively.

From (2.2), we can rewrite α, β, δ, and μ as

$$\gamma = \frac{3}{\sqrt{V} \sqrt{3K - 5S^2}},$$

$$\beta = \frac{S \sqrt{V} \gamma^2}{3}, \quad \alpha = \sqrt{\gamma^2 + \beta^2}, \quad \delta = \frac{V\gamma^3}{\gamma^2 + \beta^2}, \quad \mu = M - \frac{\beta\delta}{\gamma}. \tag{2.3}$$

As can be seen in (2.3), these parameters are not defined when $3K - 5S^2$ is not positive. Ghysels and Wang (2014) introduced the feasible domain which is the set of excess kurtosis and skewness allowed for the NIG distribution such as

$$\left\{ (K, S) : 3K - 5S^2 > 0 \right\}.$$

When the underlying distribution is NIG distribution, (K, S) should be contained in the feasible domain. However, there is no guarantee that $(\hat{K}, \hat{S})$ calculated from random samples of the NIG distribution is always in the feasible domain. In the case where $(\hat{K}, \hat{S})$ is outside of the feasible domain, it can be seen in Yoon and Song (2016) that MMEs are not defined and MLEs become unstable. For more details, refer Yoon and Song (2016).

3. Adjusted methods of parameter estimation

3.1. Exponential maximum likelihood estimator

To impose the conditions on the estimators of NIG parameters such as $-\alpha < \beta < \alpha$, $\alpha > 0$, and $\delta > 0$, we write $\alpha = \sqrt{\beta^2 + e^u}$ and $\delta = e^w$ and define a new parameterization with u, β, w, and μ as follows:

$$f_{\mathrm{NIG}}(x; u, \beta, w, \mu) = e^w \sqrt{\beta^2 + e^u} \frac{e^{e^{w + \frac{2}{u} + \beta(x - \mu)}}}{\pi \sqrt{e^{2w} + (x - \mu)^2}} K_1\left(\sqrt{\beta^2 + e^u} \sqrt{e^{2w} + (x - \mu)^2} \right),$$

where $u = \log(\alpha^2 - \beta^2)$ and $w = \log(\delta)$. We will refer the MLE under this parameterization as exp-MLE, which was used in, for instance, Kim (2019).

3.2. Epsilon estimation

By (2.2), we can easily see that the following relationship holds:

$$3K - 5S^2 = \frac{9 \sqrt{\alpha^2 - \beta^2}}{\alpha^2 \delta}. \tag{3.1}$$

Thus, it is obvious that $3K - 5S^2$ must be positive. However, when we try to fit NIG distribution to real data and compute sample moments, we sometimes observe that $3\hat{K} - 5\hat{S}^2$ is negative. In such cases, parameter estimation using NIG distribution may have a problem as indicated in Yoon and Song (2016). By Yoon and Song (2016), ordinary MLE becomes unstable and ordinary MME can not be computed when $3\hat{K} - 5\hat{S}^2 \leq 0$. To overcome this problem, Yoon and Song (2016) set up a value ε to replace $3\hat{K} - 5\hat{S}^2$ when $3\hat{K} - 5\hat{S}^2 < \varepsilon$. They called this as ε-estimation method, and we will refer the MLE and the MME with $3\hat{K} - 5\hat{S}^2$ replaced by $\max\{3\hat{K} - 5\hat{S}^2, \varepsilon\}$ as ε-MLE and ε-MME, respectively. To choose the specific value for ε, they calculated the ε-maximum likelihood estimates using 1,000 simulated observations based on a set of parameters with ε changing in a relatively wide range. Then they repeated the same procedure 1,000 times to compute the root mean square error (RMSE) of ε-MLE and chose the value of ε that gives a small RMSE overall.
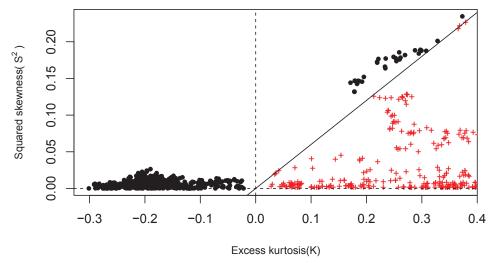
Figure 1: *Sample excess kurtosis vs sample squared skewness of S&P500 daily log returns from 2001/01/02 to 2010/12/31; each point is obtained using 250 daily log returns. Straight line: $3K - 5S^2 = 0$; +: $3K - 5S^2 > 0$; dot: $3K - 5S^2 < 0$.*

Table 1: Parameter values used in simulation for choosing ε: $3K - 5S^2 = 0.0061$ in Set 1, $3K - 5S^2 = 0.0128$ in Set 2, and $3K - 5S^2 = 0.0288$ in Set 3

	M	V	S	K	α	β	δ	μ
Set 1	−0.0011	0.0002	0.4747	0.3776	16731.2	16509.2	0.0143	−0.0881
Set 2	0.0004	0.0002	0.1338	0.0341	2906.1	2219.5	0.1563	−0.1845
Set 3	−0.0006	0.0002	0.3524	0.2166	2874.5	2589.2	0.0471	−0.0982

M = mean; V = variance; S = skewness; K = excess kurtosis.

4. Data analysis

4.1. Choice of ε

This subsection chooses the value of ε to be used in later subsections by simulation. In order to find proper sets of parameters, we checked S&P500 daily log returns from January 2, 2001 to December 31, 2010. We first calculated $3\hat{K} - 5\hat{S}^2$ with 250 consecutive daily log returns from January 2, 2001 and then dropped the first day and added a new day at the end to calculate the second $3\hat{K} - 5\hat{S}^2$. We repeated this procedure to use all the data and obtained 2265 values of $3\hat{K} - 5\hat{S}^2$ (Figure 1). In the figure, negative values of $3\hat{K} - 5\hat{S}^2$ are marked by black dots and positive values are marked by red +'s. The negative portion is 21.8%, 495 out of 2265, in which cases there might be problems with NIG parameter estimation.

Among the considered sets of 250 S&P500 daily log return data above, we select three sets with $3\hat{K} - 5\hat{S}^2$ close to 0. Then we use (2.2) to calculate values of α, β, δ, and μ and assume these as true parameter values in simulation experiments. Table 1 shows the selected parameter values in three different datasets. For each set of parameter values, we estimate ε-MLE for α, β, δ, and μ using 1,000 simulated observations with ε changing from 0.005 to 0.12 by 0.005. This procedure is repeated 1,000 times and the RMSEs of ε-MLE for α, β, δ, and μ are computed (Figure 2).

In Figure 2, RMSE of α and RMSE of β decline dramatically first and then slowly increase at
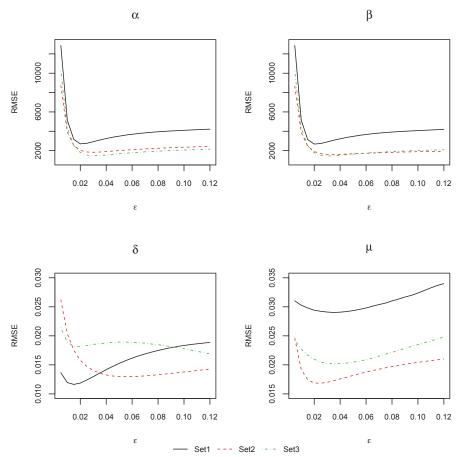
Figure 2: *RMSE of ε-MLEs by changing ε from 0.005 to 0.12 by 0.005 for three sets given in Table 1. RMSE = root mean square error; MLE = maximum likelihood estimator.*

around $\varepsilon = 0.02$. In case of δ and μ, RMSEs show different patterns depending on parameter sets but values between 0.02 and 0.03 seem to be appropriate. Thus, we decide to use 0.025 as the value of ε in numerical studies below.

4.2. Simulation

Here, we compare the performance of the ordinary MLE, the ordinary MME, ε-MLE, ε-MME, and exp-MLE by simulation. We consider two parameter settings and refer them as Scenario 1 and Scenario 2, respectively; in Scenario 1, $3K - 5S^2$ is 15.038 which is much larger than 0 and in Scenario 2, $3K - 5S^2$ is 0.038 which is close to 0. Table 2 provides the parameter values used in the two scenarios.

For each scenario, we generate 1,000 observations to calculate estimates and repeat this process 1,000 times. We consider two sets of initial values in order to also see the effect of the initial value in maximum likelihood estimation. In Case 1, the maximization procedure starts at $\alpha = 1$, $\beta = 0$, $\delta = 1$, and $\mu = 0$ and in Case 2, at ε-MME. MME is frequently used as the initial value for numerical

Table 2: Parameter values used in simulations: $3K - 5S^2 = 15.038$ in Scenario 1 and $3K - 5S^2 = 0.038$ in Scenario 2

	α	β	δ	μ	M	V	S	K
Scenario 1	77.361	0.259	0.0077	0.0002	0.0003	0.0001	0.013	5.013
Scenario 2	1539.2	102.21	0.1529	−0.0098	0.0003	0.0001	0.013	0.013

M = mean; V = variance; S = skewness; K = excess kurtosis.

Table 3: RMSE of estimators in Scenario 1: Initial values are $\alpha = 1, \beta = 0, \delta = 1$, and $\mu = 0$ in Case 1 and ε-MME in Case 2

		α	β	δ	μ
	MME	8.7150	0.5804	6.53×10^{-4}	5.02×10^{-5}
	MLE	**59.2279**	**13.3595**	2.75×10^{-3}	6.94×10^{-4}
Case 1	exp-MLE	0.2990	0.2136	5.81×10^{-5}	1.45×10^{-5}
	ε-MLE	8.6675	0.6431	4.72×10^{-4}	3.79×10^{-5}
	MLE	**1.9120**	**0.4499**	5.42×10^{-5}	3.81×10^{-5}
Case 2	exp-MLE	1.5984	0.0966	4.12×10^{-6}	6.48×10^{-6}
	ε-MLE	1.6592	0.1482	4.46×10^{-5}	7.63×10^{-6}

RMSE = root mean square error; MME = method of moments estimator; MLE = maximum likelihood estimator.

Table 4: Means and standard deviations of estimators in Scenario 1: initial values are $\alpha = 1, \beta = 0, \delta = 1$, and $\mu = 0$ in Case 1 and ε-MME in Case 2
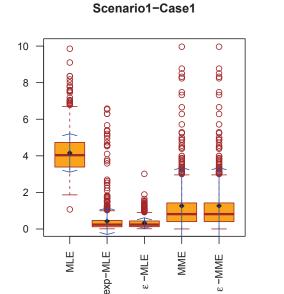
		α	β	δ	μ
True value		77.3610	0.2590	0.0077	0.0002
	MME	86.0768	−0.3211	0.0083	0.0003
		(17.9973)	(8.1433)	(0.0014)	(0.0007)
	MLE	18.1338	−13.1002	0.0049	0.0009
		(7.0908)	(5.4771)	(0.0002)	(0.0004)
Case 1	exp-MLE	77.6608	0.4729	0.0077	0.0002
		(14.6287)	(5.4163)	(0.0007)	(0.0004)
	ε-MLE	86.0293	−0.3837	0.0082	0.0003
		(9.0611)	(5.4339)	(0.0005)	(0.0004)
	MLE	79.2738	−0.1905	0.0077	0.0003
		(10.7438)	(6.1651)	(0.0006)	(0.0005)
Case 2	exp-MLE	78.9602	0.1626	0.0077	0.0002
		(10.6734)	(5.2191)	(0.0006)	(0.0003)
	ε-MLE	79.0210	0.1110	0.0077	0.0002
		(10.5841)	(5.1661)	(0.0006)	(0.0003)

Standard deviations are in parentheses. MME = method of moments estimator; MLE = maximum likelihood estimator.

procedures such as finding MLE; however, we use ε-MME instead because MME cannot be calculated in some situations. Tables 3 and 4 are RMSEs as well as the means and the standard deviations of the estimators in Scenario 1, respectively. In this scenario, MME and ε-MME are equal because $3\hat{K} - 5\hat{S}^2$ are greater than $\varepsilon = 0.025$ in all cases.

Tables 3 and 4 show the effect of initial values on estimation. In particular, the performance of the ordinary MLE improves noticeably from Case 1 to Case 2 as seen with bold-faced numbers in Table 3. ε-MLE is not affected by initial values as much as the ordinary MLE; however, the means of ε-MLE in Table 4 become closer to the true values for all parameters. Overall, exp-MLE shows the best performance regardless of initial values.

We now compare the difference between the true density and the fitted density as indicated in Figures 3 and 4. As in the previous simulation, we generate 1,000 observations each from scenarios 1
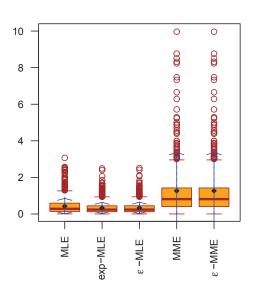
Figure 3: *Boxplots of averages of squared differences between the true density and the estimated density in Scenario 1: Initial values are $\alpha = 1, \beta = 0, \delta = 1,$ and $\mu = 0$ in Case 1 and ε-MME in Case 2. MLE = maximum likelihood estimator; MME = method of moments estimator.*

and 2 to estimate the parameters. For 1,000 equidistant points between −0.1 and 0.1 such as $x_1 = -0.1$, $x_2 = -0.0998, \ldots, x_{1000} = 0.0998$, the differences between the estimated density value and the true density value are computed. We then calculate the average of squared differences as

$$\sum_{i=1}^{1000} \frac{\left(f(x_i) - \hat{f}(x_i)\right)^2}{1000},$$

where $f(x)$ and $\hat{f}(x)$ are the true NIG density and the estimated NIG density, respectively. We obtain 1,000 averages plotted in Figures 3 and 4 by repeating this process 1,000 times. In Figure 3, boxplots of MME and ε-MME are equal because $3\hat{K} - 5\hat{S}^2$ is greater than $\varepsilon = 0.025$ in all paths. We can see that ε-MLE and exp-MLE are closer to the true density than other estimators and that the ordinary MLE performs better when we use ε-MME as the initial value.

Tables 5 and 6 are RMSEs and the means and the standard deviations of the estimators in Scenario 2, respectively. In this scenario, MME is calculated for only 544 out of 1,000 simulations since $3\hat{K} - 5\hat{S}^2$ is negative in other cases whereas ε-MME is estimated in all cases. RMSEs are much larger than those with Scenario 1, since $3K - 5S^2$ is close to 0 in this scenario. Note that the RMSEs in Table 5 and the mean and the standard deviation in Table 6 of the ordinary MLE are very large in Case 1. This indicates that the ordinary MLE can often provide extremely large values. The performance of ε-MME, exp-MLE and ε-MLE in Case 2 are similar; however, exp-MLE still shows the best result for all parameters.

Figure 4 provides the averages of squared differences between the true density and the estimated density in Scenario 2. The density estimated with exp-MLE is closer to the true density than those of ordinary MLE and ε-MLE in Case 1. When the initial value is set to be ε-MME in Case 2, all

Table 5: RMSE of estimators in Scenario 2: Initial values are $\alpha = 1$, $\beta = 0$, $\delta = 1$, and $\mu = 0$ in Case 1 and ε-MME in Case 2

		α	β	δ	μ
	MME	696.228	76.786	0.0843	0.0084
	ε-MME	76.186	41.656	0.0359	0.0003
Case 1	MLE	359389	280447.7	11.7650	14.7400
	exp-MLE	61.896	55.029	0.0065	0.0104
	ε-MLE	1025.828	61.173	0.1144	0.0078
Case 2	MLE	2248.810	61.465	0.1934	0.0013
	exp-MLE	45.489	41.594	0.0331	0.0004
	ε-MLE	94.484	48.328	0.0403	0.0003

RMSE = root mean square error; MME = method of moments estimator; MLE = maximum likelihood estimator.

Table 6: Means and standard deviations of estimators in Scenario 2: Initial values are $\alpha = 1$, $\beta = 0$, $\delta = 1$, and $\mu = 0$ in Case 1 and ε-MME in Case 2

		α	β	δ	μ
True value		1539.2	102.21	0.1529	−0.0098
	MME	843.008	25.428	0.0685	−0.0014
		(1231.938)	(935.797)	(0.0627)	(0.0403)
	ε-MME	1463.049	143.871	0.1169	−0.0095
		(777.206)	(682.236)	(0.1577)	(0.0464)
Case 1	MLE	360928.2	−280345	11.9188	14.7307
		(2017268)	(1560723)	(67.3032)	(82.4483)
	exp-MLE	1601.132	−2.814	0.1594	0.0005
		(743.039)	(682)	(0.1037)	(0.0311)
	ε-MLE	513.408	41.040	0.0384	−0.0020
		(817.254)	(728.617)	(0.0583)	(0.0474)
Case 2	MLE	3788.046	163.680	0.3464	−0.0112
		(14199.790)	(784.071)	(1.4012)	(0.0587)
	exp-MLE	1493.746	143.809	0.1197	−0.0094
		(589.716)	(11.660)	(0.0505)	(0.0265)
	ε-MLE	1444.751	150.543	0.1125	−0.0095
		(816.193)	(530.975)	(0.0645)	(0.0460)

Standard deviations are in parentheses. MME = method of moments estimator; MLE = maximum likelihood estimator.

estimators perform similarly. Note that ε-MME itself works well in this scenario and exp-MLE has some outliers in Case 1.

There are other parameter estimation methods such as the EM algorithm and MCMC method. Karlis (2002) dealt with EM algorithm for NIG distribution and Karlis and Lillestöl (2004) proposed estimation for the parameters of the NIG distribution through an MCMC scheme based on the Gibbs sampler. We also considered EM algorithm and MCMC method; however, we did not include them because the results of EM were not as good as all the estimators mentioned above and the MCMC method took so much time compared to other estimation methods.

4.3. Real data application

In a real data application, we estimate the parameters of NIG distribution for two data sets of a S&P500 daily log return and two data sets of insurance loss. The first S&P500 data set denoted as S&P500_I is collected from March 28, 2003 to April 24, 2004 and provides a negative value of $3\hat{K} - 5\hat{S}^2$. The second S&P500 data set is collected from March 07, 2006 to March 05, 2007 with $3\hat{K} - 5\hat{S}^2$ being positive and is denoted as S&P500_II. As for insurance loss, one is the U.S. indemnity loss data men-
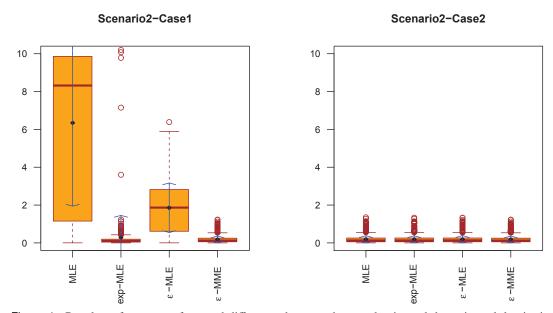
Figure 4: *Boxplots of averages of squared differences between the true density and the estimated density in Scenario 2: Initial values are $\alpha = 1, \beta = 0, \delta = 1$, and $\mu = 0$ in Case 1 and ε-MME in Case 2. MLE = maximum likelihood estimator; MME = method of moments estimator.*

Table 7: Descriptive statistics for S&P500_I, S&P500_II, U.S. indemnity loss, and Danish fire loss

	$\hat{M}$	$\hat{V}$	$\hat{S}$	$\hat{K}$	$3\hat{K} - 5\hat{S}^2$
S&P500_I	0.0009	0.00007	0.0345	−0.0256	−0.0831
S&P500_II	0.0002	0.00004	−0.4899	3.9521	10.6565
U.S. indemnity loss	8.5219	2.0446	−0.4920	1.0581	1.9637
Danish fire loss	0.7869	0.5136	1.7610	4.1790	−2.9701

M = mean; V = variance; S = skewness; K = excess kurtosis.

tioned in Frees and Valdez (1998) and the other is Danish fire loss data mentioned in McNeil (1997) and others. Both are widely used in actuarial literature and freely available through various sources such as R or MATLAB packages. U.S. indemnity loss data consists of 1,500 general liability claims each of which comprises the indemnity payment and the allocated loss adjustment expense. Danish fire loss data was collected by a Danish reinsurance company from January 03, 1980 to December 31, 1990. It contains 2,167 individual losses with a deductible of 1 million Danish Kroner. We take log of the observations in both insurance datasets before analysis, as a common practice in actuarial science (Bolance *et al.*, 2008). As in Section 4.2, initial values of $\alpha = 1, \beta = 0, \delta = 1, \mu = 0$, and ε-MME are also used here, and they are referred as Case 1 and Case 2, respectively.

Table 7 provides the descriptive statistics for four data sets used. $3\hat{K} - 5\hat{S}^2$ for S&P500_I and Danish fire loss data are negative so that ordinary MME cannot be obtained. However, fitting the NIG distribution to the S&P500_II and U.S. Indemnity loss data is done without difficulty because $3\hat{K} - 5\hat{S}^2$ is large enough. With these datasets, MME and ε-MME give exactly the same values.

Figure 5 compares a kernel density estimator with the estimated NIG density with ε-MME as the initial value whenever necessary. All of the estimators show similar results with S&P500 datasets. With U.S. indemnity loss data, the density with MME (ε-MME) is slightly off from the kernel density.

S&P500_I

S&P500_II

U.S. Indemnity loss

Danish fire loss

— Empirical --- MLE ······ MME ·-·- ε −MME -- ε −MLE ·-·- exp−MLE
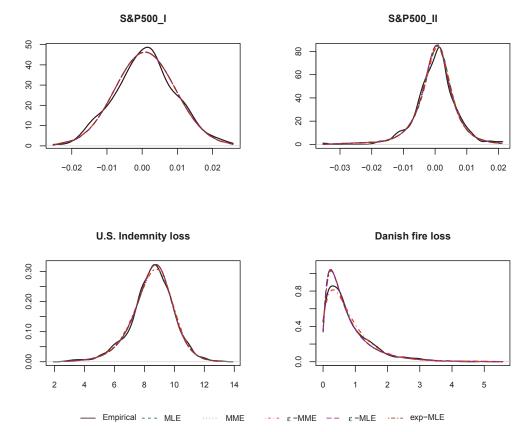
Figure 5: *Plots of empirical and estimated NIG densities with initial values of ε-MME. NIG = Normal inverse Gaussian; MLE = maximum likelihood estimator; MME = method of moments estimator.*

Table 8: *p*-values of the goodness-of-fit test for NIG distribution with S&P500_I, S&P500_II, U.S. indemnity loss, and Danish fire loss: Initial values are $\alpha = 1, \beta = 0, \delta = 1$, and $\mu = 0$

		S&P500_I		S&P500_II		U.S. indemnity loss		Danish fire loss	
		KS	AD	KS	AD	KS	AD	KS	AD
	MME	NA	NA	0.2993	0.8476	0.4602	0.5398	NA	NA
	ε-MME	0.9267	0.9756	0.2993	0.8476	0.4602	0.5398	0.0854	0.0592
Case 1	MLE	**0.0033**	**0.0051**	0.1362	**0.0380**	**0.0119**	**0.0214**	**0.0478**	**0.0127**
	exp-MLE	0.9321	0.9786	0.5584	0.9231	0.9145	0.8972	0.0649	0.0541
	ε-MLE	0.7506	0.6519	0.5386	0.9025	0.3907	0.3508	0.0817	0.0502

KS = Kolmogorov Smirnov-test; AD = Anderson-Darling test; NIG = normal inverse Gaussian; MME = method of moments estimator; MLE = maximum likelihood estimator.

With Danish fire loss data, the density with ε-MME seems to have a lower peak than other density estimates; in addition, the densities with ordinary MLE, exp-MLE, and ε-MLE have higher peak than the kernel density. Note that Danish fire loss data shows a peculiar pattern close to 0 since it has a deductible of 1 million Danish Kroner.

We run Kolmogorov-Smirnov (KS) test and Anderson-Darling (AD) test to examine if the NIG distribution is appropriate in describing the real datasets. Tables 8 and 9 show *p*-values of KS and

Table 9: *p*-values of the goodness-of-fit test for NIG distribution with S&P500_I, S&P500_II, U.S. indemnity loss, and Danish fire loss: Initial values are ε-MMEs

		S&P500_I		S&P500_II		U.S. indemnity loss		Danish fire loss	
		KS	AD	KS	AD	KS	AD	KS	AD
	MME	NA	NA	0.2993	0.8476	0.4602	0.5398	NA	NA
	ε-MME	0.9267	0.9756	0.2993	0.8476	0.4602	0.5398	0.0854	0.0592
Case 2	MLE	0.9309	0.9761	0.3323	0.8477	0.9088	0.9550	0.1251	0.0818
	exp-MLE	0.9366	0.9761	0.5631	0.9232	0.9143	0.9549	0.1887	0.1803
	ε-MLE	0.9313	0.9761	0.5487	0.9231	0.9142	0.9551	0.1682	0.1187

KS = Kolmogorov Smirnov-test; AD = Anderson-Darling test; NIG = normal inverse Gaussian; MME = method of moments estimator; MLE = maximum likelihood estimator.

AD tests with initial values of $\alpha = 1$, $\beta = 0$, $\delta = 1$, $\mu = 0$, and ε-MME, respectively. Ordinary MLE is significantly affected by the initial value and its performance is better when ε-MME is used as the initial value. In Table 8, the bold-faced *p*-values are smaller than 0.05 for the ordinary MLE. The results of the goodness-of-fit test in Table 9 improve when the initial values are set to be ε-MME. Although ε-MLE are relatively less affected by initial values than the ordinary MLE, the results mostly get improved from Case 1 to Case 2. Overall, exp-MLE shows the best goodness-of-fit in all datasets regardless of the initial value, but in Case 2, ε-MLE works similarly to exp-MLE. In summary, exp-MLE and ε-MLE with ε-MME as the initial value provide the best performance. We recommend to use ε-MME as the initial value when the ordinary MLE has to be used.

5. Conclusion

This study reviews adjusted parameter estimation methods of ε-MLE, ε-MME, and exp-MLE as well as compare them with the ordinary MLE and the ordinary MME for NIG distribution. The parameters of NIG distribution must satisfy $\alpha^2 - \beta^2 > 0$, $\alpha > 0$, and $\delta > 0$; therefore, $3K - 5S^2$ should be positive. However, $3\hat{K} - 5\hat{S}^2$ calculated from samples are not always positive and this can cause problems in estimating parameters; for example, MME cannot be obtained or maximization procedures do not converge as shown in Yoon and Song (2016). Yoon and Song (2016) suggested replacing $3\hat{K} - 5\hat{S}^2$ by $\max\{3\hat{K} - 5\hat{S}^2, \varepsilon\}$ for a fixed positive constant ε. A new parameterization that forces parameters to reside in the feasible domain of NIG distribution can also be used as in Kim (2019). We compared the performance of these estimation methods with different initial values by simulation and real data application.

From the simulation studies, the ordinary MLE was greatly influenced by the initial value and its performance improved significantly by setting the ε-MME as the initial value. However, the ordinary MLE still had large RMSE and standard deviation compared to other estimators. However, adjusted estimators were less affected by initial values than the ordinary MLE and closer to the true parameters. When ε-MME was used as the initial value, the performances of adjusted estimators were better than ordinary MLE with exp-MLE showing the best performance. In real data application, we applied adjusted and ordinary estimators to finance and insurance data sets; S&P500 daily log returns, U.S. indemnity loss data, and Danish fire loss data. Results were similar to what we have obtained in simulation experiments. Adjusted estimators and the ordinary MLE was affected by initial values. When initial values were set to be ε-MME, the results of goodness-of-fit tests improved, especially with ordinary MLE. However, adjusted estimators performed better than the ordinary MLE and particularly, exp-MLE and ε-MLE with the initial value of ε-MME showing good performance in real data application.

In the future, we would like to investigate other cases of generalized hyperbolic distribution in

terms of the feasible domain since we only deal with NIG distribution that is a special case of the generalized hyperbolic distribution. We will also study estimation methods such as EM algorithm and the MCMC method with generalized hyperbolic distribution.

References

Barndorff-Nielsen OE (1977). Exponentially decreasing distributions for the logarithm of particle size, *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **353**, 401–419.

Barndorff-Nielsen OE (1997). Processes of normal inverse Gaussian type, *Finance and Stochastics*, **2**, 41–68.

Behr A and Pötter U (2009). Alternatives to the normal model of stock returns: Gaussian mixture, generalised log *F* and generalised hyperbolic models, *Annals of Finance*, **5**, 49–68.

Bolance C, Guillen M, Pelican E, and Vernic R (2008). Skewed bivariate models and nonparametric estimation for the CTE risk measure, *Insurance: Mathematics and Economics*, **43**, 386–393.

Eberlein E and Keller U (1995). Hyperbolic distributions in finance, *Bernoulli*, **1**, 281–299.

Eriksson A, Ghysels E, and Wang F (2009). The normal inverse Gaussian distribution and the pricing of derivatives, *The Journal of Derivatives*, **16**, 23–37.

Fajardo J and Farias A (2004). Generalized hyperbolic distributions and Brazilian data, *Brazilian Review of Econometrics*, **24**, 249–271.

Figueroa-López JE, Lancette SR, Lee K, and Mi Y (2011). *Handbook of Modeling High-Frequency Data in Finance*, John Wiley & Sons, Hoboken, 3–26.

Frees E and Valdez E (1998). Understanding relationships using copulas, *North American Actuarial Journal*, **2**, 1–25.

Ghysels E and Wang F (2014). Moment-implied densities: properties and applications, *Journal of Business & Economic Statistics*, **32**, 88–111.

Hansen B (1994). Autoregressive conditional density estimation, *International Economic Review*, **35**, 705–730.

Karlis D (2002). An EM type algorithm for maximum likelihood estimation of the normal-inverse Gaussian distribution, *Statistics & Probability Letters*, **57**, 43–52.

Karlis D and Lillestöl J (2004). Bayesian estimation of NIG models via Markov chain Monte Carlo methods, *Applied Stochastic Models in Business and Industry*, **20**, 323–338.

Kim J (2019). *A study on the estimation of spliced distributions using exponential-estimation method* (Master's thesis), Korea University, Seoul.

Madan D and Seneta E (1990). The VG model for share market returns, *Journal of Financial Economics*, **63**, 511–524.

McNeil A (1997). Estimating the tails of loss severity distributions using extreme value theory, *ASTIN Bulletin*, **27**, 117–137.

Prause K (1999). *The generalized hyperbolic model: estimation, financial derivatives, and risk measures* (Ph.D. thesis), University of Freiburg, Germany.

Rydberg TH (1997). The normal inverse Gaussian Lévy process: simulation and approximation, *Communications in Statistics: Stochastic Models*, **13**, 887–910.

Yoon J and Song S (2016). A numerical study of adjusted parameter estimation in normal inverse Gaussian distribution, *The Korean Journal of Applied Statistics*, **29**, 741–752.