THE NORMAL INVERSE GAUSSIAN DISTRIBUTION: A VERSATILE MODEL FOR HEAVY-TAILED STOCHASTIC PROCESSES

Alfred Hanssen and Tor Arne Øigård

University of Tromsø, Physics Department
Electrical Engineering Group
N–9037 Tromsø, Norway
Email: {alfred,torarne}@phys.uit.no

ABSTRACT

The normal inverse Gaussian (NIG) distribution is a recent flexible closed form distribution that may be applied as a model of heavy-tailed processes. The NIG distribution is completely specified by four real valued parameters that have natural interpretations in terms of the shape of the resulting probability density function. By choosing the parameters appropriately, one can describe a wide range of shapes of the distribution. In this paper, we discuss several of the desirable properties of the NIG distribution. In particular, we discuss the cumulant generating function and the cumulants of the NIG-variables. A particularly important property is that the NIG distribution is closed under convolution. Finally, we derive a set of very simple yet accurate estimators of the NIG parameters. Our estimators differ fundamentally from estimators suggested by other authors in that our estimators take advantage of the surprisingly simple structure of the cumulant generating function.

1. INTRODUCTION

The so-called *normal inverse Gaussian (NIG) distribution* is a versatile non-Gaussian model recently introduced by Barndorff-Nielsen in 1995 [1, 2]. It was introduced as a possible model for financial data, and it has later also been attempted as a model of turbulence [2].

The NIG has to our knowledge never been applied in an electrical engineering context earlier. We will in this paper show that the NIG has several desirable and remarkable features which makes it suitable as a model of a large class of non-Gaussian noise processes. In particular, we propose that the NIG-model may be very useful for modeling impulsive noise. By examining the flexibility implied by the parameterization of the distribution, we believe that spiky noise present in radar, sonar and communication channels can be modeled by the NIG-distribution.

2. THE NORMAL INVERSE GAUSSIAN DISTRIBUTION

The normal inverse Gaussian distribution is a variance-mean mixture of a Gaussian distribution with an inverse Gaussian. A stochastic variable X is said to be normal inverse Gaussian if it has a probability density function of the form [1, 2, 5]

$$f_X(x) = \frac{\alpha\delta}{\pi}\frac{\exp[p(x)]}{q(x)}K_1\left[\alpha q(x)\right], \tag{1}$$

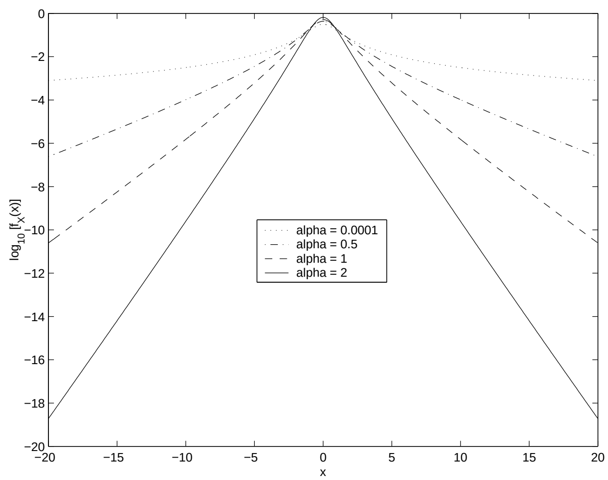

Fig. 1. NIG-density (logarithmic scale) for different values of α. Here, $\beta = \mu = 0$, and $\delta = 1$.

where $K_1(x)$ is the modified Bessel function of the second kind with index 1, $p(x) = \delta\sqrt{\alpha^2 - \beta^2} + \beta(x-\mu)$, $q(x) = ((x-\mu)^2 + \delta^2)^{1/2}$. Furthermore, $0 \le |\beta| < \alpha$, $\delta > 0$, and $-\infty < \mu < \infty$.

As seen from the definition in Eq. (1), the shape of the NIG-density is specified by a four dimensional parameter vector $(\alpha, \beta, \mu, \delta)$. This parameterization is very flexible indeed, making it possible to model a large variety of shapes and with various decay rates of the tail.

The four parameters of the NIG-distribution have natural interpretations relating to the overall shape of the density as follows. The α-parameter controls the steepness of the density, in the sense that the steepness or pointiness of the density increases monotonically with increasing α. This has implications also for the tail behavior, by the fact that large values of α implies light tails, while smaller values of α implies heavier tails. Note the similarity between this parameter and the α-parameter in the α-stable distribution [4]. Fig. 1 shows the dependency on α for $\beta = \mu = 0$ and $\delta = 1$. Note that the tails become heavier as the value of α decreases.

The β-parameter is a skewness parameter, in the sense that $\beta < 0$ implies a density skew to the left, $\beta > 0$ implies a density skew to the right, and $\beta = 0$ implies a density that is symmetric

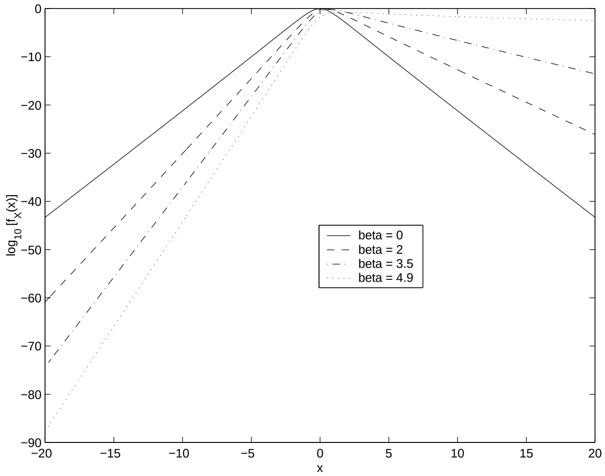

Fig. 2. NIG-density (logarithmic scale) for different values of β. Here, $\alpha = 5$, $\mu = 0$, and $\delta = 1$.

around μ, which is obviously a centrality or translation parameter. Fig. 2 shows the dependency on β. Note that the skewness increases as β increases.

Last, the δ-parameter is a scale parameter in the sense that the rescaled parameters $\alpha \to \alpha\delta$ and $\beta \to \beta\delta$ are invariant under location-scale changes of x.

3. PROPERTIES OF NIG-VARIABLES

NIG-variables obey several desirable properties that make them suitable for practical noise modeling. We will now demonstrate the attractiveness of the NIG-distribution in terms of some of its properties.

3.1. Cumulants

Barndorff-Nielsen [2] derived the moment generating function of the NIG-distribution. By generalizing his result, we readily derive the characteristic function of the NIG as

$$\Phi_X(\omega) = e^{\delta\sqrt{\alpha^2-\beta^2}} e^{-\delta\sqrt{\alpha^2-(\beta+j\omega)^2}} e^{j\mu\omega}, \qquad (2)$$

where $j = \sqrt{-1}$ is the imaginary unit, and $-\infty < \omega < \infty$.

The cumulant generating function $\Psi_X(\omega) = \ln \Phi_X(\omega)$ therefore has the following simple form

$$\Psi_X(\omega) = \delta\left[\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+j\omega)^2}\right] + j\mu\omega \qquad (3)$$

With the cumulant generating function at hand, it is now straightforward to calculate the cumulant of order n by means of

$$\kappa_X^{(n)} = (-j)^n \frac{d^n \Psi_X(\omega=0)}{d\omega^n}. \qquad (4)$$

The first four cumulants are readily found to be

$$\kappa_X^{(1)} = \mu + \frac{\beta\delta}{\sqrt{\alpha^2-\beta^2}} \qquad \kappa_X^{(2)} = \frac{\alpha^2\delta}{(\alpha^2-\beta^2)^{3/2}} \qquad (5)$$

$$\kappa_X^{(3)} = \frac{3\alpha^2\beta\delta}{(\alpha^2-\beta^2)^{5/2}} \qquad \kappa_X^{(4)} = \frac{3\alpha^2(\alpha^2+4\beta^2)\delta}{(\alpha^2-\beta^2)^{7/2}}. \qquad (6)$$

We thus see that all cumulants exist and that they are expressible as simple algebraic functions of the parameters.

It is particularly interesting to notice that the skewness and kurtosis has the following elegant closed form expressions

$$\gamma_3 = \frac{\kappa_X^{(3)}}{\left[\kappa_X^{(2)}\right]^{3/2}} = \frac{3\rho}{\sqrt{\xi}}, \qquad (7)$$

and

$$\gamma_4 = \frac{\kappa_X^{(4)}}{\left[\kappa_X^{(2)}\right]^{2}} = \frac{3\left(1+4\rho^2\right)}{\xi}, \qquad (8)$$

where $\rho = \beta/\alpha$ and $\xi = \delta\sqrt{\alpha^2-\beta^2}$.

Since $-1 < \rho < 1$ and $0 < \xi < \infty$, the expressions for skewness and kurtosis, Eqs. (7) and (8) show that we can model data that covers a very large range of non-Gaussian shapes. By combining Eqs. (7) and (8) it is easy to show that we may model variables with any simultaneous skewness and kurtosis in the region $\gamma_4 \geq 4\gamma_3^2/3$.

3.2. Exact limits

If we assume $\beta = 0$ and μ arbitrary, one can readily show that the Gaussian $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ density is a limit when either $\alpha \to \infty$ or $\delta \to \infty$, with the identification that $\sigma^2 = \delta/\alpha$.

Another important special case of the NIG is the Cauchy distribution which results when $\alpha = \beta = 0$, and μ and δ arbitrary.

3.3. Tail behavior

Asymptotically, the Bessel function behaves as

$$K_1(x) \sim \sqrt{\frac{\pi}{2x}} \exp(-x) \quad ; \quad |x| \to \infty. \qquad (9)$$

Hence, the tail of the NIG decays as

$$f_X(x) \sim |x|^{-3/2} \exp\left(\beta x - \alpha|x|\right). \qquad (10)$$

Note that Eq. (10) is invalid when $\alpha - |\beta| \ll 1$. In that special case, the tail of the NIG decays as

$$f_X(x) \sim |x|^{-2}, \qquad (11)$$

which is of course the tail behavior of the Cauchy.

3.4. Convolution property

A very attractive and useful property of the NIG that cannot be overrated, is that it is closed under convolution [1, 2]. This has far reaching consequences when considering sums of NIG variables.

Let $X_1, \ldots, X_M$ be M independent NIG-variables with common parameters α and β, but having individual location parameters $\mu_1, \ldots, \mu_M$, and individual scale parameters $\alpha_1, \ldots, \alpha_M$. Then the sum variable $Y = X_1 + \cdots + X_M$ is also NIG distributed, with parameters $(\alpha, \beta, \mu_{tot}, \delta_{tot})$, where $\mu_{tot} = \sum_{m=1}^{M} \mu_m$ and $\alpha_{tot} = \sum_{m=1}^{M} \alpha_m$.

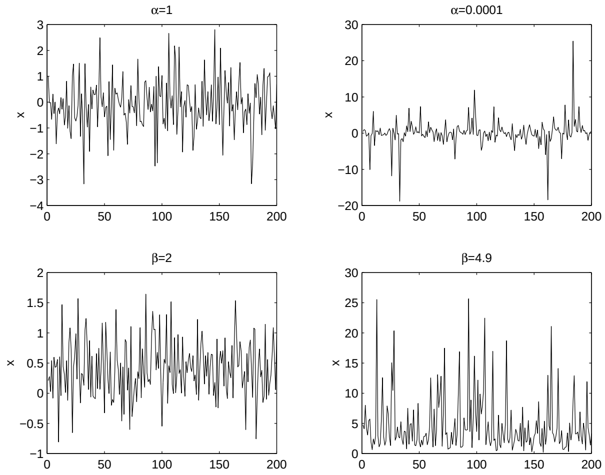

Fig. 3. Synthetic NIG data. Upper left $\alpha = 1$ and upper right $\alpha = 0.0001$, for $\beta = \mu = 0$ and $\delta = 1$. Lower left $\beta = 2$ and lower right $\beta = 4.9$, for $\alpha = 5, \mu = 0$, and $\delta = 1$.

4. GENERATION OF SYNTHETIC NIG DATA

We have generated synthetic statistically independent NIG-data by means of the well known *acceptance-rejection method* (see e.g., [3]). This method is based on drawing samples from a simpler auxiliary density $\tilde{f}_{\tilde{X}}(\tilde{x})$ from which variables can easily be made. In addition, uniform variables $y \sim \mathcal{U}[0,1]$ are generated and compared to a density that completely blankets the wanted density $f_X(x)$.

In Fig. 3 we show sequences consisting of 200 white NIG-samples in each. The upper panels have $\beta = \mu = 0$ and $\delta = 1$ as common parameters, with $\alpha = 1$ in the upper left corner, and $\alpha = 0.0001$ in the upper right corner. The lower panels have $\alpha = 5, \mu = 0$ and $\delta = 1$ as common parameters, with $\beta = 2$ in the lower left corner, and $\beta = 4.9$ in the lower right corner. Note that the $\alpha = 1$ data are symmetric and without extreme excursions from the mean, while the $\alpha = 0.0001$ data have an impulsive character suggesting heavy tails. Furthermore, note that the $\beta = 2$ data are slightly skewed, while the $\beta = 5$ data are strongly skewed towards positive values.

The theoretical probability density functions for the synthetic data shown in Fig. 3 can be found in Figs. 1 and 2.

5. ESTIMATION OF THE NIG PARAMETERS

Moment estimators exist [5], but they should be avoided due to their poor statistical behavior. Maximum likelihood estimators (MLE) have also been suggested and implemented [5]. These estimators are very complicated to handle numerically, due to the highly nonlinear form of the NIG-distribution. We recommend to avoid also the MLE due to their numerical intractability.

We have chosen to employ a totally different approach when developing estimators of the NIG-variables. The structure of the cumulant generating function (CGF), eq. (3), is very simple in terms of the four parameters that defines the NIG. This structure should obviously be exploited. We will now show that it is easy to establish a set of simple relations for the CGF that can be directly exploited when constructing accurate estimators for the parameters.

First, we expand the complex valued square root in eq. (3) explicitly, so that we can write the real and imaginary parts of $\Psi_X(\omega)$ as

$$\mathrm{Re}\Psi_X(\omega) = \delta\left[d - \frac{1}{\sqrt{2}}\sqrt{s^2 + \sqrt{s^4 + 4\beta^2\omega^2}}\right] \tag{12}$$

$$\mathrm{Im}\Psi_X(\omega) = \left[\mu + \frac{\sqrt{2}\beta\delta}{\sqrt{s^2 + \sqrt{s^4 + 4\beta^2\omega^2}}}\right]\omega, \tag{13}$$

where $d^2 = \alpha^2 - \beta^2$, and $s^2 = d^2 + \omega^2$.

We understand that the above exact expressions for $\mathrm{Re}\Psi_X(\omega)$ and $\mathrm{Im}\Psi_X(\omega)$ can be used to derive a large number of different estimators of the four NIG-parameters. It is thus possible to talk about CGF-based parameter estimators. We will now derive a set of simple and accurate CGF-based estimators of the four NIG-parameters.

5.1. Estimating δ

We first confine our attention to values of ω such that

$$\omega^2 \gg \alpha^2 - \beta^2 \quad \text{and} \quad \omega^2 \gg 4\beta^2. \tag{14}$$

Under these assumptions, we find that

$$\mathrm{Re}\Psi_X(\omega) \simeq \delta\sqrt{\alpha^2 - \beta^2} - \delta\omega. \tag{15}$$

Thus, δ is given by the negative slope of $\mathrm{Re}\Psi_X(\omega)$ for large ω, and an estimator is constructed by estimating the slope of $\mathrm{Re}\widehat{\Psi}_X(\omega)$, where $\widehat{\Psi}_X(\omega)$ is an estimate of the cumulant generating function. Formally, we may write

$$\widehat{\delta} = -\frac{d}{d\omega}\mathrm{Re}\widehat{\Psi}_X(\omega). \tag{16}$$

5.2. Estimating μ

The assumptions in Eq. (14) can now be applied to the imaginary part of the CGF, Eq. (13), which yields

$$\mathrm{Im}\Psi_X(\omega) \simeq \mu\omega. \tag{17}$$

Thus, μ is given by the slope of $\mathrm{Im}\Psi_X(\omega)$ for large ω, and an estimator is constructed by estimating the slope of $\mathrm{Im}\widehat{\Psi}_X(\omega)$. Formally, we may write

$$\widehat{\mu} = \frac{d}{d\omega}\mathrm{Im}\widehat{\Psi}_X(\omega). \tag{18}$$

Note the strong degree of mathematical symmetry between the estimators of δ and μ.

5.3. Estimating β

Combining Eqs. (12) and (13) we find that $\mathrm{Im}\Psi_X(\omega)$ can be written as

$$\mathrm{Im}\Psi_X(\omega) = \frac{\beta\omega}{\sqrt{\alpha^2 - \beta^2} - \mathrm{Re}\Psi_X(\omega)/\delta} + \mu\omega \tag{19}$$

It follows directly from (14) that

$$\left| \frac{\mathrm{Re}\Psi_X(\omega)}{\delta} \right| \gg \sqrt{\alpha^2 - \beta^2}. \tag{20}$$

Thus, we may approximate Eq. (19) by

$$\mathrm{Im}\Psi_X(\omega) \simeq \frac{\beta\omega\delta}{\mathrm{Re}\Psi_X(\omega)} + \mu\omega \tag{21}$$

It is now easy to estimate β by the use of Eq. (21). The resulting estimator reads

$$\widehat{\beta} = \frac{\mathrm{Re}\widehat{\Psi}_X(\omega)}{\omega\widehat{\delta}} \left[\widehat{\mu}\omega - \mathrm{Im}\widehat{\Psi}_X(\omega) \right]. \tag{22}$$

5.4. Estimating α

The remaining parameter, α can now readily be estimated by means of Eq. (12), to give

$$\widehat{\alpha} \simeq \left\{ \left[\frac{\mathrm{Re}\widehat{\Psi}_X(\omega)}{\widehat{\delta}} + \frac{1}{\sqrt{2}} \sqrt{\omega^2 + \sqrt{\omega^4 + 4\widehat{\beta}^2\omega^2}} \right]^2 + \widehat{\beta}^2 \right\}^{1/2}. \tag{23}$$

Again, the assumptions in (14) have been applied.

5.5. Estimating the cumulant generating function

To obtain a statistically consistent estimator of the CGF, we note that the probability density function (PDF) itself $f_X(x)$ and the characteristic function $\Phi_X(\omega)$ constitute a Fourier transform pair. We have therefore applied the following indirect method to estimate the CGF: First, we form a standard non-parametric kernel smoothed estimate of the PDF (see e.g., Silverman's book [6]), to obtain $\widehat{f}_X(x)$. The use of a Gaussian kernel function has proven to yield good results for synthetic NIG-data. Secondly, the characteristic function is estimated by means of a discrete Fourier transform of the kernel estimate $\widehat{f}_X(x)$. In practice, we perform the discrete Fourier transform by means of the Fast Fourier Transform (FFT), to yield $\widehat{\Phi}_X(\omega)$ at a set of discrete transform variables ω_i, where i is a discrete index. Finally, the estimate of the cumulant generating function is found from $\widehat{\Psi}_X(\omega) = \ln \widehat{\Phi}_X(\omega)$. Note that a phase unwrapping may be necessary to correct the phase discontinuities caused by the numerical evaluation of the complex logarithm.

5.6. Performance evaluation

We carried out a Monte Carlo simulation to assess the accuracy of the proposed estimator. The simulations are based on 1000 repetitions, and in Table 1 we show the Monte Carlo bias and standard deviation of the estimator of the scale parameter δ. The results are shown for various lengths N of the available data segments, and for two different values of δ. A relatively short range of ω-values was used for obtaining the estimate of the characteristic function: $1.4 \leq \omega \leq 2.4$, and the kernel function was chosen to be a standardized Gaussian.

From Table 1 we see that as expected, the accuracy becomes better as N increases. Both the bias and the standard deviation converges to zero as the sample size increases, so the proposed estimator is consistent. The consistency of the estimator can be shown analytically. From the table, we see that the estimation error increases with increasing values of δ. This is also confirmed by the Cramér-Rao lower bounds for the parameters [7].

N	$\delta = 1$	$\delta = 2$
100	bias: 0.5249 std: 0.8736	bias: 0.5554 std: 0.1735
1000	bias: 0.0709 std: 0.1809	bias: 0.3020 std: 0.8703
10000	bias: 0.0106 std: 0.0702	bias: 0.0196 std; 0.1366
50000	bias: 0.0021 std: 0.0352	bias: 0.0025 std: 0.0332

Table 1. *Monte Carlo simulated bias and standard deviation for $\widehat{\delta}$ for different N. Here, $\alpha = 1$ and $\beta = \mu = 0$.*

6. CONCLUSIONS

In this paper, we reviewed the normal inverse Gaussian (NIG) distribution, and discussed several of its properties. We demonstrated that the parameterization of the NIG-distribution allows for a very flexible formulation of distributions, ranging from the Gaussian to the Cauchy distribution. We also introduced a simple and accurate CGF-based estimator of the NIG-parameters.

Our conclusion is that the NIG-distribution is a versatile model for (leptokurtic) non-Gaussian variables with heavy tails (i.e., no heavier than the Cauchy, which is arguably sufficient for treating most real world impulsive noise processes). Among the many advantages of the NIG-distribution, we recall that (i) the density is given in closed form, (ii) NIG-variables are closed under convolution, (iii) the cumulant generating function is very simple, (iv) the parameterization of the model allows for a large range of probability density shapes, and (v) we have derived fast and accurate parameter estimators that take advantage of the cumulant generating function.

We believe that the NIG distribution will become useful for the modeling of impulsive noise in sonar, radar, and communication channels. We have carried out an analysis of impulsive noise in the ocean acoustic channel, which shows that the NIG outperforms the α-stable model for such noise.

7. REFERENCES

[1] O. E. Barndorff-Nielsen, "Normal inverse Gaussian processes and the modeling of stock returns", *Rep. No. 300*, Dept. of Theor. Stat., Inst. of Math., Univ. of Aarhus, Denmark, 1995.

[2] O. E. Barndorff-Nielsen, "Normal inverse Gaussian distributions and stochastic volatility modelling", Scand. J. Statist., Vol. 24, pp. 1–13, 1997.

[3] D. Gamerman, *Markov Chain Monte Carlo*, Chapman & Hall, 1997.

[4] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*, Wiley, 1995.

[5] T. H. Rydberg, "The normal inverse Gaussian Lévy process: Simulation and approximation", *Commun. Statist.–Stochastic Models*, Vol. 34, pp. 887 – 910, 1997.

[6] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, 1986.

[7] T. A. Øigård, *The normal inverse Gaussian and the stable distributions as models of heavy tailed data*, M.Eng.-Thesis, Physics Dept., Univ. of Tromsø, Tromsø, Norway, 2001.