Approximating the probability distribution of functions of random variables: A new approach*

Anders Eriksson Lars Forsberg Eric Ghysels

First draft: March 2003
This version: March 2, 2004

Abstract

We introduce a new approximation method for the distribution of functions of random variables that are real-valued. The approximation involves moment matching and exploits properties of the class of normal inverse Gaussian distributions. In the paper we we examine the how well the different approximation methods can capture the tail behavior of a function of random variables relative each other. This is obtain done by simulate a number functions of random variables and then investigate the tail behavior for each method. Further we also focus on the regions of unimodality and positive definiteness of the different approximation methods. We show that the new method provides equal or better approximations than Gram-Charlier and Edgeworth expansions.

*Anders Eriksson (corresponding author), Department of Information Science - Division of Statistics, University of Uppsala. email:anders.eriksson@dis.uu.se. Lars Forsberg, Department of Information Science - Division of Statistics, University of Uppsala. email:lars.forsberg@dis.uu.se. Eric Ghysels, Department of Economics, University of North Carolina at Chapel Hill, email:eghysels@email.unc.edu. The authors thank *The Jan Wallander and Tom Hedelius Research Foundation* and *The Swedish Foundation for International Cooperation in Research and Higher Education, (STINT)* for financial support.

1 Introduction

Many statistical models involve functional transformations of random variables. Regression models with stochastic regressors are the most common example, involving a linear transformation of random variables. Likewise, mixture models involve multiplicative transformations. In the linear model with Gaussian regressors and errors the dependent variable is also Gaussian. However, in general, when regressors and/or errors are non-Gaussian we do not know the distribution of the dependent variable. For mixture models we do not even know the distribution in the Gaussian case. In many circumstances, one is interested in the distribution of the dependent variable. In this paper we provide methods to approximate linear and multiplicative transformations of independent random variables. The results are driven by adopting a flexible class of probability laws that allows us to approximate the density of interest. Historically there have been at least three different ways of approximating an algebraic function of random variables. They are (1) the Pearson family, (2) Gram-Charlier and Edgeworth expansions and (3) the method of transformations.

Pearson (1895) established a family of frequency curves to represent empirical distributions. The so called Pearson family of distributions has proven to be useful in approximating a theoretical distribution via moment matching. However, this feature is mostly valid for the Pearson type I and type III density (known as the Beta and Gamma densities respectively). The most significant shortcoming of the Pearson type and I and type III densities is the limitation to represent densities only via two parameters. This implies that one only matches two moments.

The Gram-Charlier expansion (Charlier (1905)) and the Edgeworth expansion (Edgeworth (1896), Edgeworth (1907)) were established in the beginning of the 20th century. Both have been the most successful, and notably been linked to the bootstrap (see for example Hall (1995)). The approximation methods build on the expansion of the Gaussian density function in terms of Hermite polynomials. However, a potential drawback of such expansions is that (1) they do not always result in unimodal approximations and (2) more seriously, they do not always imply positive definiteness of the density (see Barton and Dennis (1952) and Draper and Tierny (1972)).

The main building block of the method of transformation to achieve a flexible distribution is the use of a monotonic transform to a known and well behaved distribution. The transformed random variable has a distribution that matches the characteristics of the data, such as skewness, excess kurtosis etc. This method has its drawbacks too. Johnson (1949) provided examples of classes of densities for real-valued random variables where the moment structure is too complicated to make moment matching feasible.

Following the tradition of adopting flexible functional forms for densities combined with moment matching we exploit the class of normal inverse Gaussian densities (Barndorff-Nielsen (1978)) to provide approximations to functional transformations of real-valued independent random variables. The family of normal inverse Gaussian (henceforth NIG) densities is a special case of the generalized hyperbolic distribution(GH), which is defined as a Gaussian-generalized inverse Gaussian mixing distribution. The family of NIG densities has many interesting features that are of interest for applications in areas such as turbulence and finance, among others (see Barndorff-Nielsen (1997)). Under certain regularity conditions, the class is closed under convolution, and the structure of the cumulants is particularly appealing for the purpose of moment matching.

The versatility of the class of NIG densities allows us to revisit the approximation of unknown densities via moment matching. Although we focus primarily on linear and multiplicative transformations, it should be noted that the approach proposed in this paper applies to nonlinear transformations as well. Our approximations are shown to improve upon Gram-Charlier and Edgeworth expansions for various skewed and fat-tailed distributions. The class of NIG distributions used in our approximations is a four parameter family that allows for mean, variance, skewness and kurtosis matching while maintaining the unimodal character of a distribution. For the purpose of distribution approximations, there are two main advantages to the NIG class, namely: (1) the general flexibility of the distribution and (2) the property that the parameters can be explicitly solved for in terms of the cumulants of the distribution. The latter property is appealing as it facilitates moment matching with the first four moments of an approximate NIG density.

The remainder of the paper is organized as follows. In section 2 we provide a brief discussion of the NIG class of distributions and the resulting approximation method. In section 3 we compare the NIG approximation with Edgeworth and Gram-Charlier expansions. The comparison focuses on the tail behavior for a random coeffcient model under different distributional assumptions appears in section 4. Section 5 concludes the paper.

2 Approximations and the class of normal inverse Gaussian distributions

The purpose of this section is to present the main results of the paper. In a first subsection we briefly review the NIG class of densities, and in a second subsection we present the main results regarding the approximation principle using the NIG class.

2.1 A brief review of NIG distributions

The normal inverse Gaussian distribution is characterized via a normal inverse Gaussian mixing distribution. Formally stated, let Y be a random variable that follows an inverse Gaussian law (IG) (see Sheshardi (1993)):

$$\mathcal{L}(Y) = IG\left(\delta, \sqrt{\alpha^2 - \beta^2}\right)$$

Furthermore, if X conditional on Y is normally distributed with mean $\mu + \beta Y$ and variance Y, namely: $\mathcal{L}(X|Y) = N(\mu + \beta Y, Y)$, then the unconditional density X is normal inverse Gaussian:

$$\mathcal{L}(X) = NIG(\alpha, \beta, \mu, \delta).$$

The density function for the NIG family is defined as follows:

$$f_{NIG}(x; \alpha, \beta, \mu, \delta) = \frac{\alpha}{\pi\delta} \exp\left(\delta\sqrt{\alpha^2 - \beta^2} - \beta\mu\right) \frac{K_1\left(\alpha\delta\sqrt{1 + \left(\frac{x-\mu}{\delta}\right)^2}\right)}{\sqrt{1 + \left(\frac{x-\mu}{\delta}\right)^2}} \exp(\beta x) \tag{2.1}$$

where $x \in \mathbb{R}$, $\alpha > 0$ $\delta > 0$, $\mu \in \mathbb{R}$, $0 < |\beta| < \alpha$, and $K_1(.)$ is the modified Bessel function of the third kind with index 1 (see Abramowitz and Stegun (1972)). The Gaussian distribution is obtained as a limiting case, namely when $\alpha \to \infty$. Moreover, the Fourier transform for the NIG density is given by:

$$\varphi_X(t) = \exp\left(\delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\left(\alpha^2 - (\beta + t)^2\right)}\right) + t\mu\right). \tag{2.2}$$

The NIG class of densities has the following two properties,namely (1) a scaling property:

$$\mathcal{L}_{NIG}(X) = NIG(\alpha, \beta, \mu, \delta) \Leftrightarrow \mathcal{L}_{NIG}(cX) = NIG(\alpha/c, \beta/c, c\mu, c\delta), \tag{2.3}$$

and (2) a closure under convolution property:

$$NIG(\alpha, \beta, \mu_1, \delta) * NIG(\alpha, \beta, \mu_2, \omega) = NIG(\alpha, \beta, \mu_1 + \mu_2, \delta + \omega). \tag{2.4}$$

A more convenient parameterization used throughout this paper is obtained by setting $\bar{\alpha} = \delta\alpha$ and $\bar{\beta} = \delta\beta$. This representation is a scale-invariant parameterization denoted $\overline{NIG}\left(\bar{\alpha}, \bar{\beta}, \mu, \delta\right)$, with density:

$$f_{\overline{NIG}}\left(x; \bar{\alpha}, \bar{\beta}, \mu, \delta\right) = \frac{\bar{\alpha}}{\pi\delta} \exp\left(\sqrt{\bar{\alpha}^2 - \bar{\beta}^2} - \frac{\bar{\beta}\mu}{\delta}\right) \frac{K_1\left(\bar{\alpha}\sqrt{1 + \left(\frac{x-\mu}{\delta}\right)^2}\right)}{\sqrt{1 + \left(\frac{x-\mu}{\delta}\right)^2}} \exp\left(\frac{\bar{\beta}}{\delta}x\right) \tag{2.5}$$

and the Fourier transform for the scale-invariant parameterization of the NIG-law is given by

$$\varphi_X(t) = \exp\left(\left(\sqrt{\bar{\alpha}^2 - \bar{\beta}^2} - \sqrt{\left(\bar{\alpha}^2 - \left(\bar{\beta} + \delta^2 t\right)^2\right)}\right) + t\mu\right). \tag{2.6}$$

A common reparametrization is $\bar{\varkappa} = \bar{\beta}/\bar{\alpha}$ this simplifies the expression for the cumulants throughout the paper we will use this kind of parametrization when dealing with cumulants.

2.2 Approximations using the NIG class of densities

The principle of approximation applied to the NIG class consists of constructing a non-linear system of equations for the four parameters in the NIG distribution. In particular, one sets the first and second cumulant, the skewness and the kurtosis equal to the same measures associated with the functional transformation. We present the approximation first and defer the discussion of the regularity conditions until later. It is worth noting at this stage, however, that one must assume that the relevant moments of the transformed random variable exist. Moreover, it is also assumed that one knows the first four cumulants of the function one wishes to approximate, a standard requirement in approximation theory. One of the main advantages of the NIG class, when solving the non-linear system of equations to match moments, is that one obtains explicit functions for each parameter in terms of the cumulants of the distribution to approximate.

More specifically, consider $Y = f(X_1, ..., X_n)$ where X_i are random variables and assume the expression for the first four cumulants for Y is known. Furthermore, assume that we can approximate the distribution Y with some distribution X_*

$$\mathcal{L}(X_*) = \overline{NIG}\left(\bar{\alpha}_*, \bar{\beta}_*, \mu_*, \delta_*\right),$$

with the expected value, variance skewness and kurtosis:

$$E[X_*] = \mu_* + \frac{\bar{\varkappa}_* \delta_*}{(1-\bar{\varkappa}_*^2)^{1/2}} \tag{2.7}$$

$$V[X_*] = \frac{\delta_*^2}{\bar{\alpha}_*(1-\bar{\varkappa}_*^2)^{3/2}} \tag{2.8}$$

$$S[X_*] = \frac{3\bar{\varkappa}_*}{\bar{\alpha}_*^{1/2}(1-\bar{\varkappa}_*^2)^{1/4}} \tag{2.9}$$

$$K[X_*] = 3\frac{4\bar{\varkappa}_*^2+1}{\bar{\alpha}_*(1-\bar{\varkappa}_*^2)^{1/2}}. \tag{2.10}$$

where $\bar{\varkappa}_* = \bar{\beta}_*/\bar{\alpha}_*$. In order to approximate the distribution Y we must solve for the different parameters in X_*. Therefore, let the first four cumulants for the distribution Y, denoted as $\kappa_1^Y, \kappa_2^Y \kappa_3^Y$ and κ_4^Y. We need to solve a non-linear systems of equations, a system that has an explicit solution, as shown in Appendix B.

Before we state the theoretical result, we need to discuss the regularity conditions. The first two assumptions are related to the fact that we are approximating with a unimodal distribution with the information set restricted to only four cumulants.

Assumption 2.1 *The function of random variables that you approximate should be distributed on* $\mathbb{R}$, $f(X) \in \mathbb{R}$.

Assumption 2.2 *The cumulants of* $f(X)$ *are assumed to exist up to order 4 and are known or have been estimated.*

Finally, following relation for the cumulants must be fulfilled in order to for the approximation to work properly:

Assumption 2.3 *Let* $\rho = \left(3\kappa_4^Y\left(\kappa_2^Y\right)/\left(\kappa_3^Y\right)^2 - 4\right)$. *It is assumed that* $\rho > 0$ *and* $(1-\rho^{-1}) > 0 \Leftrightarrow \rho^{-1} < 1$.

The following Lemma clarifies the restrictions imposed by Assumption 2.3:

Lemma 2.1 *Let Assumption 2.3 hold, then* $3\kappa_4^Y \kappa_2^Y / \left(\kappa_3^Y\right)^2 = 3\left(K_F S_F^{-2}\right) > 5$, *where* $K_F = \kappa_4^Y \; / \left(\kappa_2^Y\right)^2$ $S_F = \kappa_3^Y \; / \left(\kappa_2^Y\right)^{\frac{3}{2}}$ *and is the Fisherian shape coefficient of excess kurtosis and* S_F *is the Fisherian coefficient of skewness:*

Proof: See Appendix A

Given the above assumptions, the following theorem yields the parameters in the approximation distribution as functions of the cumulants of the distribution Y :

Theorem 2.1 (NIG approximation) *Let Assumptions 2.1 through 2.3 hold. Given the first four cumulants of the unknown distribution* Y *we can express the parameters generating a NIG probability distribution with the same four cumulants as* Y :

$$\overline{\alpha}_* = 3\left(4/\rho + 1\right)\left(1 - \rho^{-1}\right)^{-1/2}\left(\left(\kappa_2^Y\right)^2 / \kappa_4^Y\right) \tag{2.11}$$

$$\overline{\beta}_* = \text{signum}\left(\kappa_3^Y\right) / \sqrt{\rho} 3\left(4/\rho + 1\right)\left(1 - \rho^{-1}\right)^{-1/2}\left(\left(\kappa_2^Y\right)^2 / \kappa_4^Y\right) \tag{2.12}$$

$$\mu_* = \kappa_1^Y - \text{signum}\left(\kappa_3^Y\right) / \sqrt{\rho}\left(\left(12/\rho + 3\right)\left(\kappa_2^Y\right)^3 / \kappa_4^Y\right)^{1/2} \tag{2.13}$$

$$\delta_* = \left(3\left(\kappa_2^Y\right)^3\left(4/\rho + 1\right)\left(1 - \rho^{-1}\right) / \kappa_4^Y\right)^{1/2} \tag{2.14}$$

where $\rho = \left(3\kappa_4^Y \left(\kappa_2^Y\right) / \left(\kappa_3^Y\right)^2 - 4\right)$

Proof: See Appendix B

To conclude this section we provide an illustrative example. We do not discuss the accuracy of this approximation, see however section 3 for a simulation study regarding this approximation. The example only serves the purpose of illustrating the mechanism of the method. In particular, consider the following function of student t random variables.

$$Y = \gamma_1 X_1 + \gamma_2 X_2 \text{ where } \mathcal{L}(X_i) = t(\upsilon_i) \text{ i=1,2} \tag{2.15}$$

We know that the second and fourth cumulant equals

$$\kappa_2^Y = \gamma_1^2 v_1 / (v_1 - 2) + \gamma_2^2 v_2 / (v_2 - 2) \tag{2.16}$$
$$\kappa_4^Y = 3\gamma_1^4 v_1^2 / (v_1 - 4)(v_1 - 2) + 3\gamma_2^4 v_2^2 / (v_2 - 4)(v_2 - 2) \tag{2.17}$$

The first and third cumulant is zero, this fact implies the following limits of 2.11 and 2.14

$$\lim_{\kappa_3^Y \to 0} \bar{\alpha}_* = 3\left(\kappa_2^Y\right)^2 / \kappa_4^Y \tag{2.18}$$
$$\lim_{\kappa_3^Y \to 0} \delta_* = \sqrt{3(\kappa_2^Y)^3 / \kappa_4^Y} \tag{2.19}$$

These limits imply

$$\delta_* = \sqrt{\frac{\left(\gamma_1^2 \frac{v_1}{v_1-2} + \gamma_2^2 \frac{v_2}{v_2-2}\right)^3}{3\left[\gamma_1^4 \frac{v_1^2}{(v_1-4)(v_1-2)} + \gamma_2^4 \frac{v_2^2}{(v_2-4)(v_2-2)}\right]}}$$

and

$$\bar{\alpha}_* = \frac{\left(\gamma_1^2 \frac{v_1}{v_1-2} + \gamma_2^2 \frac{v_2}{v_2-2}\right)^2}{3\left[\gamma_1^4 \frac{v_1^2}{(v_1-4)(v_1-2)} + \gamma_2^4 \frac{v_2^2}{(v_2-4)(v_2-2)}\right]}$$

The approximate probability law can then be stated as:

$$NIG_*(\bar{\alpha}_*, 0, 0, \delta_*)$$

Thus we can use the NIG approximation to approximate the probability law for the sum of two unequally weighted student t random variables.

3 NIG approximation and its relation to Gram-Charlier and Edgeworth expansions

Here we discuss the NIG aproximation and how well it approximates a function of random variables compared to the Edgeworth and Gram-Charlier expansion. We do this by considering the regions for which Edgeworth and Gram-Charlier expansions produce unimodal and positive definite distributions and compare it with the similar region produced by the normal inverse Gaussian distribution. Furthermore, we also look at the tail behavior of the NIG approximation for some functions of random variables and compare them with the corresponding behavior for the Edgeworth and Gram-Charlier expansions. A first subsection is devoted to the regions of unimodality and positive definiteness whereas a second subsection covers the tail behavior comparison.

3.1 Regions of Unimodality and Positive Definiteness

In this subsection we derive the regions of unimodality and positive definiteness for the Edgeworth and Gram-Charlier expansions with the region of positive definiteness we mean the region where we are sure not to encounter negative probabilities. The region of unimodality is the region where the approximation density have one unique global maximum. Figure 1 such regions and was obtained via the dialytic method of Sylvester (see for instance Wang (2001)) for finding the common zeros for the Edgeworth and Gram-Charlier expansions.[1] Similar computations are reported in Barton and Dennis (1952) and Draper and Tierny (1972). Our results differ slightly from the results obtained in the earlier papers, due to nowadays' higher numerical accuracy compared to the earlier calculations.

[Insert Figure 1 somewhere here]

The region in Figure 1 are displayed in terms of the excess kurtosis and skewness coefficients for which the Gram-Charlier and Edgeworth and curves are unimodal and positive definite. Observe that we cut the expansion after reaching the fourth cumulant, which is the case in many applications (see for instance Johnson, Kotz, and Balakrishnan (1996)). Figure 1 also shows the regions in terms of skewness and kurtosis for which the normal inverse

[1] The computations and plot were generated with Maple software.

Gaussian law is defined. One immediately realizes that if one is interested in using the first four cumulants to approximate the probability distribution of a function of random variables under the assumption of unimodality one is better off using the NIG approximation. The NIG class covers a larger region with a valid probability measure as an approximation.

3.2 Tail behavior comparison in terms of fractiles- a comparison between NIG approximation, Gram-Charlier and Edgeworth expansion

In this subsection we focus on the comparison of how well the different approximation methods considered in this paper perform in terms of tail probabilities. The outline of this investigation is as follows: We start by simulating the one, fifth and tenth fractile from a function of random variables, with 5000 000 random draws. This is repeated 500 times, which yields an estimate of the true fractile. Next, we calculate the corresponding probability from the distribution functions implied by each approximation method. Finally, we compute the difference between the implied tail probability and the true one. We allow the Edgeworth and Gram-Charlier densities to have negative values however a negative tail probability or a tail probability above one is interpreted as a failure to approximate the function in question.

Some of the details of the design are as follows:

1. The function to approximate is based on a random coefficient model with an error term. The random coefficient model yields Y, which is standardized for the purpose of comparison. The standard random variable is denoted Y^*. More specifically,

$$Y = (X_1 X_2 + X_3)$$

$$Y^* = \frac{1}{\sqrt{\kappa_2^Y}} \left(X_1 X_2 + X_3 - \kappa_1^Y \right)$$

2. Next we need to assume the probability law for the random variables that enter the function. We choose three different random variables: (a) Gaussian, (b) student t and (c) a normal log normal mixing distribution (NLN) which is a skewed and leptokurtic distribution defined on $\mathbb{R}$.[2]

[2]The $NLN(\tilde{\mu}, \tilde{\sigma}, \delta)$ distribution is constructed as follows $\delta V + \sqrt{V} Z$ where $\mathcal{L}(V)$=$LN(\tilde{\mu}, \tilde{\sigma})$ and $\mathcal{L}(Z)$=$N(0,1)$

3. The final issue pertains to the choice of parameter space for the probability laws. We choose parameter spaces that imply fairly moderate excess kurtosis with and without skewness, and spaces that generate very large excess kurtosis. This is the case for model III and can be regarded as a test for how well the approximation works in a setting with extreme excess kurtosis.

The design of the comparison study is summarized in Table 1. One observation to note is that a small change in the parameter space can induce a very large change in the excess kurtosis and skewness. This is due to the fact that the excess kurtosis and skewness are nonlinear functions of the parameters we select. This effect is amplified when we consider more complicated distributional assumptions.

Table 1: Design of simulation study

Model	$\mathcal{L}(X_i)$	Case	θ_1	θ_2	θ_3	κ_2^Y	S_F^Y	K_F^Y
IA	$N(\theta_i)$	A	[1,1]	[1,1]	$[0,\frac{1}{4}]$	3.07	1.12	3.20
IB		B	[1,1]	$[\frac{1}{5},1]$	$[0,\frac{1}{4}]$	2.10	0.40	4.18
IIA	$t(\theta_i)$	A	[6]	[10]	[8]	3.21	0	7.43
IIB		B	[6]	[7]	[6]	3.60	0	9.71
IIIA	$NLN(\theta_i)$	A	$[-\frac{1}{10},1,-\frac{1}{8}]$	$[\frac{1}{8},\frac{1}{4},-1]$	$[\frac{1}{9},\frac{3}{2},-\frac{1}{300}]$	8.78	1.65	22.45
IIIB		B	$[-\frac{1}{10},1,-\frac{1}{8}]$	$[\frac{1}{8},\frac{1}{4},-1]$	$[\frac{1}{9},2,-\frac{1}{300}]$	13.64	0.083	182.17

Note that for the Gaussian probability law $\theta_i = (\mu_i, \sigma_i)$ for the student t law $\theta_i = (\upsilon_i)$ and for the NLN law $\theta_i = (\tilde{\sigma_i}, \tilde{\mu_i}, \delta_i)$.

The results are summarized in Table 2, where P denotes the true percentile whereas GC, E and NIG denote the corresponding percentile for the Gram-Charlier expansion, Edgeworth expansion and NIG approximation. The table also includes the differences between the true percentile and the percentile for each approximation method. The estimated fractiles and the associated standard error is also reported. The overall picture emerging from the Table are quite clear: when the distributional assumptions become more complicated, the performance of the Gram-Charlier and the Edgeworth expansion deteriorate more than that of the NIG approximation. Note also that for the Gram-Charlier and Edgeworth expansions the tail probabilities cease to exist for some of the fractiles. This is due to the fact that we are outside the boundaries for positive definiteness described in the previous section. Namely, tail probabilities less than zero or greater than one are obtained outside the feasible regions.

Table 2: Results comparison simulation study

Model IA

P	GC	E	NIG	P-GC	P-E	P-NIG	SE(Fractile)	Fractile
0.01	-0.010	0.007	0.003	0.020	0.003	0.007	0.002	-2.031
0.05	0.049	0.042	0.049	0.001	0.008	0.001	0.001	-1.253
0.1	0.137	0.108	0.116	-0.037	-0.008	-0.016	0.001	-0.950
Model IB								
0.01	0.020	0.020	0.012	-0.010	-0.010	-0.002	0.003	-2.690
0.05	0.029	0.031	0.058	0.021	0.019	-0.008	0.001	-1.516
0.1	0.072	0.069	0.120	0.028	0.031	-0.020	0.001	-1.031
Model IIA								
0.01	0.0425	0.0425	0.0233	-0.0325	-0.0325	-0.0133	0.0109	-2.667
0.05	0.0246	0.0246	0.0670	0.0254	0.0254	-0.017	0.0037	-1.533
0.1	0.0027	0.0027	0.1088	0.0973	0.0973	-0.0088	0.0027	-1.105
Model IIB								
0.010	0.053	0.053	0.025	-0.043	-0.043	-0.015	0.011	-2.693
0.050	0.011	0.011	0.068	0.039	0.039	-0.018	0.004	-1.520
0.100	-0.038	-0.038	0.108	NA	NA	-0.008	0.003	-1.085
Model IIIA								
0.010	0.097	0.118	0.029	-0.087	-0.108	-0.019	0.003	-2.374
0.050	-0.198	-0.204	0.070	NA	NA	-0.020	0.001	-1.311
0.100	-0.229	-0.295	0.110	NA	NA	-0.010	0.001	-0.924
Model IIIA								
0.010	1.050	1.050	0.015	NA	NA	-0.005	0.005	-2.544
0.050	-2.548	-2.548	0.033	NA	NA	0.017	0.001	-1.216
0.100	-3.907	-3.907	0.049	NA	NA	0.051	0.001	-0.811

4 The tail behavior of the NIG approximation

We continue our investigation of the NIG approximation by examining how well it fits the tails of the various functions introduced in the previous section. This is done by simulating the true density (denoted Y above) and simulating the approximating NIG density and finally compute a Quantile to Quantile plot for the 10% most extreme values for both tails i.e. a Quantile to Quantile plot only for the tails. The simulations were done with five million random draws so the Quantile to Quantile plots for the each tail consists of 500 000 observations. The results appear in Figure 2.

[Insert Figure 2 somewhere here]

The plots in Figure 2 confirm the pattern obtained in the fractile comparison with the Edgeworth and the Gram-Charlier discussed in the previous section. In particular, the tail behavior worsens when we impose assumptions regarding the random variables in the random coefficient model that imply more excess kurtosis and skewness. This is not surprising since the role of the higher moments for the behavior of the function of random variables increases in importance.

5 Concluding remarks

We introduced an approximation to unknown distributions via the NIG class and showed it to be a powerful tool to improve the calculations of tail probabilities when the information set is restricted to the first four cumulants. Using NIG approximations generates lesser approximation errors than using Gram-Charlier and Edgeworth expansions, especially when approximating a function with exhibits combinations of skewness and kurtosis that falls outside the region of positive definiteness of the Gram-Charlier and Edgeworth expansions.

References

Abramowitz, M., and I. A. Stegun (1972): *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. Dover publications Inc., New York.

Barndorff-Nielsen, O. E. (1978): "Hyperbolic Distributions and Distributions on Hyperbolae," *Scandinavian Journal of Statistics*, 5, 151–157.

——— (1997): "Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling," *Scandinavian Journal of Statistics*, 24, 1–13.

Barton, D., and K. Dennis (1952): "The Conditions Under Which Gram-Charlier and Edgeworth Curves are Positive Definite and Unimodal," *Biometrika*, 39, 425–428.

Charlier, C. V. (1905): "Uber Die Darstellung Willkurlicher Funktionen," *Arkiv fur Matematik, Astronomi och Fysik*, 9(20).

Draper, N., and D. Tierny (1972): "Regions of Positive and Unimodal Series Expansion of the Edgeworth and Gram-Charlier Approximations," *Biometrika*, 59(2).

Edgeworth, F. Y. (1896): "The Asymmertical Probability Curve," *Philosophical Magazine 5th Series*, 41.

——— (1907): "On the Representaion of Statistical Frequency by a Series," *Journal of the Royal Statistical Society, Series A*, 80.

Hall, P. (1995): *The Bootstrap and Edgeworth Expansion*. Springer-Verlag New York Inc., New York.

Johnson, N. L. (1949): "System of Frequency Curves Generated by Methods of Translation," *Biometrika*, 36, 149–176.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1996): *Continuous univariate distributions vol 1*. John Wiley and sons, New York.

Pearson, K. (1895): "Contributions to the Mathematical Theory of Evolution. II. Skew Variations in Homogenous Material," *Philosophical transactions of the Royal Society of London, Series A*, 186.

Sheshardi, V. (1993): *The Inverse Gaussian Distribution - a Case Study in Exponential Families*. Oxford university press, Oxford.

Wang, D. (2001): *Elimination Theory Methods and Practice in Mathematics and Mathematics-Mechanisation*. Shandong Education Publishing House, Jinan.

A Proof of Lemma 2.1

Proof.

The implied domain for $3\kappa_4^Y\left(\kappa_2^Y\right)/\left(\kappa_3^Y\right)^2$ for each case follows below:

i) $\rho > 0$

This implies that $\left(3\kappa_4^Y\left(\kappa_2^Y\right)/\left(\kappa_3^Y\right)^2 - 4\right) > 0$. This is fullfilled when the following is inequality is obtained:

$$3\kappa_4^Y\left(\kappa_2^Y\right)/\left(\kappa_3^Y\right)^2 > 4$$

$$\Leftrightarrow$$

$$3\left(K_F S_F^{-2}\right) > 4$$

ii) $\rho^{-1} < 1$

This implies that $\left(3\kappa_4^Y\left(\kappa_2^Y\right)/\left(\kappa_3^Y\right)^2 - 4\right)^{-1} < 1$. This is fulfilled when the following is inequality is obtained:

$$3\kappa_4^Y\left(\kappa_2^Y\right)/\left(\kappa_3^Y\right)^2 < 4 \vee 3\kappa_4^Y\left(\kappa_2^Y\right)/\left(\kappa_3^Y\right)^2 > 5$$

$$\Leftrightarrow$$

$$3\left(K_F S_F^{-2}\right) < 4 \quad \vee \quad 3\left(K_F S_F^{-2}\right) > 5$$

In order to $\rho > 0 \wedge \rho^{-1} < -1$ then $3\left(K_F S_F^{-2}\right) > 5$ ■

B Derivation of the approximation formulas

Proof. The problem can be described as finding a unique set of parameters that generates a particular set of the first four cumulants for the function of random variables, here denoted Y. This problem narrows down to solving a system of nonlinear equations.

State the system of nonlinear equations to solve as:

$$\mu_* + \frac{\bar{\varkappa}_* \delta_*}{\left(1 - \bar{\varkappa}_*^2\right)^{\frac{1}{2}}} = \kappa_1^Y \tag{B.20}$$

$$\frac{\delta_*^2}{\bar{\alpha}_* \left(1 - \bar{\varkappa}_*^2\right)^{\frac{3}{2}}} = \kappa_2^Y \tag{B.21}$$

$$\frac{3\bar{\varkappa}_*}{\bar{\alpha}_*^{\frac{1}{2}} \left(1 - \bar{\varkappa}_*^2\right)^{\frac{1}{4}}} = \frac{\kappa_3^Y}{\left(\kappa_2^Y\right)^{\frac{3}{2}}} \tag{B.22}$$

$$\frac{4\bar{\varkappa}_*^2 + 1}{\bar{\alpha}_* \left(1 - \bar{\varkappa}_*^2\right)^{\frac{1}{2}}} = \frac{\kappa_4^Y}{\left(\kappa_2^Y\right)^2} \tag{B.23}$$

B.23 yields:

$$3\frac{4\bar{\varkappa}_*^2 + 1}{\bar{\alpha}_* \left(1 - \bar{\varkappa}_*^2\right)^{\frac{1}{2}}} = \frac{\kappa_4^Y}{\left(\kappa_2^Y\right)^2} \Leftrightarrow$$

$$\bar{\alpha}_* = 3\frac{4\bar{\varkappa}_*^2 + 1}{\left(1 - \bar{\varkappa}_*^2\right)^{\frac{1}{2}}} \frac{\left(\kappa_2^Y\right)^2}{\kappa_4^Y} \tag{B.24}$$

and B.24 in the square of B.22 yields

$$\frac{3^2 \bar{\varkappa}_*^2}{3\frac{4\bar{\varkappa}_*^2+1}{\left(1-\bar{\varkappa}_*^2\right)^{\frac{1}{2}}} \frac{\left(\kappa_2^Y\right)^2}{\kappa_4^Y} \left(1 - \bar{\varkappa}_*^2\right)^{\frac{1}{2}}} = \frac{\left(\kappa_3^Y\right)^2}{\left(\kappa_2^Y\right)^3} \Leftrightarrow$$

$$\frac{3\bar{\varkappa}_*^2}{\left(4\bar{\varkappa}_*^2 + 1\right)} = \frac{\left(\kappa_3^Y\right)^2}{\kappa_4^Y \left(\kappa_2^Y\right)} \Leftrightarrow$$

$$\frac{4}{3} + \frac{1}{3\bar{\varkappa}_*^2} = \frac{\kappa_4^Y \left(\kappa_2^Y\right)}{\left(\kappa_3^Y\right)^2} \Leftrightarrow$$

$$\bar{\varkappa}_*^2 = \frac{1}{\varrho} \Leftrightarrow \tag{B.25}$$

$$\bar{\varkappa}_* = \frac{\text{signum}\left(\kappa_3^Y\right)}{\sqrt{\varrho}} \tag{B.26}$$

where :$\varrho = \left(3\kappa_4^Y \left(\kappa_2^Y\right) / \left(\kappa_3^Y\right)^2 - 4\right)$

B.25 in B.24 yields:

$$\bar{\alpha}_* = 3\frac{4/\varrho + 1}{\sqrt{(1-\varrho^{-1})}}\frac{\left(\kappa_2^Y\right)^2}{\kappa_4^Y} \tag{B.27}$$

B.27 and B.25 in B.21 yields:

$$\frac{\delta_*^2}{3\frac{4/\varrho+1}{(1-\varrho^{-1})^{\frac{1}{2}}}\frac{\left(\kappa_2^Y\right)^2}{\kappa_4^Y}\left(1-\varrho^{-1}\right)^{\frac{3}{2}}} = \kappa_2^Y \Leftrightarrow$$

$$3\left(4/\varrho + 1\right)\left(1 - \varrho^{-1}\right)\frac{\left(\kappa_2^Y\right)^3}{\kappa_4^Y} = \delta_*^2 \Leftrightarrow$$

$$\sqrt{3\left(4/\varrho + 1\right)\left(1 - \varrho^{-1}\right)\frac{\left(\kappa_2^Y\right)^3}{\kappa_4^Y}} = \delta_* \tag{B.28}$$

B.26 and B.28 in B.20 yields:

$$\mu_* + \frac{\frac{\text{signum}\left(\kappa_3^Y\right)}{\sqrt{\varrho}}\sqrt{3\left(4/\varrho+1\right)\left(1-\varrho^{-1}\right)\frac{\left(\kappa_2^Y\right)^3}{\kappa_4^Y}}}{\sqrt{(1-\varrho^{-1})}} = \kappa_1^Y \Leftrightarrow$$

$$\kappa_1^Y - \frac{\text{signum}\left(\kappa_3^Y\right)}{\sqrt{\varrho}}\sqrt{\left(12/\varrho + 3\right)\frac{\left(\kappa_2^Y\right)^3}{\kappa_4^Y}} = \mu_* \tag{B.29}$$

■

C Figures

,

Figure 1: Regions of positive definiteness and unimodality

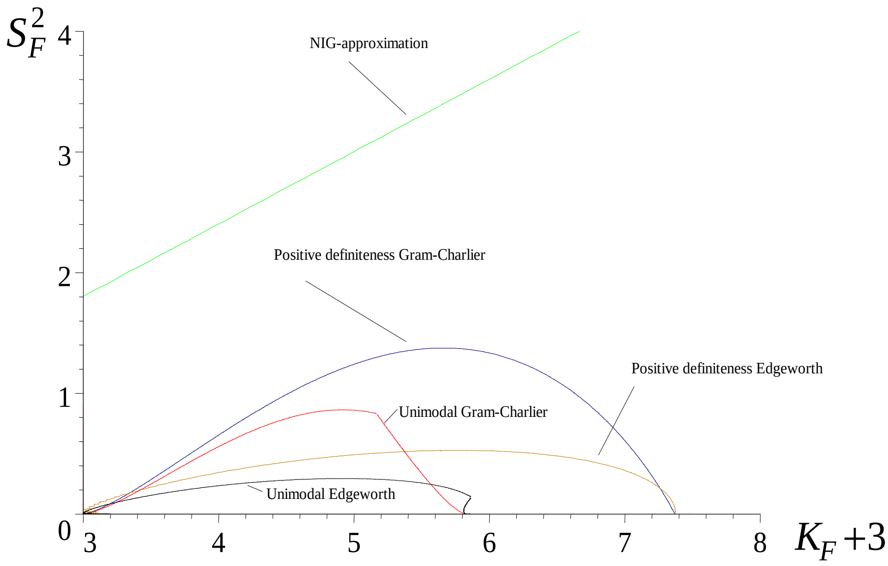
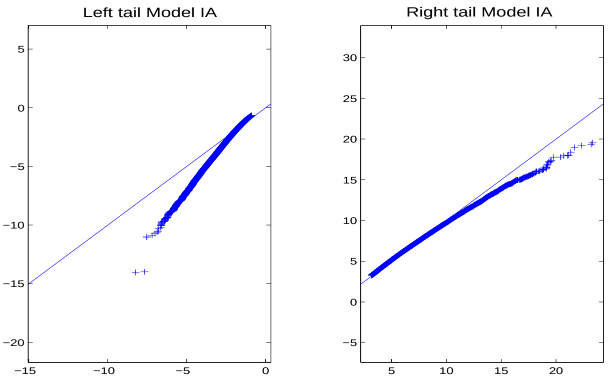
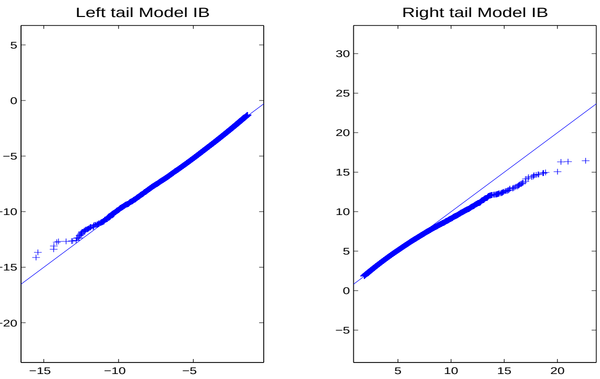


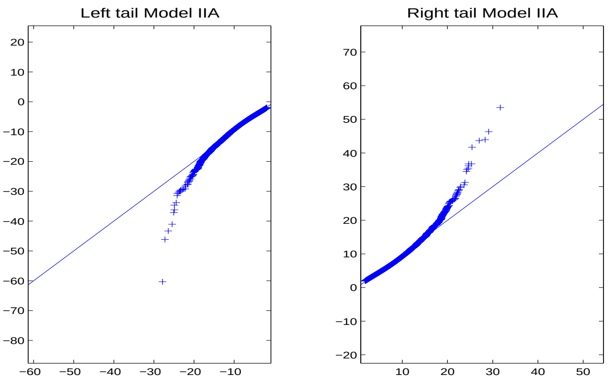
Figure 2: Result tail behavior of the NIG approximation



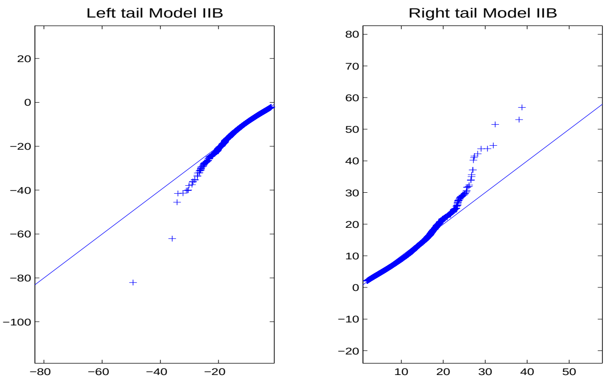
(a) Tail behavior Model IA



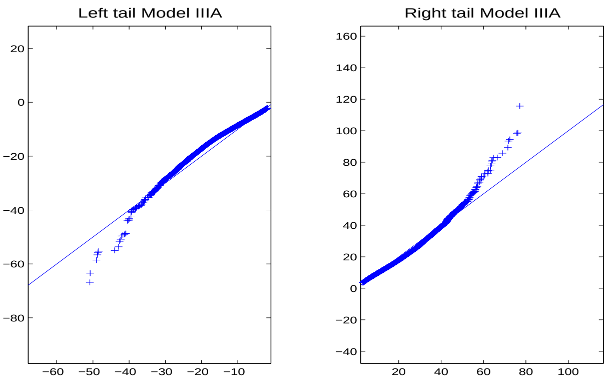
(b) Tail behavior Model IB



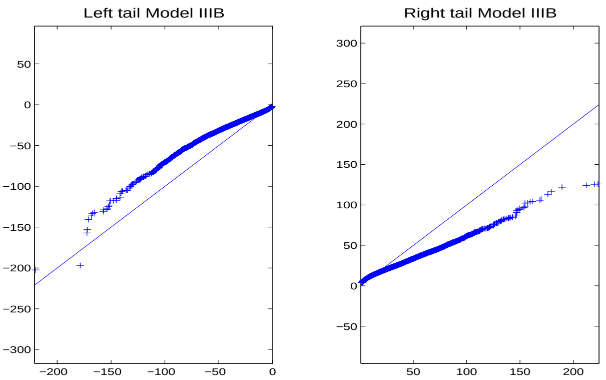
(c) Tail behavior Model IIA



(d) Tail behavior Model IIB



(e) Tail behavior Model IIIA



(f) Tail behavior Model IIIB