Some crude approximation, calibration and estimation procedures for NIG-variates

Jostein Lillestöl

Department of Finance and Management Science
The Norwegian School of Economics and Business Administration
Bergen, Norway

and

Sonderforchungsbereich 373
Humboldt Universität zu Berlin

First draft November 8, 2001
Revised November 8, 2002

Abstract

In this paper we explore some crude approximation, calibration and estimation procedures for Normal Inverse Gaussian (NIG) variates of potential use in risk management. Among others we treat in some detail the calibration of bivariate NIG consistent with marginal NIG.

KEY WORDS: Normal Inverse Gaussian distribution, risk management.

E-mail: jostein.lillestol@nhh.no

1 Background and outline

In finance it is an empirical fact that return distributions are often skewed and have heavier tails than the normal distribution. Risk management based on normal assumptions may therefore lead to underestimation of the risk. Theoretical researchers have tried to remedy this by offering other classes of distributions, first the stable Parertian class and more recently the generalized hyperbolic class. Nevertheless risk management in practice is still mostly based on normal assumptions, for a variety of reasons, among them: The lack of consensus among theorists, the mathematics is not understood, the computations are more demanding, the results are not easily communicated, and finally, the feeling that the methods fail to address issues just as important as skewness and heavy tails. For practical use there is a need to take a pragmatic view in order to overcome some of the reasons for not choosing one of the available alternatives to the normal model.

The purpose of this paper is to point out and explore some possible pragmatics, having applications to financial returns and risks in mind. A desirable property of the return distribution is that weighted sum of returns have distribution within the same class. This may put undesirable restrictions on available distribution classes, unless a more pragmatic view is taken. In Section 2 of this paper we explore in further detail an approximation suggested by Lillestöl (2000) concerning the distribution of weighted sum of returns from the univariate normal inverse Gaussian (NIG) distribution. In Section 2 we also explore some tail-based estimates suggested by Venter and de Jongh (2002). In Section 3 we explore possible pragmatics related to the multivariate NIG-family, among others: relate univariate and multivariate parameters, extend tail-based estimation to the bivariate case, look into the use of normal copulas as an alternative, and finally, examine an exchangeable multivariate NIG-structure. .

Before we start, let us give a brief account of our framework and some of the main issues. The framework is the generalized hyperbolic class (GH) of distributions, having the hyperbolic (H) and normal inverse Gaussian (NIG) as special cases, see Barndorff-Nielsen (1997) and Eberlein and Keller (1995), where details on the univariate and multivariate version of these distributions are given. They and others have demonstrated that these distributions fit a variety of financial data extremely well. These distribuitions also fit into a process context with generalized hyperbolic marginals, and finanicial theory is developed as alternative to the theory based on the Wiener process as the driving process, see e.g. Prause (1999a). The fact that the hyperbolic class is infinitely divisible is mentioned as an advantage, since we are then able to draw conclusions for different time horizons. However, this must be used with care. There are strong indications that skewness and kurtosis may not be the same for different time horizons, and the skewness may even change sign. This is reported, among others, by Stehle and Grewe (2001) who found that the monthly rates of return of 18 German stock mutual funds were negatively skewed, while

the annual returns were positively skewed.

The parameter estimation of generalized hyperbolic distributions can be estimated by likelihood methods, e.g. by the 'hyp' program developed by Blæsild and Sörensen (1992). The computational burden in the multivariate case is heavy. The 'hyp' program can, in a reasonable amount of time, only handle up to dimension 3. The computational burden becomes much easier by restriction to symmetric distributions. They have the convenient feature, once the multivariate relationships are settled, that the marginals are transparent without further computation, just as in the normal case, but contrary to the skew case. Bauer (2000) has demonstrated that the common approach to Value-at-Risk computations carry over to the class of elliptic distributions. The symmetric generalized hyperbolic distributions are within this class, and offers the additional conveniency for computation, just as fast as in the multivariate normal case.

The limitation to symmetric distributions for technical reasons is not much of a sacrifice for some financial assets, but not for others. The limitation cannot easily be removed unlesss one takes a more pragmatic approach in some other respect. To stay within the framework of GH is perhaps to allow too much generality, both subclasses H and NIG are general enough to fit a wide class of financial data well.

We prefer NIG to the competing class H, because of its nice additive features in relation to portfolios, see Lillestöl (2000). The marginalization features are maybe awkward, i.e. to infer from the joint distibution to the marginal distribution and vice versa. Moreover, there are indications in the literature (e.g. Prause, 1999b) that NIG fits the tails of distributions slightly better than H, which is appreciated by Value-at-Risk evaluators. The hyperbolic class H may be preferred for other reasons, pragmatic or not. An advantage may be faster ML-estimation due to fewer Bessel functions to compute. However, within the Bayesian framework the estimation can be done quite easily by Markov chain Monte Carlo methods, as shown by Karlis and Lillestöl (2002).

To sum up this introduction, the generalized hyperbolic class is able to pick up several stylished facts in financial data that are not accounted for by traditional methods.

2 Univariate NIG-variates

2.1 Framework

The Normal Inverse Gaussian (NIG) distribution introduced by Barndorff-Nielsen (1997) is a promising alternative for modelling financial data exhibiting skewness and fat tails. Various aspects of this are explored by him and his associates, see the reference list in the expository paper by Barndorff-Nielsen and Shephard (2001). Lillestøl (2000) has explored some facets of the additive properties of the NIG distribution useful for risk analysis. Although the NIG-family is closed

under convolution, non-equally weighted linear combinations of independent NIG-variates are not NIG.

We will consider a random variate X having a Norman inverse Gauusian distribution denoted by $NIG(\alpha, \beta, \mu, \delta)$. The distribution is characterized by 4 parameters $(\alpha, \beta, \mu, \delta)$, where α is related to steepness, β to asymmetry, and μ and δ are related to location and scale respectively, for short referred to below as the location and scale parameter. The NIG-distribution has a fairly complicated density $f(x; \alpha, \beta, \mu, \delta)$, but its moment-generating function is simple, namely

$$M_X(u) = exp(u\mu + \delta(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + u)^2}))$$

where $\alpha \geq \beta$. The Cauchy distribution is obtained as limiting case when $\alpha \to 0$ and the normal distribution is obtained as $\alpha \to \infty$ together with $\delta \to \infty$ so that $\delta/\alpha \to \sigma^2$. The distribution has semi-heavy tails which can be expressed as

$$f(x; \alpha, \beta, \mu, \delta) = \{ \begin{array}{ll} \sim C|x|^{-3/2} e^{bx} & x \to -\infty \\ \sim C|x|^{-3/2} e^{-ax} & x \to \infty \end{array}$$

where $a = \alpha - \beta$ and $b = \alpha + \beta$. From the moment-generating function it follows that (let $\gamma = \sqrt{\alpha^2 - \beta^2}$ for short)

$$
\begin{aligned}
EX &= \mu + \delta \cdot \frac{\beta}{\gamma} \\
varX &= \delta \cdot \frac{\alpha^2}{\gamma^3} \\
Skewness &= 3 \cdot \frac{\beta}{\alpha} \cdot \frac{1}{(\delta\gamma)^{1/2}} \\
Kurtosis &= 3 \cdot (1 + 4(\frac{\beta}{\alpha})^2) \cdot \frac{1}{\delta\gamma}
\end{aligned}
$$

The class of NIG-distributions has the following properties:

(i) If $X \sim NIG(\alpha, \beta, \mu, \delta)$ then $Y = kX \sim NIG(k^{-1}\alpha, k^{-1}\beta, k\mu, k\delta)$.

(ii) If $X_1 \sim NIG(\alpha, \beta, \mu_1, \delta_1)$ and $X_2 \sim NIG(\alpha, \beta, \mu_2, \delta_2)$ are independent then the sum $Y = X_1 + X_2 \sim NIG(\alpha, \beta, \mu_1 + \mu_2, \delta_1 + \delta_2)$.

From the above we realize that it may be difficult to compare different distributions with respect to proximity to the Gaussian or Cauchy by judging α and β regardless the chosen scale. A useful property for risk analysis is the fact that a sum of independent NIG-variates with common α and β, but different location and scale parameters, is itself NIG with parameters obtained by summing the location and scale parameters and keeping the others fixed. However, we also want to handle unequal weights. The assumption of independence may be valid for credit returns, but of course not for stock returns.

2.2   An approximation for sums of independent NIG-variates

We are mainly interested in the distribution in the context of credit risks or returns on financial assets. Consider therefore r joint returns $\mathbf{X} = (X_1, X_2, \ldots, X_r)$ and the return $Y = \mathbf{w}'\mathbf{X}$ on a portfolio of credit returns $\mathbf{w} = (w_1, w_2, \ldots, w_r)$. In the case of independent NIG-returns with common α and β parameter and weights that are all equal to $1/r$, we have that

$$Y = \bar{X} \sim NIG(r\alpha, r\beta, \bar{\mu}, \bar{\delta})$$

where the bars denote the average of the individual greeks. In the case of non-equal weights we do not have exact NIG. However, we may catch the main features by approximating as follows:

$$Y \approx NIG(\alpha_w, \beta_w, \mu_w, \delta_w)$$

where

$$\mu_w = \sum_i w_i \mu_i$$

$$\delta_w = \sum_i w_i \delta_i$$

$$\alpha_w = \frac{\sum_i w_i \delta_i}{\sum_i w_i^2 \delta_i \alpha_i^{-1}}$$

$$\beta_w = \frac{\sum_i w_i \delta_i \beta_i \alpha_i^{-1}}{\sum_i w_i^2 \delta_i \alpha_i^{-1}}$$

These approximations are obtained by matching terms (admittedly somewhat ad hoc) in the expressions for the expectation and the exponent of the moment-generating function. This involves a Taylor series expansion with accuracy that depends on the absolute value of the ratio β/α. We therefore expected the approximation to get better the closer we are to symmetric distribtions. We will see in an example below that this is not neccesarily so.

If we introduce the notation $\sigma_i^2 = \delta_i/\alpha_i$, not to be confused by variance, but motivated by its limiting property, we see that we can write

$$\alpha_w = \frac{\sum_i w_i \sigma_i^2 \alpha_i}{\sum_i w_i^2 \sigma_i^2}$$

$$\beta_w = \frac{\sum_i w_i \sigma_i^2 \beta_i}{\sum_i w_i^2 \sigma_i^2}$$

Note that if we let $a_i = \alpha_i - \beta_i$ and $b_i = \alpha_i + \beta_i$ denote parameters that determines the left-tails and right-tails of each distribution respectively, we get the same weighted sum formula for a_w and b_w as well.

We have investigated by simulations how well these formulas for NIG-parameter determination approximate the exact distribution. We did not expect it to be

very tight in general, but reasonably good at least for some cases of practical interest. We provide some results in the following example that expose both possibilities and limitations. Whether it is useful in a financial context, will depend on the available alternatives, one of them is not use any information on skewness and heavy tails at all.

Example 1

Consider the variates $X_i \sim NIG(2, \beta_i, 0, 1)$ with β_i's given in Table 1 and their equally weighted linear combinations.

	r	$(\beta_1, \ldots, \beta_r)$	approximate
a.	2	(1, 1)	(4, 2, 0, 1)
b.	2	(1, -1)	(4, 0, 0, 1)
c.	2	(1, 0)	(4, 1, 0, 1)
d.	3	(1, 1, -1)	(6, 1, 0, 1)

Table 1: Parameters for simulated examples

For each of the situations a-d in the table we have simulated the variates and computed the average, and then simulated the variate according to the approximation (case a. is exact and is included for comparison). We repeated the simulations n=1000 times and judged the approximations by plotting each pair of order statistics against each other, i.e. a kind of QQ-plot, see Figure 1. The fit is good when the points stay close to the equiangular line. The plots indicate good fit for case a, as expected, and for c and possibly also for d. For the case b the fit was apparently not good. Being the only symmetric case of the four, this may seem surprising.

The fit can be measured in a variety of different manners. We can of course use the common two-sample Kolmogorov-Smirnov statistic for which tables for determining p-values are readily available or we could use an Anderson-Darling type statistic, which pays more attention to the tails. Another possibility is to accumulate the absolute values of the differences between the order statistics, i.e.

$$D = \sum_{i=1}^{n} |\overline{X}_{(i)} - \widetilde{X}_{(i)}|$$

obtained from the computed averages $\overline{X}_i$ of the NIG-simulations and the simulated proxies $\widetilde{X}_i$. An alternative statistic would be to take $\widetilde{X}_{(i)} = \widetilde{G}^{-1}(i/(n+1))$, where $\widetilde{G}$ is the cumulative proxy NIG-distribution, which requires computation by numerical integration. These test statistics are measures of overall fit. For many applications it is more important to have a good fit in the tails. This is so for Value at Risk computations in finance, where the focus is on the lower
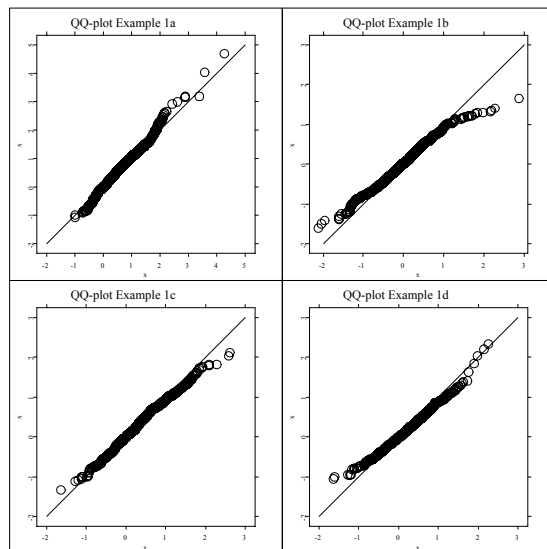
Figure 1: qq-plots

tail. A test statistic for this case is suggested by Venter and de Jongh (2002) as follows

$$D_{LT} = \sum_{i=1}^{n} |\log(\widetilde{G}(\overline{X}_{(i)})) - \log(i/(n+1))|$$

Since both log terms are close to zero as the subscripts are getting larger, the contribution in the sum comes mainly from the low order statistics. The corresponding upper tail statistic is

$$D_{UT} = \sum_{i=1}^{n} |\log(1 - \widetilde{G}(\overline{X}_{(i)})) - \log((n+1-i)/(n+1))|$$

Approximate p-values for the test statistics D, D_{LT} and D_{UT} can be obtained as parametric bootstraps, which are quite reliable in this context, see Stute et.al. (1993) or Davison and Hinkley (1999). The p-values for the test are given in Table 2. From this table we see that the statistics confirm the good fit in case a and c. The bad fit in case b does not show up in the KS-statistic, but there is an indication in D and in particular in the upper tail statistic D_{UT}. Case d may seem somewhat peculiar. The KS- and D-statistic indicate bad fit, but there is no indication of lack of fit in the tails.

Whether the results above was just "bad luck" for case b and perhaps "good luck" for the skew case c is investigated by repeated simulations of each of the four cases. We see the same picture in general. There is a tendency for the points of the QQ-plot to be steeper than the equiangular line in the cases b, c

	KS	P-KS	P-D	P-LT	P-UT
a.	0.0521	0.13	0.17	0.41	0.47
b.	0.0420	0.34	0.08	0.10	0.05
c.	0.0300	0.76	0.65	0.95	0.99
d.	0.0631	0.05	0.02	0.73	0.58

Table 2: Simulated examples: P-values model fit

and d, more so for b than c and d. This means that the approximations have tendency to possess lighter tails, and more so when variates of opposite skewness are added. In applications in practice variates are mostly skewed in one direction, e.g. in finance where large losses may incur. It is of some interest to compare the lower and upper fractiles for the true distribution and its approximation, as well as the fractile computed from the normal distributions with the same expectation and variance. They are given in the Table 3, where the true distribution is simulated based on n=100.000 observations. The approximate NIG is simulated as well, in order to avoid inversion of integrals involving Bessel functions.

	Fractile	0.01	0.05	0.95	0.99
a.	True	-0.637	-0.300	1.679	2.390
	NIG	-0.639	-0.301	1.695	2.411
	Normal	-0.886	-0.443	1.598	2.021
b.	True	-1.600	-0.993	0.997	1.575
	NIG	-1.230	-0.815	0.813	1.234
	Normal	-1.443	-1.021	1.021	1.443
c.	True	-0.936	0.556	1.261	1.862
	NIG	-0.910	0.546	1.163	1.696
	Normal	-1.022	-0.638	1.215	1.599
d.	True	-1.021	-0.612	1.041	1.532
	Nig	-0.781	0.487	0.869	1.243
	Normal	-0.986	-0.641	1.026	1.371

Table 3: Fractiles of true and approximate distribution

We see that the Normal 5% and 95% fractiles are not far off in any of the examples, but that 1% and 99% fractiles are off. This confirms what is generally known that 5% and 95% fractiles for NIG (and also finance data) are fairly well approximated by the corresponding Normal ones, and that divergences turn up in the more extreme fractiles. The exact case a is of course superfluous, but is included in order to compare with the normal approximation and to get an impression of simulation accuracy. In the case c with positive β we see, as expected, that the upper-fractiles are too low and the lower fractiles too high for both NIG and Normal, and that the NIG fractiles are closer to the true

extreme fractiles than the corresponding Normal ones. Note, however, that the symmetric fractile differences are about the same. The symmetric case b is disturbing. Neither Normal nor NIG is anywhere near the true 1% and 99% fractile, and the NIG-fractiles are further off than the Normal. Given the results in case c, we expected reasonably good results in case d as well. This turned out not to be the case, and we infer that the suggested NIG approximation is not likely to work if a sum contains summands that are skewed in opposite directions. However for larger portfolios of returns that are mainly skewed in one direction our experience indicates that the NIG approximation works as in case c. On the other hand, the normal approximation becomes better for the case of more summands as well.

Our hope of a good approximation in general which also works for the tails is not fullfilled, perhaps because it tries to be an overall approximation with moderate success. However, approximation adapted to a specific tail may be obtained along different lines of reasoning.

2.3   Tail based estimation

The estimation of NIG-parameters can be done by maximum likelihood methods. The 'hyp' program developed at Aarhus University by Blaesild and Sörensen (1992) is available for this purpose. A similar program is developed at Freiburg University by Eberlein et al. (1998). These programs also cover the multivariate case. Similar programs for the univariate cases exist elsewhere, e.g. at Potchefstroom University (Venter and de Jongh, 2001). The estimation is challenging since some of the parameters are hard to separate, the problem being that a flat-tailed distribution with a big scale is hard to distinguish from a fat-tailed distribution with small scale. The likelihood function with respect to these parameters then becomes very flat, and may have local mimima. Good starting values and security for convergence of the iterations are therefore essential for practical use. The estimation can also be done using empirical Bayes methods using the EM-algorithm as shown by Karlis (2002), and his program produces results in agreement with those mentioned above. Bayes methods using Markov chain Monte Carlo methods have also been tried, see Lillestöl (2001) and Karlis and Lillestöl (2002). Here the estimation problem essentially may be splitted in two, the estimation of inverse Gaussian parameters and the estimation of heteroscedastic regression.

A pragmatic approach to the estimation of NIG-parameters in the univariate case may be the one suggested by Venter and de Jongh (2002). Departing from the approximate expressions for the tail of the NIG-density given in the previous section, they derive the following approximation for $a = \alpha - \beta$ and $b = \alpha + \beta$:

$$a \sim \frac{1}{2} \frac{q_{1-\varepsilon} + E(X|X > q_{1-\varepsilon})}{E(X^2|X > q_{1-\varepsilon}) - q_{1-e}E(X|X > q_{1-e})}$$

$$b \sim -\frac{1}{2} \frac{q_{\varepsilon} + E(X|X < q_{\varepsilon})}{E(X^2|X < q_{\varepsilon}) - q_e E(X|X < q_e)}$$

8

where q_ε and q_{1-e} are the left and right ε-fractile of the distribution respectively. Estimates are then obtained from the order statistics. $X_{(1)}, X_{(2)}, ..., X_{(n)}$. After the choice of a suitable ε we can estimate the q's by the corresponding fractiles in the empirical cdf and the expectations by averaging over the observations and squared observations beyond the appropriate fractile. The estimates of α and β are then obtained from a and b by half their sum and half their difference respectively. We see also that the product of a and b estimates γ^2. We will name these estimates "Tail Based Estimates", TBE for short. Although Venter and de Jongh originally suggested this procedure for preliminary estimates to be used as starting values for ML-estimation, it is tempting to stick to it in practice for the following reasons: It is very transparent, and involves directly the expression $E(X|X < q_\varepsilon)$ related to "shortfall", which is of prime importance to risk managers, i.e. answers the question "if return is bad, how bad can we expect it to be?".

When α , β and γ are estimated, we can get estimates of δ and μ by just replacing the mean and the variance in the expression in Section 2 by their empirical counterparts. We will see how this estimation approach works in some examples using simulated data. Again we will expose weak points as well as some comforting.

Example 2

Consider a series n=400 NIG-observations simulated from $(\alpha, \beta, \mu, \delta) = (2, 1, 2, 1)$. This is a situastion of some challenge: Considerable skewness in conjunction with the small α and only moderate sample size for identifcation. We got the estimates given in Table 4.

Parameters					
Estimate	α	β	μ	δ	γ
MLE	2.490	1.343	1.876	1.049	1.855
TBE-5%	4.120	3.069	0.766	1.599	2.754
TBE-1%	3.360	1.988	-0.112	3.626	2.709

Table 4: Example Tail-Based estimates

We see that the ML-estimate did reasonably well, but that the TB-estimates are far off. We have repeated this simulation 100 times in order to get an impression of the distribution of the TB-estimates. It is not clear which one of the two TB-estinates is the better. The experience is that for 1 % one frequently gets too few observations in the tails to get useful estimates. If we increase the sample size to about 1000 it seems that 1 % is a viable alternative. The results for TBE-5% are shown in the graphs of Figure 2.

We see that both α and β are systematically overestimated. On the other hand μ is underestimated, while δ is overestimated (they become negatively correlated by their definition). In some contexts we are not primarily interested in precise estimates as long as we can fairly represent the features beyond second
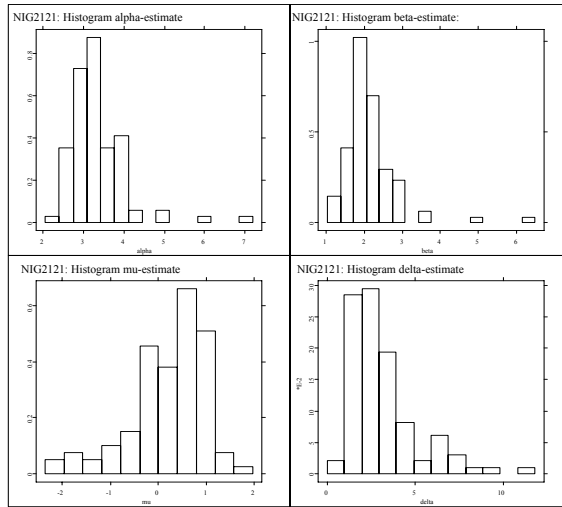
9

Figure 2: Histograms of parameter estimates

moments. We see that they are given mainly in terms of the ratio β/α and the product $\delta\gamma$. The histograms for the estimates of these two are in Figure 3, and we see that they both are overestimated but less so. Since the skewness is expressed by a ratio involving these two, this is rather satisfactory.
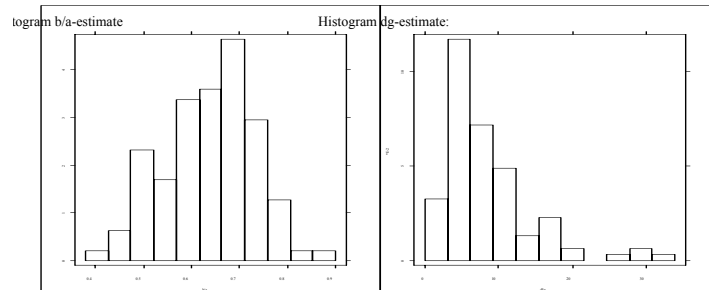

Figure 3: Histograms of parameter estimates

The way the TB-estimates of μ and δ are defined we are secured that the fitted NIG has estimate of expectation and variance that corresponds to using the first and second order moments only. In the risk management context of Value at Risk this may be a major step, obtained by simple means, e.g. for simulation of scenarios that are more realistic wrt. extreme events. Of course one could simulate directly from the empirical cdf, forsaking the opportunity to do parameters comparisons and vary these, i.e. to put in more or less skewness

and heavy tail at will.

3 Multivariate NIG-variates

3.1 Framework

An approximation for i.i.d. variates may be of some use for evaluating credit risk, although some weak dependencies may be expected, for instance due to swings in the economy. For portfolio risk involving equities and/or derivatives the correlations are the key issue. Until recently the only feasible parametric approach for fast computation has been based on multinormal assumptions, thus negliecting skewness and heavy tails. Now considerable efforts are made to provide a wider choice of distributions.

A vector of returns $\mathbf{X}$ is distributed multivariate NIG $(\alpha$, $\boldsymbol{\beta}$, $\boldsymbol{\mu}$, δ ,$\boldsymbol{\Phi})$ where α and δ are scalars, $\boldsymbol{\beta} = (b_1, b_2, \ldots, b_r)$ and $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_r)$ are vectors and $\boldsymbol{\Phi} = (\phi_{ij})$ is positive definite matrix with determinant 1. The moment generating function is

$$M_{\mathbf{X}}(\mathbf{u}) = exp(\mathbf{u}'\boldsymbol{\mu} + \delta(\sqrt{\alpha^2 - \boldsymbol{\beta}'\boldsymbol{\Phi}\boldsymbol{\beta}} - \sqrt{\alpha^2 - (\boldsymbol{\beta} + \mathbf{u})'\boldsymbol{\Phi}(\boldsymbol{\beta} + \mathbf{u})}))$$

The expectation vector of $\mathbf{X}$ is

$$EX = \boldsymbol{\mu} + \delta(\alpha^2 - \boldsymbol{\beta}'\boldsymbol{\Phi}\boldsymbol{\beta})^{-1/2}\boldsymbol{\beta}\boldsymbol{\Phi}$$

and the covariance matrix is

$$\Sigma = \delta(\alpha^2 - \boldsymbol{\beta}'\boldsymbol{\Phi}\boldsymbol{\beta})^{-1/2}(\boldsymbol{\Phi} + (\alpha^2 - \boldsymbol{\beta}'\boldsymbol{\Phi}\boldsymbol{\beta})^{-1}\boldsymbol{\Phi}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{\Phi})$$

Consequently $\boldsymbol{\Phi}$ relates to the covariance in a fairly complicated manner involving all other parameters as well. Among others we see that $\boldsymbol{\Phi}$ diagonal is not sufficient for $\boldsymbol{\Sigma}$ to be diagonal and vice versa, unless in the symmetric case when $\boldsymbol{\beta}$ is zero. In some cases we may assume that $\boldsymbol{\beta}'\Phi\boldsymbol{\beta}$ is negligible compared to α^2 and use as approximation

$$\Sigma \approx \frac{\delta}{\alpha}(\boldsymbol{\Phi} + \frac{1}{\alpha^2}\boldsymbol{\Phi}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{\Phi}) \approx \frac{\delta}{\alpha}\boldsymbol{\Phi}$$

Then the second term in the middle is likely to be negligible as well, and we may just as well use the even cruder approximation. This amounts to assuming that Φ diagonal represents approximately uncorrelated returns.

3.2 Relating univariate and multivariate parameters

We are mainly interested in the return $Y = \mathbf{w}'\mathbf{X}$ on a portfolio $\mathbf{w}$. The moment generating function is

$$\begin{aligned} M_Y(u) &= M_{\mathbf{X}}(u\mathbf{w}) \\ &= exp(u\mathbf{w}'\boldsymbol{\mu} + \delta(\sqrt{\alpha^2 - \boldsymbol{\beta}'\boldsymbol{\Phi}\boldsymbol{\beta}} - \sqrt{\alpha^2 - (\boldsymbol{\beta} + u\mathbf{w})'\boldsymbol{\Phi}(\boldsymbol{\beta} + u\mathbf{w})})) \end{aligned}$$

This is one-dimensional $NIG(\alpha_w, \beta_w, \mu_w, \delta_w)$ where

$$
\begin{aligned}
\mu_w &= \mathbf{w}'\boldsymbol{\mu} \\
\delta_w &= \phi_w \cdot \delta \quad \text{where} \quad \phi_w = (\mathbf{w}'\boldsymbol{\Phi}\mathbf{w})^{1/2} \\
\beta_w &= \phi_w^{-2}\mathbf{w}'\boldsymbol{\Phi}\boldsymbol{\beta} \\
\gamma_w &= \phi_w^{-1}\gamma \quad \text{where} \quad \gamma = (\alpha^2 - \boldsymbol{\beta}'\boldsymbol{\Phi}\boldsymbol{\beta})^{1/2} \\
\alpha_w &= (\gamma_w^2 + \beta_w^2)^{1/2}
\end{aligned}
$$

The marginal distribution of the component X_i's are obtained by letting $w_i = 1$ and $w_j = 0$ for $j \neq i$. We then get $\mu_w = \mu_i$ and (note that $\phi_i^2 = \phi_{ii}$)

$$
\begin{aligned}
\delta_i &= \phi_i \cdot \delta \\
\beta_i &= \phi_i^{-2}\sum_j \phi_{ij}b_j \\
\gamma_i &= \phi_i^{-1}\gamma \\
\alpha_i &= (\gamma_i^2 + \beta_i^2)^{1/2}
\end{aligned}
$$

So, if we go to the multivariate setting outlined in the previous section, we are rewarded by getting all marginal and linear combinations univariate NIG. Note, however, that independent univariate NIG-variates are not jointly multivariate NIG in the sense above, which is contrary to the case of the multinormal distribution. Note also that the alfa-scalars here do not correspond to an alfa-parameter common to all the marginals. We see that the marginal α_i's are affected jointly by $\boldsymbol{\beta}$ and $\boldsymbol{\Phi}$. It is worthwhile to note that $\phi_i^2(\alpha_i^2 - \beta_i^2)$ must be constant for all i. This makes it difficult to interpret parameters and a bit awkward to establish a joint model specification from given marginal specifications. In practice we may want to do that in order to establish simulation schemes that corresponds to common knowledge, which is mostly about the marginals for features beyond second order properties.

Example 3

Consider the bivariate case when $\boldsymbol{\Phi}$ is diagonal, which means non-negative correlation, the size depending on the skewnesses. In this case it follows that $\beta_i = b_i$ for $i = 1, 2$ (this holds for any dimension). Suppose we want to have equal marginal α-parameters. If the diagonal elements ϕ_i^2 are different we get $\alpha_i^2 = (\phi_1^2 b_1^2 - \phi_2^2 b_2^2)/(\phi_1^2 - \phi_2^2)$ while $\alpha^2 = (\phi_1^4 b_1^2 - \phi_2^4 b_2^2)/(\phi_1^2 - \phi_2^2)$. Note that $b_1 = b_2 = b$ now implies $\alpha_i^2 = b^2$ and $\alpha^2 = (\phi_1^2 + \phi_2^2)b^2$. Note also that in the case of diagonal $\boldsymbol{\Phi}$ and equal marginal α-parameters, the diagonal elements are equal if and only if the skewnesses for the marginals are equal, and then the common α_i is given by $\alpha_i^2 = \alpha^2 - \beta^2$, where β is their common skewness. If we instead require that the marginal skewnesses are equal, i.e. $\beta_i = b_i = b$, we get the restriction $\phi_1^2\alpha_1^2 - \phi_2^2\alpha_2^2 = (\phi_1^2 - \phi_2^2)b^2$.

As a numerical example take $\delta = 1$ and $\mu = (0,0)'$ and

$$\Phi = \left[\begin{array}{cc} \frac{4}{3} & 0 \\ 0 & \frac{3}{4} \end{array} \right]$$

If we add the restriction of a common α-parameter for the marginal, all parameters are uniquely determined and examples of this are presented in the first two columns of Table 5. If we take the common β-parameters equal, say to b, we still have a choice beetween different α_i's for given b (but their squares are linearly related). As an example take $b = 1$ and $\alpha_1 = 2$ and compute the rest. We then get the righthand side column of the table.In dimension more than two it gets more complicated.

	$\beta = (3/2, 1)$	$\beta = (1, 1/2)$	$\beta = (1, 1)$
α_1	1.96	1.40	2.00
α_2	1.96	1.40	2.52
α	2.43	1.68	2.47
δ_1	1.15	1.15	1.15
δ_2	0.87	0.87	0.87
γ_1	1.27	0.98	1.73
γ_2	1.69	1.31	2.31
EX_1	1.37	1.18	0.86
EX_2	0.51	0.33	0.37
$varX_1$	2.19	2.40	2.60
$varX_2$	0.69	0.76	0.45
$Corr$	0.39	0.25	0.36
$Skew_1$	1.89	2.01	1.40
$Skew_2$	1.26	1.01	0.84

Table 5: Parameter determination

In practice it is more likely to have opinions on the marginal α-parameters, based on experience on where the kind of data at hand should be placed on the vertical axis between the Cauchy and the Normal distribution. Then it is partly a question whether skewness or correlation is the dominant feature. It seems perhaps more convenient to start with a matrix Φ, and then, for the chosen α, try out a reasonable $\boldsymbol{\beta}$-vector, by computing marginals.

In this section we have focused on coherent model specification and not on estimation. We have mentioned earlier the problems of estimation, in particular in the multivariate case. There exists various possibilities for a pragmatic solution to this, depending on the context. We will discuss two possible approaches in the following.

3.3 Tail based estimation

It is possible to extend the idea of tail based estimation to get estimates for bivariate NIG by combining the results of Section 2.3 and 3.2. We have to add something that can pick up the correlation structure. One possibility is to look at the tail behaviour of the sum and difference. Let $\mathbf{X} = (X_1, X_2)'$ be distributed $NIG\,(\alpha\,,\,\boldsymbol{\beta}\,,\,\boldsymbol{\mu}\,,\,\delta\,,\,\boldsymbol{\Phi})$ where $\beta = (b_1, b_2)'$. Then let $\mathbf{Y} = (Y_1, Y_2)'$ where

$$
\begin{aligned}
Y_+ &= X_1 + X_2 \\
Y_- &= X_1 - X_2
\end{aligned}
$$

Y_+ and Y_- are distributed $NIG(\alpha_+, \beta_+, \mu_+, \delta_+)$ and $NIG(\alpha_-, \beta_-, \mu_-, \delta_-)$ respectively. From the marginals of $\mathbf{X}$ and $\mathbf{Y}$ we can estimate $\alpha_i, \beta_i, \gamma_i, \delta_i, \mu_i$ for $i = 1, 2$ and $\pm$ as we did in Section 3.2. These univariate NIG-parameters are related by (using compressed notation and a ϕ-notation similar to that of Section 3.2)

$$
\begin{aligned}
\mu_{\pm} &= \mu_1 \pm \mu_2 \\
\delta_{\pm} &= \delta\,\phi_{\pm} \\
\phi_{\pm}^2 &= \phi_1^2 + \phi_2^2 \pm 2\phi_1\phi_2\rho \\
\beta_{\pm} &= \phi_{\pm}^{-2}(b_1\phi_1^2 \pm b_2\phi_2^2 + (b_1 \pm b_2)\phi_1\phi_2\rho) \\
\gamma_{\pm} &= \phi_{\pm}^{-1}(\alpha^2 - (b_1^2\phi_1^2 + b_2^2\phi_2^2 \pm 2b_1b_2\phi_1\phi_2\rho))^{1/2} \\
\alpha_{\pm} &= (\gamma_{\pm}^2 + \beta_{\pm}^2)^{1/2}
\end{aligned}
$$

We need the estimates of $\alpha, b_1, b_2, \phi_1, \phi_2$ and ρ, which can be obtained from these equations combined with those of Section 2.3. It is helpful to note that

$$
\begin{aligned}
\tau &= \frac{\phi_2}{\phi_1} = \frac{\delta_2}{\delta_1} = \left(\frac{\gamma_2}{\gamma_1}\right)^{-1} \\
\theta &= \frac{\phi_+}{\phi_-} = \frac{\delta_+}{\delta_-} = \left(\frac{\gamma_+}{\gamma_-}\right)^{-1} \\
\rho &= -\frac{1+\tau^2}{2\tau} \cdot \frac{1-\theta^2}{1+\theta^2}
\end{aligned}
$$

The third equation is obtained by division of the two expressions for $\phi_{\pm}^2$ above using the first and second equation and solving for ρ. Remembering that $\phi_1^2\phi_2^2 = (1-\rho^2)^{-1}$, we can solve for ϕ_1 and ϕ_2 and then finally obtain b_1, b_2, α and δ. However, this kind of artificial data augmentation leads to an overidentified situation. For instance, we can either take γ's or δ's as basis for estimating τ and θ. This can be resolved by different means. One possibility is to use both and "symmetrice" by taking geometric means, and similarly for δ and γ. We then get the proxy formulas

$$\tilde{\tau} = \sqrt{\frac{\delta_2}{\delta_1} \cdot \frac{\gamma_1}{\gamma_2}}$$

$$\tilde{\theta} = \sqrt{\frac{\delta_+}{\delta_-} \cdot \frac{\gamma_-}{\gamma_+}}$$

$$\tilde{\delta} = \sqrt{\frac{\delta_1}{\phi_1} \cdot \frac{\delta_2}{\phi_2}}$$

$$\tilde{\gamma} = \sqrt{\gamma_1 \phi_1 \cdot \gamma_2 \phi_2}$$

Example 4

We have simulated 400 observations according to the parameters of Example 3 (Table 5 right column) i.e.

$$\Phi = \left[\begin{array}{cc} \frac{4}{3} & 0 \\ 0 & \frac{3}{4} \end{array} \right]$$

$\beta = (1,1), \mu = (0,0)', \delta = 1$ and $\alpha = 2.47$, which means $\gamma = 2.00$. The scatter diagram is given in Figure 4 and the smoothed density plot in Figure 5.
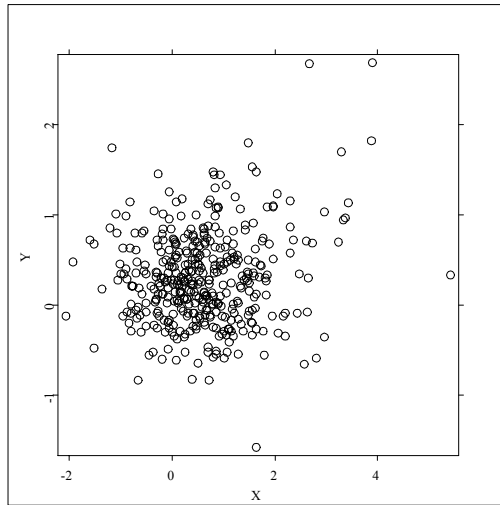

Figure 4: Scatterplot of original data n=400

The crude estimates obtained by using 5% tails are given in Table 6.
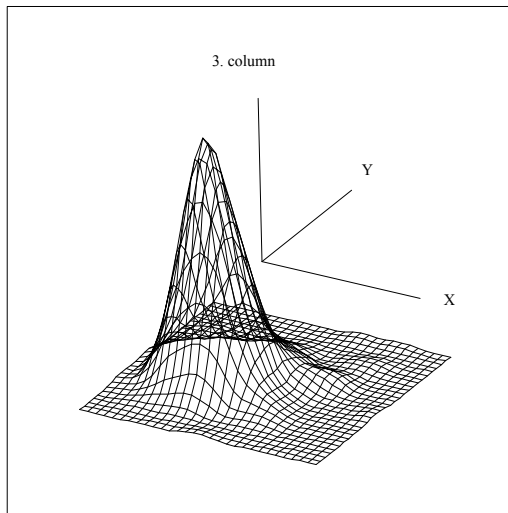
15

Figure 5: Smoothed density estimate

We see that some estimates are surprisingly good, and some are a bit off. The most pronounced deviation is the diminished skewness of the first component and the enlarged disparity between the matrix diagonal terms. The corresponding estimates using the 1 % tails gives similar results except that α now is overestimated to 3.142, the disparity of the skewnesses is about the same, but reversed (!), and a slightly larger matrix off-diagonal term occurs. Although there are some discrepancies from the true values, the result is not bad, and not worse than expected based on experience from univariate estimation. We also computed estimates based on a simulated sample size of n=1000. Now 1 % tails are preferred and results are substantially as above, but with somewhat less dis-

Parameter estimates				
i	α_i	β_i	μ_i	δ_i
1	1.502	0.538	0.148	1.189
2	2.807	1.081	0.046	0.598
+	1.238	0.537	0.287	1.258
−	1.341	0.172	0.117	1.419
—	ρ	ϕ_{11}	ϕ_{22}	ϕ_{12}
—	0.034	1.915	0.523	0.033
δ	b_1	b_2	α	γ
0.843	0.511	1.048	2.248	1.909

Table 6: Parameter estimates of bivariate NIG

16

parity between the skewnesses. The findings are supported by repeated (though not extensive) simulations. The general experience is that the estimates of α and the skewnesses may show some instability, in particular for the smallest sample size, which seems to be balanced off by a reasonably good estimate of γ.

3.4 The use of copulas for NIG-data

Dependence in finance is, as mentioned above, mostly handled by normality and linear correlation methods. An approach for handling non-normal data without relying on linear correlation is offered by copulas, see Nelsen (1999). A copula is a device to parametrize dependence structures according to given marginals. In the bivariate case we have

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$$

where $C(u_1, u_2)$ is a copula function, which is a bivariate cumulative distribution on the unit square indexed by a parameter θ that accounts for possible covariation. Some copulas are given in Table 7 with catalogue numbers according to Nelson (1999).

i	$C_i(u_1, u_2)$	Parameter region
3	$\frac{u_1 u_2}{1-\theta(1-u_1)(1-u_2)}$	$\theta \in [-1, 1]$
4	$\exp\left(-\left[(-\ln(u_1)^\theta + \ln(u_2)^\theta\right]^{1/\theta}\right.$	$\theta \in [1, \infty]$
6	$1 - \left[(1-u_1)^\theta + (1-u_2)^\theta + (1-u_1)^\theta \cdot (1-u_2)^\theta\right]^{1/\theta}$	$\theta \in [1, \infty]$
9	$u_1 u_2 \exp(-\theta \ln u_1 \ln u_2)$	$\theta \in [0, 1]$
10	$\frac{u_1 u_2}{1-(1-(1-u_1)^\theta(1-u_2)^\theta)^{1/\theta}}$	$\theta \in [0, 1]$
12	$\left[1 + (u_1^{-1} - 1)^\theta + (u_2^{-1} - 1)^\theta\right]^{-1}$	$\theta \in [1, \infty]$

Table 7: Some copulas numbered according to Nelsen (1999).

Ideally we should go for copulas with NIG-marginals. However, findings in the literature seem to indicate that simple copulas based on on normal marginals most often outperform linear correlation on real (heavy tailed) data. Thus instead of going for the best, we could see how normal copulas behave on NIG-data. As an example we will again look at the situation described in Example 3 (righthand column), where the diagonal $\boldsymbol{\Phi}$-matrixs in fact corresponds to linear correlation of 0.199. Again we use the simulated dataset of n=400 observations plotted in Figure 4, which has empirical correlation of 0.149. We tried the copulas in Table 7 as well as a few others. The copula and dependence parameter may be chosen by visual comparison of the scatterplot in Figure 4 with scatterplots of data simulated for different choices in accordance with the observed marginal structure, i.e. in the case of normal copulas just the means and the standard deviations. For a chosen copula, the dependence parameter θ may be estimated by maximum likelihood, exact or by some approximate technique. It

turned out that copula 4, 6 and 12 were the ones that looked best by visual inspection and also behaved well numerically using the quantlets VaRsimcopula and VaRfitcopula in XploRe. However in all three cases the maximum likelihood estimate of θ turned out to be 1, corresponding to independence for case 4 and 6, while the interval search option in the first two cases gave parameter values slightly different from 1, although visual inspection of simulated data suggested a slightly larger value. As an illustration we may compare the scatterdiagram of the original data with that simulated by copula 6 by taking $\theta = 1.25$.
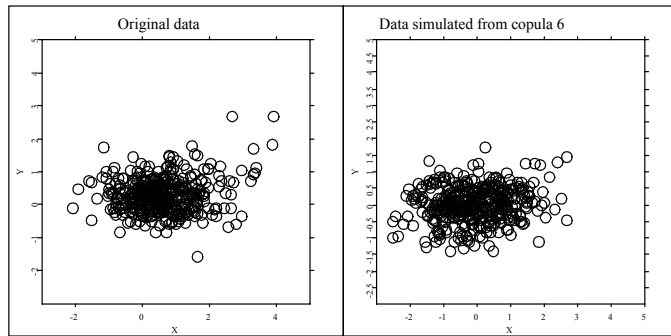

Figure 6: Scatterplot comparison of fitted copula 6

This does not look bad, but a closer examination reveals that the peakedness of the original data is not reflected by the simulated data according to the fitted copula 6. Note also that the plots support the earlier remarks in Section 2.2 that normal methods applied to heavier tailed data quite often are able to reproduce 5% fractiles correct, but not 1% fractiles, which are most relevant for risk analysis. The findings in this example, which are not at all surprising, are confirmed by repeated simulations. Besides this we draw the tentative conclusion that normal copulas may also tend to neglect weak correlation in NIG-type data that could be of importance in financial VaR-type calculations. The use of copulas with NIG-marginals will be explored in a separete paper.

3.5   Reduction to bivariate case by principal components

Let us noe return to the multivariate case. In risk management the correlations between the returns of the various assets that can go into a portfolio is crucial. The success of the multinormal distribution in that all you need in this context is the pairwise correlations besides expectations and variances. It is not easy to establish and represent the added information neccessary for the corresponding analysis based on NIG assumptions. Suppose we have large number of assets

that can potentially go into a portfolio. One possibility is to do a principal component analysis of the covariance matrix, and then use a small number of principal components to establish the main risk features in the market. If the original (large) return vector was multivariate NIG, then the principal components, as linear combinations, are univariate NIG (and uncorrelated). The NIG-parameters of each of these can then be estimated. However, since uncorrelatedness is not independence in the NIG case, we must be careful. If we stick to just the first principal component there is no problem. If we want to keep two, we have to consider these as bivariate NIG and estimate parameters accordingly. In view of the problems of multivariate NIG estimation for dimension more than three, it seems fruitless to keep more than three principal components. Each asset return can then be expressed linearly by the low order principal component(s) plus a remainder term consisting of the omitted ones. For risk management one can neglect the remainder, but scale up the expression so that we get the "correct" variance. We then have to do a parameter correction according to property (i) of univariate NIG in Section 2.1. This way one may be able to pick up both the correlation structure in the market and that returns are skewed, if so.

3.6 Approximation by exchangeable structure

In some cases it may be helpful to assume an exchangeable correlation structure, i.e. assume that the components of $\boldsymbol{\beta} = (b, b, \ldots, b)$ are all equal and that $\boldsymbol{\Phi} = (\phi_{ij})$ has equal diagonal and equal off-diagonal elements. In this case we just have to specify the dimension r and the ratio c between the off-diagonal and the diagonal elements (which has to be greater than $-1/(r-1)$ to achieve a positive definite matrix). Then the a and b of the multivariate distribution is uniquely determined by the specification of the (common) marginal α_i and β_i. The formulas for the diagonal element in order to achieve determinant 1 is

$$d = (1-c)^{-1}(1 + \frac{c}{1-c}r)^{-1/r}$$

If we let $p = 1 + (r-1)c > 0$ we have (the details are given in Lillestöl (1998) and not repeated here)

$$
\begin{aligned}
b &= \beta_i/p \\
\alpha &= d^{1/2}(\alpha_i^2 - \beta_i^2(1 - rp^{-1}))^{1/2}
\end{aligned}
$$

We see that the components of the $\boldsymbol{\beta}$-vector do not depend on α_i, and is just a rescaling up or down according to whether c is negative or positive. The influence on the α is more complicated. Some examples may provide a feeling for the relation between the joint and marginal parameters. It is easily checked that the common correlations are given by

$$\frac{c+z}{1+z}$$

where $z = db^2 p^2 (\alpha^2 - db^2 pr)^{-1} > 0$. For $c = 0$, that is diagonal Φ, the correlation is positive. Zero correlation requires negative $c = -z$. Note however that this does not correspond to independence. This gives an equation for c that can be solved numerically. The case $r = 2$ is particaularly simple. We then get the cubic equation $c^3 - c - k^2/(1 - k^2) = 0$ where $k = \beta_i/a_i$. Solutions are given in Table 8.

			Off-diagonal ratios c for given k				
k	0.0	0.1	0.2	0.3	0.4	0.5	0.6
c	0.000	-0.010	-0.042	-0.099	-0.198	-0.395	-

Table 8: Off-diagonal ratios for uncorrelated case

Example 5

Consider first the bivariate case r=2. In Table 9 we have tabulated b for varying c and β_i (nonnegative w.l.g.) valid for any α_i.

We see that the components of the $\boldsymbol{\beta}$-vector are up 25% from the individual β_i for c=-0.2 and down 16.7% for c=0.2. In Table 10 we have tabulated the common α for varying c and β_i and α_i.

In order to get some impression of how the dimension r affects α and b, we provide Table 11. The most striking feature of the table is the rapid increase in α for negative c's as β increases. This is important since we have to take a negative c in order to get uncorrelated components.

				Table of b for r=2 varying c and β_i (any α_i)					
$c =$	-0.3	-0.2	-0.1	0.0	0.1	0.2	0.3	0.4	0.5
$\beta_i = 0$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	1.43	1.25	1.11	1.00	0.91	0.83	0.77	0.71	0.67
2	2.86	2.50	2.22	2.00	1.82	1.67	1.54	1.43	1.33
3	4.29	3.75	3.33	3.00	2.73	2.50	2.31	2.14	2.00
4	5.71	5.00	4.44	4.00	3.64	3.33	3.08	2.86	2.67
5	7.14	6.25	4.56	5.00	4.55	4.17	3.85	3.57	3.33
6	8.57	7.50	6.67	6.00	5.46	5.00	4.62	4.29	4.00

Table 9: Common skewness parameter b

Table of α for r=2 varying c, α_i and β_i									
$c =$	-0.4	-0.3	-0.2	-0.1	0.0	0.1	0.2	0.3	0.4
$\alpha_i = 2, \ \beta_i = 0$	2.09	2.05	2.02	2.01	2.00	2.01	2.02	2.05	2.09
1	2.63	2.48	2.37	2.29	2.24	2.20	2.18	2.18	2.20
$\alpha_i = 4, \beta_i = 0$	4.18	4.10	4.04	4.01	4.00	4.01	4.04	4.10	4.18
1	4.47	4.33	4.23	4.16	4.12	4.11	4.12	4.16	4.23
2	5.26	4.96	4.74	4.58	4.47	4.40	4.37	4.36	4.40
3	6.35	5.86	5.49	5.21	5.00	4.85	4.74	4.68	4.65
$\alpha_i = 8, \beta_i = 0$	8.36	8.19	8.08	8.02	8.00	8.02	8.08	8.19	8.36
2	9.45	8.65	8.45	8.32	8.25	8.22	8.25	8.33	8.47
4	10.52	9.91	9.48	9.16	8.94	9.80	8.73	8.73	8.80
6	12.71	11.71	10.97	10.42	10.00	9.69	9.48	9.35	9.31

Table 10: Common tail parameter α

Table of α for $\alpha_i = 4$ varying r and c and β_i								
$c =$	-0.2	-0.1	0.0	0.1	0.2	0.3	0.4	0.5
$r = 3, \beta_i = 0$	4.10	4.02	4.00	4.02	4.07	4.17	4.30	4.49
1	4.58	4.35	4.24	4.20	4.22	4.28	4.39	4.56
2	5.80	5.22	4.90	4.71	4.62	4.60	4.65	4.76
3	7.39	6.42	5.83	5.46	5.22	5.09	5.04	5.08
$r = 5, \ \beta_i = 0$	4.37	4.05	4.00	4.03	4.12	4.26	4.46	4.73
1	6.91	4.89	4.47	4.35	4.35	4.43	4.59	4.82
2	11.56	6.82	5.66	5.17	4.96	4.90	4.95	5.11
3	16.63	9.17	7.21	6.31	5.83	4.59	5.49	5.55
$r = 10, \beta_i = 0$	-	4.30	4.00	4.06	4.20	4.40	4.66	5.02
1	-	11.53	5.00	4.57	4.53	4.63	4.83	5.14
2	-	21.82	7.21	5.84	5.38	5.25	5.30	5.51
3	-	32.37	9.850	7.49	6.57	6.16	6.01	6.06

Table 11: Common tail parameter α

4 Acknowledgment

References

Barndorff-Nielsen, O.E. (1997): Normal inverse Gaussian processes and stochastic volatility modelling. *Scandinavian Journal of Statistics vol. 24,* 1-13.

Barndorff-Nielsen, O.E. and N. Shephard (2001): Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in Financial economics. *J. Royal Statist. Soc. B, 63 ,* 167-241.

Bauer, C. (2000): Value at Risk using hyperbolic distributions. *Journal of Economics and Business, 52,* 455-467.

Blaesild, P. and M.K.Sörensen (1992): 'hyp' - A computer program for analyzing data by means of the hyperbolic distribution. *Research Report No 248, Department of Theoretical Statistics, University of Aarhus, Denmark.*

Davison, A.C. and D.V.Hinkley (1999): *Bootstrap Methods and their Application. Cambridge University Press, Cambridge.*

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

Eberlein, E. , A. Ehret, O. Lübke, F. Özkan, K. Prause, S. Raible, R. Wirth and M. Wiesendorfer Zahn (1998): Freiburg Financial Data Tools. *Freiburg: Mathematische Stokastik, Universität Freiburg.*

Karlis, D. (2002): An EM type algorithm for maximum likelihood estimation for the Normal Inverse Gaussian distribution, *Statistics and Probability Letters*, 57, 43-52

Karlis, D. and J. Lillestöl (2002): Bayesian estimation of NIG-models via Markov chain Monte Carlo methods. *Preprint*

Lillestöl, J. (1998): Fat and skew? Can NIG cure? On the prospects of using the Normal inverse Gaussian distribution in finance, *Discussion paper 1998/11, Department of Finance and Management Science, The Norwegian School of Economics and Business Administration.*

Lillestöl, J. (2000): Risk analysis and the NIG distribution. *The Journal of Risk, 2 ,* 41-56.

Lillestöl, J. (2001): Bayesian Estimation of NIG-parameters by Markov chain Monte Carlo Methods, *Discussion paper 2001/3, Department of Finance and Management Science, The Norwegian School of Economics and Business Administration.* Earlier version as *Discussion paper 112/2000, Sonderforschungsbereich 373, Humboldt Universität zu Berlin*

Nelsen, R.R. (1999): *An introduction to copulas. Springer, New York.*

Prause, K. (1999a): The generalized hyperbolic model: Estimation, financial derivatives and risk measures. *Dissertation Albert-Ludwigs-Universität Freiburg.*

Prause, K. (1999b): How to use NIG laws to measure market risk. *FDM Preprint 65, University of Freiburg.*

Stehle, R. and O. Grewe (2001): The long-run performance of German stock mutual funds. *Draft, Humboldt Universität zu Berlin.*

Stute, W., W.G. Manteiga and M.P. Quindimil (1993): Bootstrap based goodness-of-fit tests. *Metrika, 40,* 243-256.

Venter, J.H. and P.J. de Jongh (2002): Risk estimation using the Normal Inverse Gaussian distribution. *The Journal of Risk, 4,* 1-23.