Fat and skew: Can NIG cure? On the prospects of using the Normal Inverse Gaussian distribution in finance

Jostein Lillestøl
Institute of Finance and Management Science
The Norwegian School of Economics and Business Adminstration
Bergen, Norway

March 6, 1998
Revised June 15, 1998

Abstract

This paper explores the possibility of using the Normal Inverse Gaussian (NIG) distribution introduced by Barndorff-Nielsen (1997) in various problem areas in finance where distributions often are found to be non-normal due to skewness and fat tails. More specificly we discuss problems of risk analysis and portfolio choice in a NIG context. We also briefly look into some aspects of NIG-modelling and estimation, but numerics and empirics will be pursued elsewhere.

E-mail: jostein.lillestol@nhh.no

1 Background

In empirical finance it is frequently observed that asset returns have distributions with fat tails, and they are often skew. [1] Moreover certain nonlinear dependence structures occur. Other features are observed as well, depending on the context. In order to model financial data we need a repertoire of distributions and modelling techniques which are able to represent these stylized facts, and which are at the same time analytically tractable. The literature is by now immense, within three interconnected areas:

1. Distributions (stable Paretian, generalized beta of second kind etc)
2. Time series model (GARCH, SVM etc)
3. Process models (Diffusion and jump processes etc)

Recently a new family of distributions named normal inverse Gaussian (NIG) is brought to the attention of workers in empirical finance by Barndorff-Nielsen. Research so far is promising. It fits data very well, is analytically tractable, and may be basis for (state space) time series modelling and process modelling as well, see Aase (1997).

In the next section we summarize some of the features of the NIG family, with emphasis on properties that can be useful in the financial context. In the following sections we develop some results which may be of use in risk analysis.

2 The normal inverse Gaussian distribution

The normal inverse Gaussian distribution is characterized by 4 parameters $(\alpha, \beta, \mu, \delta)$, where α is related to steepness, β to symmetry, and μ and δ are related to location and scale respectively, for short referred to below as the location and scale parameter. The distribution arises as the marginal distribution of X in (X, Z) where [2]

$$\begin{aligned} X \mid Z = z &\sim N(\mu + \beta z, z) \\ Z &\sim IG(\delta, \sqrt{\alpha^2 - \beta^2}) \quad \text{where} \quad 0 \leq \mid \beta \mid < \alpha \end{aligned}$$

[1] See for instance the scientific review paper on Value at Risk by Duffie (1997) and worries from practice in $CreditMetrics^{TM}$.

[2] Here $N(\theta, z)$ is the normal distribution with variance z and $IG(\delta, \gamma)$ is the inverse Gaussian distribution with density given in Johnson et. al. (1995).

Its moment generating function is

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

from which we can derive (let $\gamma = \sqrt{\alpha^2 - \beta^2}$ for short)

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

It is well worth noting that

$$\mu = EX - \beta(1 - (\frac{\beta}{\alpha})^2) varX$$

It is also seen that a sum of independent NIG-variates with common α and β, but different location and scale parameters, is itself NIG obtained by summing the location and scale parameters and keeping the others fixed. We illustrate the feasible (α, β)-combinations in Figure 1.

We see that $\beta = 0$ gives symmetric distributions where

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

The Cauchy distribution is obtained [3] for $\alpha = 0$ and the normal distribution is obtained as $\alpha \to \infty$. The latter is seen by letting $\alpha \to \infty$ and $\delta \to \infty$ so that $\delta/\alpha \to \sigma^2$. In fact there is no need for β to be zero to achieve a normality limit. Even when β itself follows a limiting process, we get a normal limit as long as β tends to a finite limit β_∞. The normal limit then corresponds to $N(\mu + \beta_\infty \cdot \sigma^2, \sigma^2)$.

The $NIG(\alpha, \beta, \mu, \delta)$ density is given by

$$g(x; \alpha, \beta, \mu, \delta) = a(\alpha, \beta, \mu, \delta) q(\frac{x - \mu}{\delta})^{-1} K_1(\delta\alpha q(\frac{x - \mu}{\delta})) e^{\beta x}$$

[3] Seen by looking at the characteristic function $M_X(it)$.

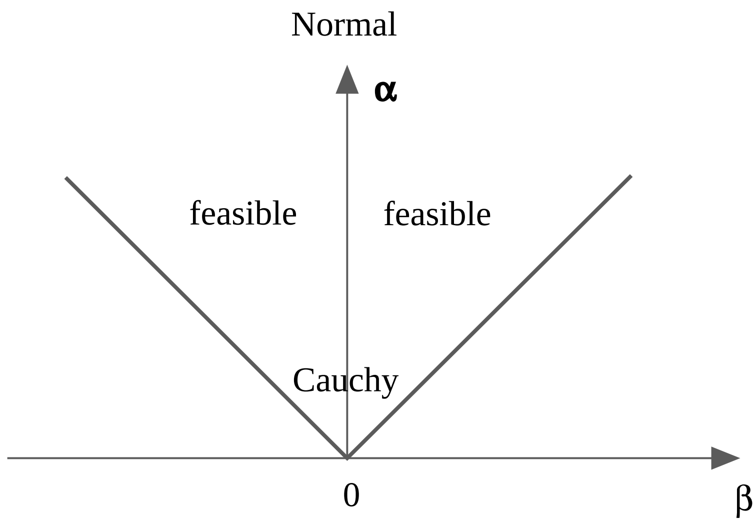

Figure 1: The feasible (α, β) map.

where $q(x) = \sqrt{1+x^2}$ and $a(\alpha, \beta, \mu, \delta) = \pi^{-1}\alpha exp(\delta\sqrt{\alpha^2 - \beta^2} - \beta\mu)$ and K_1 is the modified Bessel function of second kind (by some called third order) and index 1. [4] The distribution has semiheavy tails, ie.

$$g(x; \alpha, \beta, \mu, \delta) \sim const \mid x \mid^{-3/2} e^{-\alpha|x|+\beta x} \quad \text{as} \quad x \to \pm\infty$$

A homogeneous (i.e stationary increment) Levy process (i.e. continuous in probability) X_t with $NIG(\alpha, \beta, \mu, \delta)$ marginals can be defined by

$$\begin{aligned} M_t(u; \alpha, \beta, \mu, \delta) &= M(u; \alpha, \beta, \mu, \delta)^t \\ &= M(u; \alpha, \beta, \mu t, \delta t) \end{aligned}$$

and may be replaced by a random time change of Brownian motion, that is

$$X_t = \mu_t + B_{Z_t}$$

where B_t is Brownian motion with drift β and diffusion coefficient 1 and Z_t is a homegenous Levy process with $IG(\delta, \sqrt{\alpha^2 - \beta^2})$ marginals.

Barndorff-Nielsen has shown that X_t is a superposition of weighted independent Poisson processes with small jumps dominating. He has also explored a class of processes with NIG marginals and IG-marginals of Ornstein-Uhlenbeck type with background driving process being a homogeneous Levy

[4] See Ambramowitz & Stegun (1972).

process. This may be useful for modelling in continuous time, say of financial processes. However the likelihood analysis of discrete observations from the processes is challenging, as it is for the common stochastic differential.

It is fairly easy to simulate NIG-variates, see Appendix.

3 NIG-returns and its convenience utility

The trade-off between high return and risk is important in finance. So called mean-variance analysis has its theoretical basis in the case of normal variates and/or quadratic utility function, or special cases of matching utility functions and distributions (e.g. log utility and lognormal distribution). Knowing that these assumptions are unrealistic it is still widely use for convenience, see Levy & Markowitz (1979) and Kroll, Levy & Markowitz (1984).

If we model returns by the NIG-family of distributions a convenient alternative may be to start from a utility function of constant absolute risk aversion, a desirable property according to Arrow (1971). This means that

$$U(X) = 1 - exp(-\lambda X)$$

The expected utility of X being $NIG(\alpha, \beta, \mu, \delta)$ becomes

$$\begin{aligned} EU(X) &= 1 - Eexp(-\lambda X) \\ &= 1 - exp(-\lambda\mu + \delta(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta - \lambda)^2})) \end{aligned}$$

where we have to add the restriction $\alpha + \beta > \lambda$. We see that NIG-prospects may be ranked by their value of the expression

$$H(\lambda, \alpha, \beta, \mu, \delta) = \lambda\mu - \delta(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta - \lambda)^2})$$

In the case of no skewness $\beta = 0$ this is reduced to

$$H(\lambda, \alpha, 0, \mu, \delta) = \lambda\mu - \delta(\alpha - \sqrt{\alpha^2 - \lambda^2})$$

In the normal case the corresponding well known formula is [5]

$$H = \lambda\mu - \frac{1}{2}\lambda^2\sigma^2 = \lambda(\mu - \frac{1}{2}\lambda\sigma^2)$$

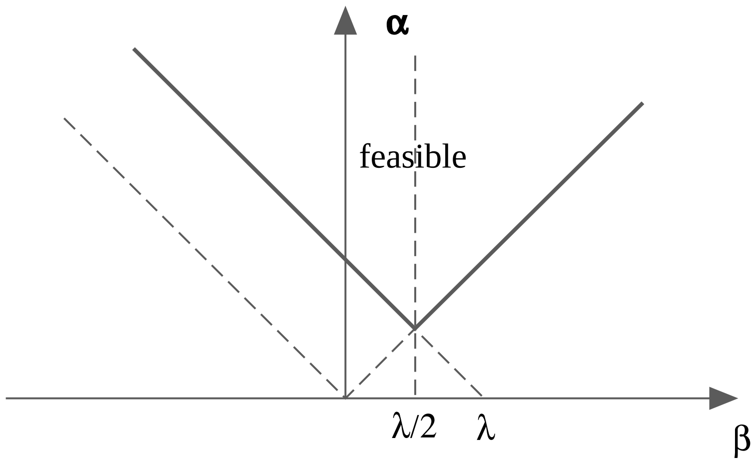

Figure 2: Feasible combinations of (α, β)

The parameter restrictions now in effect are illustrated in Figure 2.

Roughly [6] the restriction can be interpreted as those return distributions for which increase in return level can compensate increased risk for the given level of risk aversion We see that increased risk aversion requires more steepness unless the distribution has a minimum asymmetry towards longer right tails.

A better understanding of how H depends on the various parameters involved is obtained by deriving a first order approximation. We get

$$\begin{aligned} H &\approx \lambda(\mu + \delta\frac{\beta}{\gamma}) - \lambda^2\frac{1}{2}\frac{\delta}{\gamma} \\ &= \lambda(EX - \lambda\frac{1}{2}(1 - (\frac{\beta}{\alpha})^2)varX) \end{aligned}$$

For the normal case the approximate formula is exact. We see that approximately the ranking amounts to a trade-off between expectation and variance, the latter having a correction depending on the steepness parameter α and the symmetry parameter β. For a given variance a higher expectation is required to compensate a smaller α and a higher β. However a discussion of the dependence on the parameters based on the latter for-

[5]The certainty equivalent $\mu - \frac{1}{2}\lambda\sigma^2$ applies in a total "all or nothing" context. In a market context, the first order marginal condition implies the certainty equivalent $\mu - \lambda\sigma^2$.

[6]Roughly since the expectation and variance formulae are somewhat more involved.

mula may be misleading, since the expectation and variance depend on these parameters.

An interesting borderline case is $\lambda = 2\beta$ (dotted above) where

$$H(2\beta, \alpha, \beta, \mu, \delta) = \lambda \cdot \mu = 2\beta\mu$$

i.e. the utility is not affected by δ and α at all. This may seem odd, but recalling the formula for μ in terms of expectation and variance we again get the approximate formula for H above as an exact formula, so that volatility does matter.

We can now examine the indifference curve relationships. Let the curves be indexed by the level h. In the normal case we have

$$\mu = \frac{h}{\lambda} + \frac{1}{2}\sigma^2\lambda$$

i.e. straight lines in the (μ, σ^2) plane with slope increasing with the risk avension, or parabolas in the (μ, σ) plane. In the close to normal (symmetric) case the indifference curves are approximated by

$$\mu = \frac{h}{\lambda} + \frac{1}{2}\frac{\delta}{\alpha}\lambda$$

i.e. increased volatility in the δ sense is compensated by increased steepness. For skew distributions in the neighborhood of the normal, the indifference curves are close to

$$\mu = \frac{h}{\lambda} + \frac{1}{2}\frac{\delta}{\alpha}(\lambda - 2\beta)$$

i.e. increased volatility in the δ sense is compensated by increased positive skewness.

4 Independent portfolio NIG-returns

We will be interested in r joint returns $X = (X_1, X_2, \ldots, X_r)$ and the return $Y = w'X$ on a portfolio $w = (w_1, w_2, \ldots, w_r)$. In the case of independent NIG-returns with common α and β parameter, and equal weights we have that

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

This means that

$$Skewness = 3 \cdot \frac{\beta}{\alpha} \cdot \frac{1}{(\bar{\delta}\gamma)^{1/2}} \cdot \frac{1}{\sqrt{r}}$$

and furthermore

$$\bar{\mu} = \frac{h}{\lambda} + \frac{1}{2}\frac{\bar{\delta}}{r\alpha}(\lambda - 2r\beta)$$

In the case of non-equal weights we do not have exact NIG. Hopefully we catch the main features by approximating as follows [7]

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

where

$$\begin{aligned}
\mu_w &= \sum_i w_i \mu_i \\
\delta_w &= \sum_i w_i \delta_i \\
\alpha_w &= \frac{\sum_i w_i \delta_i}{\sum_i w_i^2 \delta_i \alpha_i^{-1}} \\
\beta_w &= \frac{\sum_i w_i \delta_i \beta_i \alpha_i^{-1}}{\sum_i w_i^2 \delta_i \alpha_i^{-1}}
\end{aligned}$$

In the case of equal α's and β's we have the simpler formulas

$$\begin{aligned}
\alpha_w &= \frac{\sum_i w_i \delta_i}{\sum_i w_i^2 \delta_i} \cdot \alpha \\
\beta_w &= \frac{\sum_i w_i \delta_i}{\sum_i w_i^2 \delta_i} \cdot \beta
\end{aligned}$$

It will be of interest to investigate how well this approximate the exact distribution. 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.

[7] The approximation is obtained by matching terms (admittedly somewhat ad hoc) in the expressions for the expectation and the exponent of the momentgenerating function.

5 Multivariate returns and portfolios

In order to model vector correlated returns $\mathbf{X}$ we may turn to the family of multivariate NIG $(\alpha\,, \boldsymbol{\beta}\,, \delta\,, \boldsymbol{\mu}\,, \boldsymbol{\Phi})$ distributions 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 properties of marginalization, conditioning and linear transformation are given in Blæsild (1981). The marginal and linear combinations are both univariate NIG. However, we note that independent univariate NIG-variates are jointly not multivariate NIG in the sense above!

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

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

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 $\phi_i^2 = \phi_{ii}$)

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

Note that the alfa-scalars here do not correspond to an alfa-parameter common to all the marginals. We see that the marginal α_is are affected jointly by

$\boldsymbol{\beta}$ and $\boldsymbol{\Phi}$. This makes it difficult to interpret parameters and a bit awkward to establish a joint model specification from given marginal specifications.

The covariance matrix of X is [8]

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

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. Some insight is gained by looking at special cases. If $\boldsymbol{\Phi}$ is the unity matrix we get

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

Consequently Σ_{ij}'s are affected by $\beta_j\beta_k$ and Σ is diagonal in this case only if the β's are zero. Then $\boldsymbol{\Sigma} = \frac{\delta}{\alpha}\mathbf{I}$, which is in agreement with the limiting case of $\delta/\alpha \to \sigma^2$. If $\boldsymbol{\beta}'\Phi\boldsymbol{\beta}$ is negligible compared to α^2

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

because then the only omitted term is likely to be negligible as well. Then Φ diagonal corresponds to approximate uncorrelated returns.

It is of interest to explore how the parametes change from the individual marginal to the resulting weighted combination. In general we have

$$\begin{aligned} D &= \frac{\delta_w}{\delta_i} = \frac{\phi_w}{\phi_i} = F \\ G &= \frac{\gamma_w}{\gamma_i} = (\frac{\phi_w}{\phi_i})^{-1} = F^{-1} \\ C &= \frac{Skewness}{Skewness_i} = \frac{\beta_w/\beta_i}{\alpha_w/\alpha_i} = \frac{B}{A} \end{aligned}$$

where the definitions of A, B, C, D, F and G are self-explanatory. We see that the change in skewness depends on the change in α and β alone. Moreover it follows that the change in standard deviation, here denoted by S, becomes $S = AF^2$, which is equal to F in the symmetric case.

In finance analysts are used to so-called (μ, σ) maps, among others to illustrate efficient frontiers. It would be of interest to see if something similar

[8] The covariances is most easily obtained from the cumulant generating function.

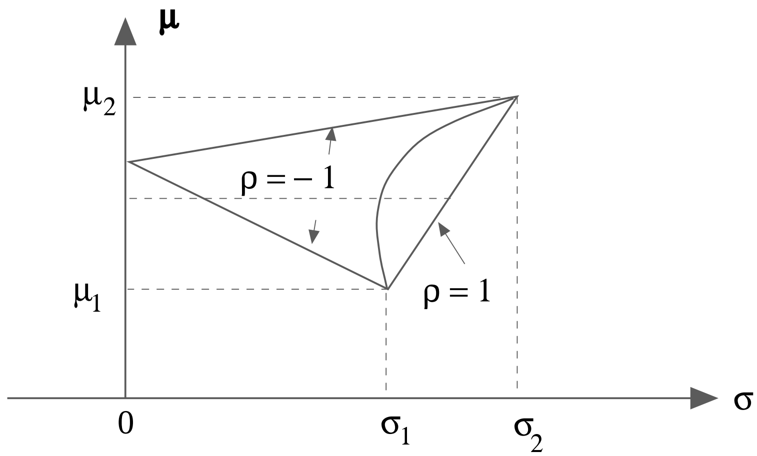

Figure 3: A (μ, σ) map for given δ and α

pertains with NIG-parameters. Some insight may be gained by looking at the bivariate case. [9]

$$\Phi = \begin{pmatrix} \phi_1^2 & \phi_1\phi_2\rho \\ \phi_1\phi_2\rho & \phi_2^2 \end{pmatrix}$$

where again $\phi_i^2 = \phi_{ii}$ and ρ is introduced in order to mimic variance and correlation. However with the convention that $det\Phi$ is equal to 1 we have $\phi_1^2\phi_2^2 = (1-\rho^2)^{-1}$. We now get

$$\phi_w^2 = w_1^2\phi_1^2 + w_2^2\phi_2^2 + 2w_1w_2\phi_1\phi_2\rho$$

and note that the cases $\rho = \pm 1$ lead to a complete square.

In the symmetric case when $\boldsymbol{\beta} = 0$ we have $\beta_w = 0$, $\alpha_w = \gamma_w = \phi_w^{-1}\alpha$ and $\delta_w = \phi_w\delta$. Now $EY = \mu_w = w_1\mu_1 + w_2\mu_2$ and $varY = \phi_w^2\frac{\delta}{\alpha}$ while $varX_i = \phi_i^2\frac{\delta}{\alpha}$. If we let $\sigma = \phi(\frac{\delta}{\alpha})^{1/2}$ we see that a (μ, σ) map, for fixed δ and α, will have essentially the same features as the common (μ, σ) map in terms of expectation and standard deviation, see Figure 3. However our parameters correspond with the usual ones only in the symmetric case. In order to take skewness into account, we could modify the required μ by an additive factor. The results of the previous section suggest the factor $\beta \cdot \sigma^2$.

[9] Admittedly the use of greek letters here conflicts with the use of beta and gamma in the portfolio literature.

If we stick to the common (μ,σ) map in terms of expectation and standard deviation, and want to take into account the skewness, we may scale up the required expectation for given standard deviation by an additive factor. The preceeding section suggests that as a first approximation we may use the factor $\lambda\frac{1}{2}(\frac{\beta}{\alpha})^2\sigma^2$. This depends on λ, but a possible "parameter free" choice is $\lambda = 2\cdot\beta$. It remains to be seen how this works and whether it is useful at all.

It would clearly be of interest to see how well the approach of this section match with valuation schemes based on economic equilibrium considerations, for instance extensions of CAPM to accomodate skewness, see Kraus & Litzenberger (1976), (1983).

6 Portfolio choice and equilibrium considerations

We will consider equilibrium conditions for a portfolio of multivariate NIG-returns in conjunction with a riskfree asset using the exponential utility above. onsider first the case of an individual investor. Let the initial wealth be W_0 and final wealth be

$$W = W_0(1 + \sum_{i=0}^{r} w_i R_i)$$

where

$$\begin{aligned} R_0 &= \text{return on the riskfree asset} \\ R_i &= \text{returns on risky asset no.i} \end{aligned}$$

The problem of maximizing the expected utility $EU(W) = 1 - Eexp(-\lambda W)$ subject to the budget restriction is now seen to be equivalent to maximizing

$$\lambda(w_0R_0 + \mathbf{w}'\boldsymbol{\mu}) - \delta((\alpha^2 - \boldsymbol{\beta}'\boldsymbol{\Phi}\boldsymbol{\beta})^{1/2} - (\alpha^2 - (\boldsymbol{\beta} - \lambda\mathbf{w})'\boldsymbol{\Phi}(\boldsymbol{\beta} - \lambda\mathbf{w}))^{1/2})$$

subject to

$$w_0 + \mathbf{w}'\mathbf{e} = 1$$

The Lagrangian becomes (with θ being the Lagrange multiplier)

$$L = \lambda(w_0R_0+\mathbf{w}'\boldsymbol{\mu}) - \delta((\alpha^2-\boldsymbol{\beta}'\boldsymbol{\Phi}\boldsymbol{\beta})^{1/2} - (\alpha^2-(\boldsymbol{\beta}-\lambda\mathbf{w})'\boldsymbol{\Phi}(\boldsymbol{\beta}-\lambda\mathbf{w}))^{1/2}) - \theta(w_0+\mathbf{w}'\mathbf{e}-1)$$

By putting the expressions obtained by differentiating with respect to the w_i's equal to zero we get that $\lambda R_0 = \theta$ and

$$\boldsymbol{\mu} - R_0\mathbf{e} = \delta(\alpha^2 - (\boldsymbol{\beta} - \lambda\mathbf{w})'\boldsymbol{\Phi}(\boldsymbol{\beta} - \lambda\mathbf{w}))^{-1/2}\boldsymbol{\Phi}(\lambda\mathbf{w} - \boldsymbol{\beta})$$

If we introduce the shorthand $\boldsymbol{\psi} = \boldsymbol{\Phi}(\lambda\mathbf{w} - \boldsymbol{\beta})$, with components ψ_i we get

$$\frac{\mu_i - R_0}{\mu_j - R_0} = \frac{\psi_i}{\psi_j}$$

This ratio does not depend on α and δ. Note also that we can write

$$\mu_i - R_0 = \frac{\delta^2}{\delta_w}(\alpha_w^2 - (\beta_w - \lambda)^2)^{-1/2}\psi_i$$

Here the subscripts w refer to the portfolio of the risky assets with formulae given in the previous section (but now with sum of weights one minus the fraction invested in the riskfree asset). The corresponding terms obtained by dividing the weights by $1 - w_0$ will be denoted by subscript P, and consequently $\mu_w = \mathbf{w}'\boldsymbol{\mu} = (1 - w_0)\mu_P$. We now write

$$\mu_R = w_0 R_0 + \mathbf{w}'\boldsymbol{\mu} = w_0 R_0 + (1 - w_0)\mu_P$$

and note that $\mu_R - R_0 = (1 - w_0)(\mu_P - R_0)$. From the above we now get

$$\frac{\mu_i - R_0}{\mu_R - R_0} = \frac{\psi_i}{\psi_w}$$

where $\psi_w = \mathbf{w}'\boldsymbol{\Phi}(\lambda\mathbf{w} - \boldsymbol{\beta})$.

We also take a brief look at the market equilibrium conditions for the case of investors having identical probability beliefs. The exponential utility then leads to identical compositions of risk portfolios [10].

For the market to clear, the optimal proportions of risk assets for each investor must be those of the market risk asset portfolio m. This leads to

$$\mu_i - R_0 = \frac{\psi_i}{\psi_m}(\mu_m - R_0)$$

where the ψ's are given by the formulae above, but with components of m summing to one and λ replaced by $(1 - w_0)\lambda$. To characterize the solution we may just leave out $(1 - w_0)$.

The above formulae parallels the classic ones, but recall again that μ's are not expectations in the skew case. The formulae may be explored from different viewpoints, and we will only make some brief comments here. In

[10] Cass & Stiglitz have shown that a neccessary and sufficient condition for this is that each invetors risk tolerance is a linear function of wealth, that is $-U_i'/U_i'' = a_i + bW_i$ with the same cautiousness b for all investors, in our case $b = 0$.

the case of an equally weighted market portfolio of r assets having all components of β equal to b, we get

$$\frac{\psi_i}{\psi_w} = \frac{r\sum_j \phi_{ij}}{\sum_{ij}\phi_{ij}} = \frac{r\phi_i^2\beta_i}{\sum_i \phi_i^2\beta_i}$$

which does not depend on λ and b at all. Note however that the individual skewnesses may differ through differing ϕ's.

7 Exchangeable returns

Of some interest (at least for exploring the aspects of NIG modelling) is the exchangeable case where $\beta = (b, b, \ldots, b)$ and

$$\Phi = \begin{pmatrix} d & cd & cd & \cdots & cd \\ cd & d & cd & \cdots & cd \\ \vdots & \vdots & & & \vdots \\ cd & cd & cd & \cdots & d \end{pmatrix}$$

The matrix can be written $\mathbf{\Phi} = d(c\mathbf{E} + (1-c)\mathbf{I})$ where $\mathbf{I}$ is the rxr identity matrix and $\mathbf{E}$ is the rxr matrix of ones. In order to have $det\mathbf{\Phi} = 1$ we must have

$$det\Phi = d^r(1-c)^{r-1}(1 + c(r-1)) = 1$$

We see that a neccessary requirement is $c > -1/(r-1)$. For given c we then have

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

which tends to

$$d = (1-c)^{-1}exp(-\frac{c}{1-c}) \quad \text{as} \quad r \to \infty.$$

In the exchangeable case

$$\boldsymbol{\beta}'\mathbf{\Phi}\boldsymbol{\beta} = rpdb^2$$

$$\mathbf{\Phi}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{\Phi} = p^2d^2b^2\mathbf{E}$$

where $p = 1 + (r-1)c$. If we assume equal weights we have

$$\begin{aligned} \phi_w^2 = \mathbf{w}'\mathbf{\Phi}\mathbf{w} &= pdr^{-1} \\ \mathbf{w}'\mathbf{\Phi}\boldsymbol{\beta} &= pdb \\ \beta_w &= br \\ \beta_i &= bp \end{aligned}$$

Moreover set $dp^2b^2\alpha^{-2} = t^{-1}$. Then

$$\begin{aligned} \alpha_w &= d^{-1/2}(rp^{-1})^{1/2}\alpha \\ \alpha_i &= d^{-1/2}(1 + t^{-1}(1 - rp^{-1}))^{1/2}\alpha \end{aligned}$$

We see that [11]

$$\begin{aligned} B &= \frac{\beta_w}{\beta_i} = rp^{-1} > 1 \\ A &= \frac{\alpha_w}{\alpha_i} = (rp^{-1})^{1/2}(1 + t^{-1}(1 - rp^{-1}))^{-1/2} > 1 \\ C &= \frac{Skewness}{Skewness_i} = (rp^{-1}(1 + t^{-1}(1 - rp^{-1}))^{1/2} > 1 \end{aligned}$$

Moreover $G = B^{1/2}$, $F = B^{-1/2}$ and that now $S = AF^2 = A/B$. The inequalities for B and C came as a surprise. By letting $r \to \infty$ we get

$$\begin{aligned} B &\to c^{-1} \\ A &\to c^{-1}(c + (1 - c)(\frac{\beta_i}{\alpha_i})^2)^{1/2} \\ C &\to (c + (1 - c)(\frac{\beta_i}{\alpha_i})^2)^{-1/2} \\ S &\to (c + (1 - c)(\frac{\beta_i}{\alpha_i})^2)^{1/2} \end{aligned}$$

It is at first sight somewhat surprising that A does not tend to infinity and C does not tend to zero. However we knew that S would not tend to zero for correlation.

It is of interest to explore how the natural parameters of the joint distribution are determined from the natural marginal parameters and the covariance structure. In general this is not easy, but some insight is gained in the exchangeable case.

With r and c given, p and d is determined. The common individual beta determines $b = p^{-1}\beta_i$, which in turn determines $\beta_w = br$ without knowledge of other parameters. The common individual alpha determines

$$\alpha = d^{1/2}(\alpha_i^2 - p^2b^2(1 - rp^{-1}))^{1/2}$$

which in turn determines $\alpha_w = d^{-1/2}(rp^{-1}))^{1/2}\alpha$.

[11] In case of A this is seen by noting that $x = 1$ and $x = t$ solves the equation $x(1 + t^{-1}(1 - x)) = 1$ and looking at the cases $t > 1$ and $t < 1$. Similarly for C.

Going back to the expression for the covariance matrix, we see that the off-diagonal elements in the exchangeable case are (here $\gamma = (\alpha^2 - rpdb^2)^{1/2}$)

$$\delta\gamma^{-1}(cd + \gamma^{-2}b^2d^2p^2)$$

while the diagonal elements are

$$\delta\gamma^{-1}(d + \gamma^{-2}b^2d^2p^2)$$

The correlations therefore become

$$\frac{c + \gamma^{-2}b^2dp^2}{1 + \gamma^{-2}b^2dp^2}$$

For $c = 0$, that is diagonal Φ, the correlation is positive. Zero correlations requires negative c. Note however that this does not correspond to independence.

8 Estimation of NIG-parameters

The estimation of parameters of the NIG distribution from sampled data may be based on the likelihood-function. The expression becomes fairly complicated, and the numerical and programming challenges are demanding, but may be handled, see Blæsild & Sørensen (1992) and later extensions. Another possibility is to use the the method of moments, which here amounts to equatiing the expressions in section 2 for the mean, variance, skewnwss and kurtosis in section 2 to their empirical counterparts. We then get four equations which may be solved for the four parameters, in fact exact expressions are easily obtained.

Given its simple expression, it seems worthwhile to explore estimation schemes based on the momentgenerating function, see for instance Epps, Singleton & Pulley (1982). One possibility is the generalized method of moments (GMM), which is adaptable to numerous different situations, and well known to workers in financial econometrics.

We consider here only the case of n independent NIG-variates $X_1, X_2, \ldots, X_n$. Our moment equations will be

$$\frac{1}{n}\sum_{i=1}^{n} e^{uX_i} = exp(u\mu + \delta(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + u)^2}))$$

Now let the logarithm of the left hand side be denoted by $v(u)$. By choosing four different $u = u_i (i = 1, 2, 3, 4)$ and letting $v_i = v(u_i)$ we get the following four estimating equations for the four unknowns:

$$u_i\mu + \delta(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + u_i)^2}) = v_i \quad i = 1, 2, 3, 4$$

These equations may be written on the following generic form:

$$1 + au - bv = r \quad \text{with} \quad r = r(c, d) = \sqrt{1 - cu - du^2}$$

and where the coefficients a, b, c, d in terms of NIG-parameters are $a = \mu/\delta\gamma$, $b = 1/\gamma\delta$, $c = 2\beta/\gamma^2$, $d = 1/\gamma^2$ and To prepare for the numerical solution, it may be worthwhile to reduce the four equations to two by eliminating a and b. If we introduce the following shorthand notation

$$\begin{aligned} p_{ij} &= 1 - \frac{u_i v_j}{u_j v_i} \\ q_{ij} &= 1 - r_i - \frac{u_i}{u_j}(1 - r_j) \end{aligned}$$

we get after some simple algebra

$$\frac{q_{12}}{p_{12}} = \frac{q_{13}}{p_{13}} = \frac{q_{14}}{p_{14}} = v_1 b = 1 - r_1 + au_1$$

Since $q_{ij} = q_{ij}(c, d)$, the two first equalities may be used to solve for c and d. The last two equalities give a and b after resubstitution of c and d [12]. Given a, b, c, d, we can now obtain the estimates of the NIG-parameters by substitution in $\mu = a/b$, $\delta = b^{-1}d^{1/2}$, $\gamma = d^{-1/2}$, $\beta = \frac{1}{2}cd^{-1}$ and finally $\alpha = \sqrt{\gamma^2 + \beta^2}$. The numerics and estimation on simulated and real data will be pursued elsewhere, in order to see how well this procedure is a viable alternative to using the likelihood directly.

[12] Note that the subscript in the last two expressions is the common first index of the first three, and that we alternatively could have provided identities in terms of 2, 3 or 4 as the first subscript.

Appendiks: Simulation of NIG-variates

Let V be a chisquare variate with 1 degree of freedom and compute the roots with respect to Z of

$$V = \frac{(\gamma Z - \delta)^2}{Z}$$

They are given by

$$Z = \frac{\delta}{\gamma} + \frac{1}{2\gamma^2}(V \pm \sqrt{V^2 + 4\gamma\delta V})$$

Let Z_1 and Z_2 be the minus and plus root respectively, and note that $Z_2 = \delta^2/Z_1$.

Let

$$\begin{aligned} Z &= Z_1 \quad \text{with probability} \quad \frac{\delta}{\delta + Z_1} \\ &= Z_2 \quad \text{with probability} \quad \frac{Z_1}{\delta + Z_1} = \frac{\delta}{\delta + Z_2} \end{aligned}$$

Then Z is $IG(\delta, \gamma)$ [13] and consequently we get a $NIG(\alpha, \beta, \mu, \delta)$ variate by taking $\gamma = \sqrt{\alpha^2 - \beta^2}$ and generate a U being $N(0, 1)$ and then compute

$$X = \mu + \beta Z + \sqrt{Z} \cdot U$$

[13] See Michael et. al. (1976)

References

Aase, K. (1995) An equilibrium approach to derivative securities: Stochastic volatility and survival. *Working paper no. 2/96 Institute of Finance and Management Science, The Norwegian School of Economics and Business Administration.*

Abramowitz, M. & Stegun, I.A. (1972) Handbook of Mathematical Functions. Dover.

Alexander, C. (1997) Evaluating the latest empirical reserch on incorporating fat-tails and asymmetry into your risk measurement. Notes from talk given at RiskForum Geneva 1997.

Arrow, K. (1971) Essays on the theory of risk bearing. Markham Publ. Chicago.

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

Blæsild, P. (1981) The two-dimensional hyperbolic distribution and related distributions, with applications to Johansens's bean data. *Biometrika 68, 251-253.*

Blæsild, P. & Sørensen, M.K. (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.*

Duffie, D. & Pan, J. (1997) An overview of Value at Risk. *Journal of Derivatives, 4, 7-49.*

Epps, T.W., Singleton, K.J., & Pulley, L.B. (1982) A test of separate families of distributions based on the empirical moment generating function. *Biometrika, 69, 391-399.*

Hansen, L. (1982) Large sample properties of generalized method of moments estimators. *Econometrica, 50, 1029-1056.*

Johnson, N.L., Kotz, S. & Balakrishnan, N. (1995) *Continuous univariate distributions, vol. 2.* Wiley.

Kraus, A. & Litzenberger, R.H. (1976) Skewness preferences and the valuation of risky assets. *The Journal of Finance vol XXXI, 1085-1099.*

Kraus, A. & Litzenberger, R.H. (1983) On the distributional conditions for a comsumption-oriented three moment CAPM. *The Journal of Finance vol XXXVIII, 1381-1389.*

Kroll, Y., Levy, H. & Markowitz, H.M. (1984) Mean-variance versus direct utility maximization. *Journal of Finance, 39, 47-61.*

Levy, H. & Markowitz, H.M. (1979) approximating expected utility by a function of mean and variance. *American Economic Review, 69, 308-317.*

Michael, J.R., Schuhany, W.R. & Haas, R.W. (1976) Generating random variables using transformations with multiple roots. *The American Statistician, 30, 88-89.*

Morgan, J.P. (1997) $CreditMetrics^{TM}$.

Rydberg, T.H. (1996) The normal inverse Gaussian Levy process: simulation and approximation. *Research Report no.344 Department of Theoretical Statistics, University of Aarhus.*