

Some new bivariate IG- and NIG-distributions for
modelling covariate financial returns ¹

Jostein Lillestøl ²

The Norwegian School of Economics and Business Administration,

Bergen, Norway

December 13, 2006

with correction February 28, 2007

¹The author thanks Jonas Andersson and Leif Sandal for helpful comments and Martin Ridout for careful reading, which uncovered a mistake.

²Contact information: Jostein Lillestøl, Institute of Finance and Management Science, The Norwegian School of Economics and Business Administration, Helleveien 30, N-5045 Bergen, Norway, tel: +4755959248, e-mail: jostein.lillestol@nhh.no

Some new bivariate IG- and NIG-distributions for modelling covariate financial returns

Abstract

The univariate Normal Inverse Gaussian (NIG) distribution is found useful for modelling financial return data exhibiting skewness and fat tails. Multivariate versions exist, but may be impractical to implement in finance. This work explores some possibilities with links to the mixing representation of the NIG distribution by the IG-distribution. We present two approaches for constructing bivariate NIG distribution that take advantage of the correlation between the univariate latent IG-variables that characterizes the marginal NIG-distribution. These are readily available from the marginal estimation, either by maximum likelihood via the EM-algorithm or by Bayesian estimation via Markov chain Monte Carlo methods. A context for implementation in finance is given.

Introduction

Classical financial theory is heavily dependent on mean-variance criteria and normal assumptions. Despite the arguments against, it may still be preferred by many practitioners due to its beauty and ease of implementation. However, in the last decades both theorists and practitioners have explored alternatives that face the empirical facts that return distributions are mostly heavy tailed and often skewed. Among classes of distributions explored are the stable, elliptic and (generalized) hyperbolic classes. The non-parametric path is also explored. The choice between them or whether to go non-parametric is not easy. Among desirable features are: ease of aggregation over space (portfolios) and time, ease of estimation and implementation. The classical theory did well on this, and provided also parameters with meaningful interpretation for both economists and practitioners. So far, the alternative theories do not match this in all respects, but progress is being made, both with respect to improvement of each alternative and the knowledge base for making choices among them.

This author has some affinity towards the Normal Inverse Gaussian distribution (NIG), see Barndorff-Nielsen (1997) and Lillo (2000). This is a four parameter subclass of the five-parameter generalized hyperbolic family, which is advocated by many authors, see Eberlein and Keller (1995), Rydberg (2000). The NIG-distribution fits well observed (marginal) returns and has some nice theoretical properties. So much for the univariate case.

In recent years extreme losses due to correlated returns are brought to the forefront. In order to model this, various approaches may be taken: Pick a class of multivariate distributions, say multivariate stable, elliptic or hyperbolic, or construct a multivariate distribution from given marginals using copulas. Although natural multivariate versions of distributions that fits well univariate data are available, they are not necessarily viable alternatives for risk estimation in practice, see, however, Aas et.al (2006). For instance,

the multivariate NIG of Barndorff-Nielsen (1997) has the defect that the key parameter of each univariate marginal depends on the joint skewness parameter vector. Moreover the estimation has to be done jointly. In many cases one has reasonably good knowledge of the marginals and just want to embed this in a joint setting, without having to reestimate all parameters jointly. Moreover, by separating the marginal-modelling from dependence-modelling one can look at the issues separately, e.g. the character of the tail dependence. In both respects the copula approach may be preferred, see Cherubini et.al. (2004) and McNeil et.al. (2005).

In this paper we explore the possibility of modelling joint distributions with NIG-marginals based on the mixture representation of Normals with the Inverse Gaussian distribution. We will follow two different approaches, both depart from classes of bivariate Inverse Gaussian distributions (IG). Each approach has its own merits and drawbacks. Various definitions of bivariate IG-distribution exist in the literature. A "natural" one is given by Kocherlakota (1986), who dismisses a previous suggestion by Al-Hussain and Abd-El-Hakim (1981), since it is not "as natural" and not easily simulated. We shall see that this is easily overcome.

Our first approach is based on defining bivariate IG-distributions in terms of their moment generating function. Our second approach departs from a suggested new scheme for bivariate simulation, based on the idea of Michael et.al. (1976). In judging the merits of the two approaches one may look for its ease of deriving moments, handling joint tail events and provide expressions for (transforms of) the distribution of linear combinations e.g. portfolios of returns. Another criterion would be the ease of bivariate simulation. Since each approach offers classes of models, we also have the problem of identifying a suitable one for the data at hand. Given our limited objective, it may be sufficient to pick one as an "all purpose" bivariate distribution. So far, we regard the second approach as the most promising, and this is explored more extensively in this paper.

The main objective of the paper is to provide approaches to the multivariate case that does not go far beyond the univariate case in estimation complexity, and is easily implemented as a bi-product from the univariate estimation. Implicitly we hope that the argument in favor of the NIG-distribution is strengthened.

The paper deals mainly with the case of independent identically distributed variables. Of course this may seem as a serious limitation when modelling financial return data, which are mostly time series exhibiting volatility clustering. In practice this can be dealt with in different ways: (1) ignore it (2) extend the model (3) estimate the devolatilized series followed by revolatilation. The first possibility may be preferred in the cases where correlation and heavy tails are the key issues and the objective is to improve upon even more simplistic methods based on independence and normality. As for the second possibility, extending the model may be done in a variety of ways. See Barndorff-Nielsen (1998) for NIG theory in the process context and Barndorff-Nielsen (1999) and Andersson (2001) for theory in the volatility clustering time series context. Given our objective, the natural one is to restrict ourselves to models where the time dependence parameters are determined by the marginal series. Such an extension of this work is also fairly straightforward. In practice one may prefer the very pragmatic approach, when the alternative is to neglect either the heavy tails and/or the correlations altogether. This may rule out the second possibility. The third possibility is in fact preferred by many risk analysts.

1 The NIG distribution and its extension

The distribution of a NIG-variate X is characterized by four parameters: $\mu, \delta, \alpha, \beta$ which relates mainly to location, scale, peakedness/heavy tail and skewness respectively. The density function of X is fairly complicated involving Bessel functions. However, the

distribution has a simple moment generating function given by

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

Here $\delta > 0$ and $0 \leq |\beta| < \alpha$. We will find it useful to write $\gamma = \sqrt{\alpha^2 - \beta^2}$ and $\phi = \delta \cdot \gamma$, and different parameterizations of the distribution using these and/or other parameters are available. From the moment generating function it is easily derived that

$$EX = \mu + \delta \cdot \frac{\beta}{\gamma} \quad \text{var}X = \delta \frac{\alpha^2}{\gamma^3}$$

and simple formulas are available for skewness and kurtosis as well. The distribution also has a convenient mixture representation as the X -marginal of (X, Z) where

$$\begin{aligned} X | Z = z &\sim N(\mu + \beta z, z) \\ Z &\sim IG(\delta, \gamma) \end{aligned}$$

Here $IG(\delta, \gamma)$ is the well known Inverse Gaussian distribution (also named Wald distribution), see Johnson, Kotz and Balakrishnan (1995) or Seshadri (1993). This means that we can write

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

where W is standard normal independent of Z . Various interpretations of Z may be given: As just a latent unobservable variable, or as a variable for data augmentation in the context of estimation. In modelling financial data it may represent volatility, although it also affects the skewness in cases of $\beta \neq 0$.

The estimation of NIG-parameters can be done by maximum likelihood methods

directly, or via the EM-algorithm, see Karlis (2002), or by Bayesian estimation via Markov chain Monte Carlo, see Karlis and Lillestøl (2004). The latter two produce a Z-series as a bi-product of the estimation process. Our approach to the bivariate case is to choose models having simple links between joint distribution and the correlation between the marginal latent variables. This correlation may be estimated as a bi-product of the marginal estimation by the methods mentioned. This procedure is akin to the method of inference functions for margins (IFM), see McLeish and Small (1988). For the bivariate case (X_1, X_2) we will have in mind the representation where

$$X_i = \mu_i + \beta_i \cdot Z_i + \sqrt{Z_i} \cdot W_i$$

with standard normal W_i 's, but where both pairs (Z_1, Z_2) and (W_1, W_2) may be correlated, with (Z_1, Z_2) being some kind of bivariate IG. This may be contrasted to the construction of Barndorff-Nielsen (1997), based on the representation formula where $Z_i \equiv Z$ and the W_i 's multivariate normal. This leads to a model with some attractive features and some drawbacks. One attractive feature is that Z may be viewed as a single latent variable, and that linear combinations are univariate NIG. A drawback is the complicated relationship between marginal and joint parameters, and that joint re-estimation is required when adding variables. Our construction gives rise to a wider range of opportunities, but admittedly also to some challenges, among them identification issues.

For the given model specification we have in general

$$\text{cov}(X_1, X_2) = \beta_1 \cdot \beta_2 \cdot \text{cov}(Z_1, Z_2) + E(Z_1 Z_2)^{1/2} \cdot \text{cov}(W_1, W_2)$$

where the first term disappears if at least one of the marginals are symmetric, and the second term disappears when the W_i 's are uncorrelated. For the correlation we may

write

$$\rho(X_1, X_2) = \frac{\beta_1\beta_2}{\alpha_1\alpha_2}\rho(Z_1, Z_2) + \kappa \cdot \frac{\gamma_1\gamma_2}{\alpha_1\alpha_2}\rho(W_1, W_2)$$

where $\kappa = E(Z_1Z_2)^{1/2}/(EZ_1 \cdot EZ_2)^{1/2}$ with $0 \leq \kappa \leq 1$ and equal to one for independence. This formula reveals how the correlation of the observable variables are affected by the correlation of the unobservables. It also gives some insight to limiting cases. If the α 's tends to infinity while the β 's stay fixed, corresponding to the marginals tending to normality, we see that the first term vanishes leaving the second term as $\kappa \cdot \rho(W_1, W_2)$. Note also the case when $W_1 = W_2 = W$ and their correlation is one. To distinguish the correlations involved, we will, when needed, use the notation $\tau_1 = \rho(Z_1, Z_2)$, $\tau_2 = \rho(W_1, W_2)$ and $\rho = \rho(X_1, X_2)$.

Useful expressions for the conditional expectation $E(X_2|X_1 = x_1)$ and conditional variance $var(X_2|X_1 = x_1)$ are not readily available in general.

2 A context for applications

The following is a possible context for application of such modelling: Suppose we want to establish an online system, where we store relevant information on return series, raw data and derived distributional parameter estimates of each series, in order to estimate correlation parameters according to our model, without having to re-estimate everything as new series come along. Suppose we use a simulation-based estimation process and keep the Z-series for later use. Instead of keeping all of them for every asset, it may be sufficient to keep a representative one for each asset, say by averaging the generated series. Each pair of such series (\hat{Z}_1, \hat{Z}_2) may then be correlated based on the joint stretch of available instants.

Example: Ford vs. GM

Consider the weekly logarithmic returns for Ford and GM equities for 20 years from mid 1983 to mid 2003, see the scatterplot in Figure 1. The computed correlation between the returns is $\rho(X_F, X_{GM}) = 0.6657$.

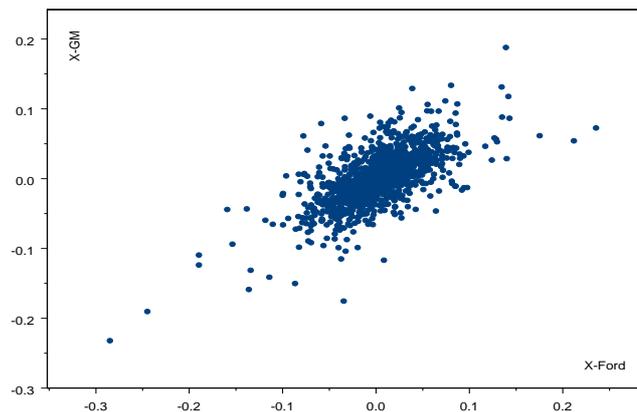


Figure 1: Scatterplot weekly log-returns GM vs. Ford 1983-2003

The parameter estimates of the marginal NIG-distributions turned out to be as in Table 1. The estimation is Bayesian using a fairly uninformative prior and sampling from the posterior by a Markov chain Monte Carlo scheme as described in Karlis and Lillestøl (2004). The estimates are based on 1000 sample repeats.

	μ	β	γ	δ	ϕ
Ford	0.0001	0.6963	27.4502	0.0550	1.5037
GM	0.0019	-0.9241	33.3621	0.0550	1.8347

Table 1: NIG parameter estimates Ford and GM

Figure 2 shows a stretch of the two retained representative Z-series. The correlation between the two Z-series is $\rho(\hat{Z}_F, \hat{Z}_{GM}) = 0.5223$. From the observed series (X_F, X_{GM}) and retained Z-series we can compute estimates $(\hat{W}_F, \hat{W}_{GM})$ of the innovation series

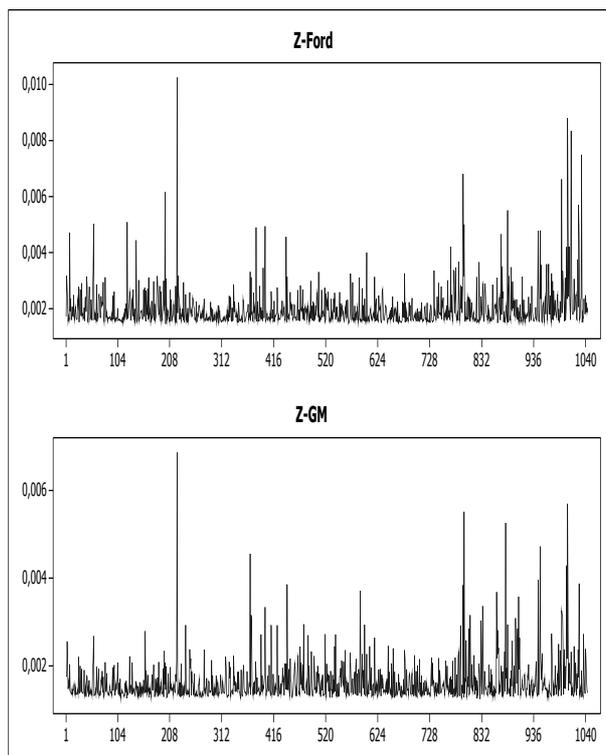


Figure 2: Z-series Ford and GM

(W_F, W_{GM}) . The correlation of these series is $\rho(\hat{W}_F, \hat{W}_{GM}) = 0.6110$. Figure 3 shows scatterplots of the pair of Z-series and W-series respectively.

The data is analyzed as if the consecutive returns are independent and identically distributed. In fact, the data show no autocorrelation, but has some weak volatility clustering. The estimated innovation series have distributions that look normal, however with somewhat shorter tail than expected. In a sense the Z-series may have picked up too much of the tail variation.

3 Bivariate IG from moment generating functions

A number of different multivariate extensions of the IG-distribution have been suggested through the years. However, most of them have marginals that are not univariate IG.

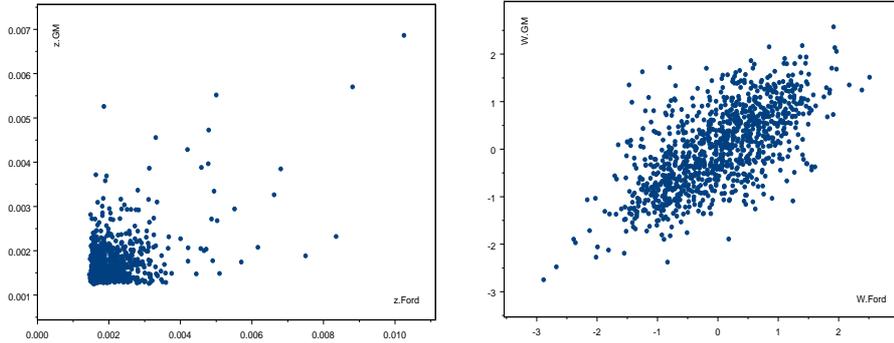


Figure 3: Scatterplots Z2 vs Z1 and W1 vs. W2 GM vs Ford

One exception is discussed by Barndorff-Nielsen et.al. (1992), where references to some of the others may be found.

Our first approach to the bivariate case is to define the distribution of $Z = (Z_1, Z_2)$ in terms of its moment generating function (mgf)

$$M(t_1, t_2) = Ee^{t_1 Z_1 + t_2 Z_2}$$

with IG-marginals $M_i(t)$ corresponding to $Z_i \sim IG(\delta_i, \gamma_i)$ for $i=1,2$. The moment generating function of the IG-distribution is (letting $\phi = \delta \cdot \gamma$)

$$M_Z(t) = Ee^{tZ} = e^{\phi(1-(1-2\gamma^{-2}t)^{1/2})}$$

The clue is now to define $M(t_1, t_2)$ with a parameter directly related to its correlation, with as few restrictions as possible. This leads to easy joint parameter estimation from data generated when estimating the marginal parameters. Note that our approach is not margin-free, since the Z_i 's are estimated from IG-assumptions. For short let $M_i(t) = \exp(\phi_i g_i(t))$, where $g_i(t) = 1 - (1 - 2\gamma_i^{-2}t)^{1/2}$. In Table 2 we give three examples of generating functions of bivariate IG-distributions with the desired features, together

with a correlation structure that fits the suggested estimation scheme.

	Moment generating function $M(t_1, t_2)$	$\rho(Z_1, Z_2)$
1.	$\exp(\phi_1 g_1(t_1) + \phi_2 g_2(t_2))(1 + \theta \cdot (1 - \exp(\phi_1 g_1(t_1))) \cdot (1 - \exp(\phi_2 g_2(t_2))))$	$\theta \cdot (\phi_1 \cdot \phi_2)^{1/2}$
2.	$\exp(\phi_1 g_1(t_1) + \phi_2 g_2(t_2)) / (1 - \theta \cdot (1 - \exp(\phi_1 g_1(t_1))) \cdot (1 - \exp(\phi_2 g_2(t_2))))$	$\theta \cdot (\phi_1 \cdot \phi_2)^{1/2}$
3.	$\exp(1 - [(1 - \phi_1 \cdot g_1(t_1))^{1-\theta} + (1 - \phi_2 \cdot g_2(t_2))^{1-\theta} - 1]^{1/(1-\theta)})$	$-\theta \cdot (\phi_1 \cdot \phi_2)^{1/2}$

Table 2: Moment generating functions and correlations for bivariate IG

Here θ is a positive covariate parameter with $\theta = 0$ corresponding to independence. The first two examples give positive correlation, the last one negative correlation. The bivariate densities may be obtained by numerical Laplace inversion. Various bivariate possibilities for Laplace inversion are described in Abate and Whitt (2006). In Figure 4 we illustrate the densities of Example 2 and Example 3 with correlation 0.5 and -0.5 respectively, with all marginal parameters equal to one. The inversion method used is an Euler-Euler algorithm.

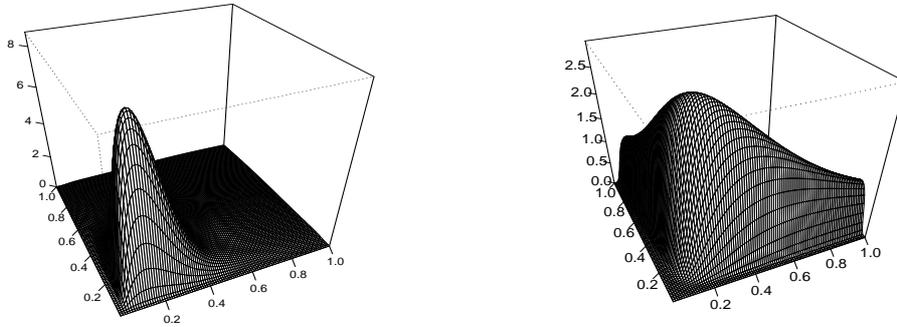


Figure 4: Bivariate IG-distributions: Example 2 (left), Example 3 (right)

Several families of both kinds exist, and it easy to identify the distribution within each family, but of course are left with the problem of which one to choose. This and the question on how to simulate from these distributions are under investigation.

This approach can be extended readily beyond the bivariate case in different ways: The simplest is to take the γ -parameter to be common to all pairs of variables. Then it can be chosen from experience or by averaging over pairs of correlations. On the other hand we can model the bivariate relationship separately, at the risk of some inconsistencies. Finally one may have a full multivariate scheme with correlation parameters for each pair.

Before we go on to our second approach it may be worthwhile to take a side view to copulas. Having expressions for moment generating functions it is possible to construct copulas which are derived from mixture representations. Following Marshall & Olkin (1988) we may obtain a joint cumulative distribution of $X = (X_1, X_2)$ having the desired marginal cumulative NIG-distributions $F_1(x)$ and $F_2(x)$ by

$$F(x_1, x_2) = L(L_1^{-1}(F_1(x_1)), L_2^{-1}(F_2(x_2)))$$

where $L(t_1, t_2) = M(-t_1, -t_2)$ is the joint Laplace-transform and L_i^{-1} is the inverse of the marginal Laplace-transform $L_i(t_i) = M_i(-t_i)$ $i=1,2$. In our case we have $L_i^{-1}(u) = \frac{1}{2}((1 - \phi_i^{-1} \ln(u_i))^2 - 1)\gamma_i^2$. It is easily checked that this leads to well known Archimedian copula functions $C(u_1, u_2)$, so that $F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$. See Nelsen (1999) for the general theory of copulas and Genest and Rivest (1993) on the identification of Archimedian copulas. This construction is, however, derived from a mixture representation of the distribution of (X_1, X_2) on (Z_1, Z_2) of frailty type, different from the one we take as point of departure. More specifically

$$F(x_1, x_2) = \int \int F_1(x_1)^{z_1} F_2(x_2)^{z_2} dG(z_1, z_2)$$

where G denotes a bivariate IG-distribution.

An advantage of the copula approach is of course that we have explicit expressions

for the probabilities of joint tail events. On the other hand, the handling of linear combinations like portfolios is not that easy.

4 Bivariate IG defined by simulation scheme

Our second approach defines a bivariate NIG-distribution from new classes of bivariate IG-distributions which are easily simulated. Different parameterizations exist for the univariate IG-distribution, and are used depending on the context. Here we use $IG(\zeta, \phi)$ with parameters $\zeta = \delta/\gamma$ and $\phi = \delta \cdot \gamma$. The density is then given by

$$f(z) = \left(\frac{\phi\zeta}{2\pi}\right)^{1/2} z^{-3/2} e^{-\frac{1}{2} \frac{\phi(z-\zeta)^2}{\zeta z}}$$

where

$$EZ = \zeta \quad var Z = \sigma^2 = \frac{\zeta^2}{\phi}$$

The "natural" definition of bivariate IG-distribution by Kocherlakota (1986) gives rise to a fairly complex density, involving an infinite series. However the moment and the conditional structures are transparent, with links to the bivariate chisquare distribution with one degree of freedom. Based on the idea of Michael et.al. (1976) this also gives the opportunity to simulate (Z_1, Z_2) from a bivariate IG-distribution $IG(\zeta_1, \zeta_2, \phi_1, \phi_2, \rho)$ from the two roots of each of the equations

$$V_i = \frac{\phi_i(Z_i - \zeta_i)^2}{\zeta_i Z_i}$$

where (V_1, V_2) is bivariate central chisquare with one degree of freedom, obtained from the standard bivariate normal with correlation ρ . By letting Z_i^- and Z_i^+ be the minus

and plus roots respectively, we take

$$\begin{aligned} Z_i &= Z_i^- & \text{with probability} & \quad p_i^- = \frac{1}{1 + Z_i^-/\zeta_i} \\ &= Z_i^+ & \text{with probability} & \quad p_i^+ = \frac{1}{1 + Z_i^+/\zeta_i} = \frac{Z_i^-/\zeta_i}{1 + Z_i^-/\zeta_i} \end{aligned}$$

The ease of simulation is taken as an argument in favor of this definition of the bivariate IG-distribution over an earlier suggestion by Al-Hussain and Abd-El-Hakim (1981). However, the idea of generating variables from equations with multiple roots suggested by Michael et.al. (1976) may be extended to cover a variety of bivariate cases by simulating (Z_1, Z_2) from independent univariate chisquare variates. We may simply use a scheme for the joint probabilities p_{12}^{ij} for $ij \in \{--, -+, +-, ++\}$ of the form given in Table 3.

	Z_2^-	Z_2^+	
$Z_1^-:$	$p_1^- \cdot p_2^- \cdot (1 + \theta \cdot g(Z))$	$p_1^- \cdot p_2^+ \cdot (1 - \theta \cdot g(Z))$	
$Z_1^+:$	$p_1^+ \cdot p_2^- \cdot (1 - \theta \cdot g(Z))$	$p_1^+ \cdot p_2^+ \cdot (1 + \theta \cdot g(Z))$	

Table 3: Simulation scheme for bivariate IG

In this simulation scheme $g(Z)$ is a suitable function of $Z = (Z_1^-, Z_1^+, Z_2^-, Z_2^+)$ and θ is a correlation measure, with natural restrictions to ensure probability values between zero and one, and where $\theta = 0$ corresponds to independence. A convenient choice is

$$p_{12}^{ij} = p_1^i \cdot p_2^j \cdot [1 + \theta \cdot \text{sign}(Z_1^i - \zeta_1) \cdot \text{sign}(Z_2^j - \zeta_2) \cdot \left(\frac{Z_1^i}{\zeta_1} \cdot \frac{Z_2^j}{\zeta_2}\right)^{1/2} \cdot h_1(V_1) \cdot h_2(V_2)]$$

where $h(V)$ indicates a function that depends on just the corresponding chisquare-variate and not on the root chosen. Note that this means positive or negative product of signs according to whether the conforming or opposing roots are chosen. Using the fact that $Z_i^- \cdot Z_i^+ = \zeta_i^2$ it is easy to check that this fits the scheme above with correct marginals.

The joint density will now have form

$$f(z_1, z_2) = f_1(z_1) \cdot f_2(z_2)(1 + \theta \cdot k_1(z_1) \cdot k_2(z_2))$$

with

$$k_i(z_i) = \text{sign}(z_i - \zeta_i) \cdot \left(\frac{z_i}{\zeta_i}\right)^{1/2} \cdot h_i(v_i(z_i))$$

and where $v_i(z_i) = \frac{\phi_i(z_i - \zeta_i)^2}{\zeta_i z_i}$. We see that we necessarily have

$$\int_0^\infty k_i(z_i) f(z_i) dz_i = E k_i(Z_i) = 0$$

and that

$$\text{cov}(Z_1, Z_2) = \theta \cdot A_1 \cdot A_2$$

where A_i is given by

$$A_i = \int_0^\infty z_i k_i(z_i) f(z_i) dz_i = E(Z_i k_i(Z_i))$$

Choices of form $h_i(v) = v^{a_i-1} \cdot \exp(-b_i \cdot v/2)$ for suitable $a_i \geq 1$ and $b_i \geq 0$ are convenient in theory. This also provides modelling flexibility that goes beyond a common fallacy in risk management, that marginal distributions and the correlations uniquely determine the joint distribution, which is true for elliptic families, but wrong in general. However, the interpretation of the a - and b -parameters may not be transparent, for two reasons: They affect both the dependency of the z_i ' via $v_i(z_i)$ for $i = 1, 2$ and the constant θ . We return to this later.

With the above choice of h -functions we are able to express the integral in terms of the Gamma-function. We have (omitting subscripts)

$$A = \frac{\zeta}{\phi^{1/2}} \cdot c(a, b) = \sigma \cdot c(a, b)$$

where

$$c(a, b) = \frac{\Gamma(2a)}{\Gamma(a + 1/2) \cdot 2^{a-1/2} \cdot (1 + b)^a}$$

This may be simplified in two cases:

$$\begin{aligned} c(a, b) &= \frac{(a-1)!}{\sqrt{2\pi}} \cdot \left(\frac{2}{1+b}\right)^a && \text{for } a \text{ integer} \\ &= \frac{1 \cdot 3 \cdot 5 \cdots (2a-1)}{(1+b)^a} && \text{for } a = n + 1/2 \text{ with } n \text{ integer} \end{aligned}$$

The correlation may now be written as

$$\tau_1 = \rho(Z_1, Z_2) = c(a_1, b_1) \cdot c(a_2, b_2) \cdot \theta$$

The cases with $a_i = n_i + 1/2$ for $n = 1, 2, \dots$ will be of particular interest. We will refer to this as the "half-integer case", where we have

$$k_i(z_i) = \left(\frac{z_i - \zeta_i}{\sigma_i}\right) v_i(z_i)^{n_i-1} e^{-\frac{1}{2}b_i v_i(z_i)}$$

For the case of common $n = 1$ and $b = 1$, i.e. $h(v) = v^{1/2} \exp(-v/2)$, we get the distribution of Al-Hussain and Abd-El-Hakim (1981) where $\theta = 8\rho$.

It is also interesting to note that if we take $h \equiv 1$, for which $\rho(Z_1, Z_2) = \theta \cdot \frac{2}{\pi}$, we get a density decreasing from the origin towards the means, and then having a temporary bump.

The moment generating function is in general

$$M(t_1, t_2) = M_1(t_1) \cdot M_2(t_2) + \theta \cdot K_1(t_1) \cdot K_2(t_2)$$

where the $M_i(t)$'s are the univariate moment generating functions given, and the $K_i(t)$'s

are similar integrals using the $k_i(t)$'s as weights, i.e. (omitting subscripts) $M(t) = \int \exp(tz)f(z)dz$ and $K(t) = \int \exp(tz)k(z)f(z)dz$. For the half-integer case we have

$$K(t) = \int_0^\infty \left(\frac{z-\zeta}{\sigma}\right)v(z)^{n-1}e^{-\frac{1}{2}bv(z)+tz}f(z)dz$$

which in the case $n = 1$ may be simplified to

$$K(t) = \phi(a(t)^{-1/2} - a(0)^{-1/2}) \cdot e^{-((a(t)a(0))^{1/2}-a(0))}$$

where $a(t) = \phi(1 + b) - 2\zeta t$.

Since the Z_i 's may be interpreted as volatilities (at least in the symmetric case), it may be of interest to look more closely at their conditional expectations. The conditional expectation of Z_2 given $Z_1 = z_1$ is in general

$$E(Z_2|Z_1 = z_1) = \zeta_2 + \theta \cdot A_2 \cdot \text{sign}(z_1 - \zeta_1) \cdot \left(\frac{z_1}{\zeta_1}\right)^{1/2} \cdot h_1(v_1(z_1))$$

For the choice of h-functions of the type above, we get for the half-integer case

$$E(Z_2|Z_1 = z_1) = \zeta_2 + c(n_1 + 1/2, b_1)^{-1} \cdot \tau_1 \frac{\sigma_2}{\sigma_1} \cdot (z_1 - \zeta_1) \cdot v_1(z_1)^{n_1-1} e^{-\frac{1}{2}b_1 v_1(z_1)}$$

where the case of $n_1 = 1$ is notably simpler, and with $b = 1$ is reduced to

$$E(Z_2|Z_1 = z_1) = \zeta_2 + 2^{3/2} \cdot \tau_1 \frac{\sigma_2}{\sigma_1} \cdot (z_1 - \zeta_1) \cdot e^{-\frac{1}{2}v_1(z_1)}$$

By series expansion we get the following approximate expression when z_1 is close to ζ_1 :

$$E(Z_2|Z_1 = z_1) \approx \zeta_2 + 2^{3/2} \cdot \tau_1 \frac{\sigma_2}{\sigma_1} \cdot \left[(z_1 - \zeta_1) - \frac{1}{6} \cdot \frac{\phi_1}{\zeta_1} (z_1 - \zeta_1)^3 \right]$$

In comparison with the linear minimum squared loss predictor, the overshoot of the first order term by the factor $2^{3/2}$ is compensated by a third order term. This term is, in a sense, related to the skewness of the distribution, but due to a relationship between the second and third moment of the IG-distribution, it has expectation $\sigma_1^2/2$.

Example: Simulation

We have simulated 100 observations from a bivariate IG-distribution corresponding to $n = 1$ and $b = 1$, i.e. the distribution of Al-Hussain and Abd-El-Hakim. The scatterplot in Figure 5 shows the case for marginal IG-distribution with $\phi = 2$ and $\delta = 1$ and (Z_1, Z_2) -correlation equal to $\tau_1 = 0.5$. In Figure 6 we show the corresponding (X_1, X_2) -plots for the cases of NIG-parameters $\mu = 0$ and $\beta = 0$ and (W_1, W_2) -correlations equal $\tau_2 = 0$ and 0.5 respectively.

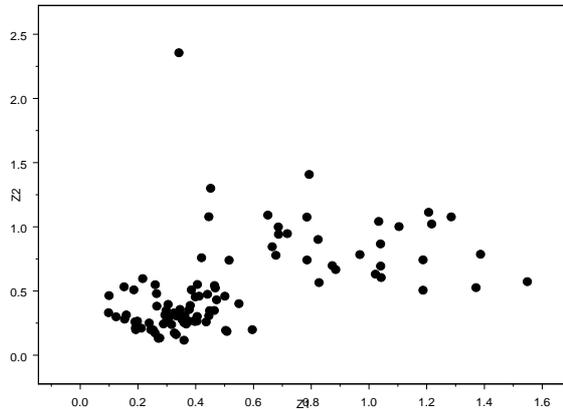


Figure 5: Scatterplot Z_2 vs Z_1 simulated data

We now return to the question of covariate model choice by determining the parameters a_i and b_i , either by informed calibration or direct estimation from data. We can get some insight to their interpretation by studying the the excess probability mass in the dominant covariate direction compared to the independent case. This can be done by plots of the "mass transfer" function $D(z) = \theta \cdot k_1(z) \cdot k_2(z) \cdot f_1(z) \cdot f_2(z)$ for given

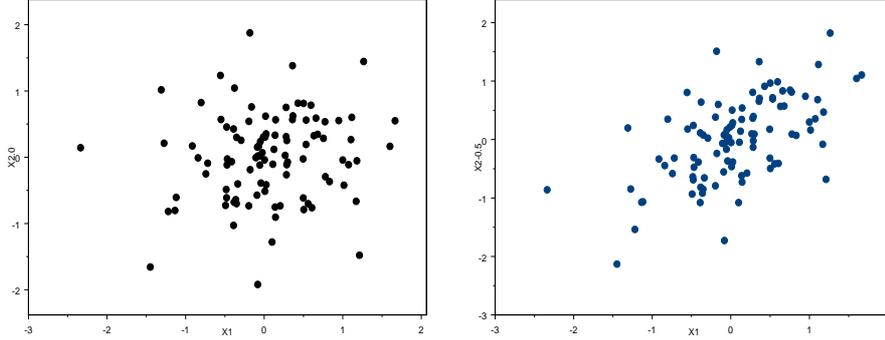


Figure 6: Scatterplots X_2 vs X_1 for W -correlations 0 and 0.5

choices of τ_1 and marginal parameters. Assuming positive τ_1 the most striking features are as follows: For $a = 1$, probability mass is added in the region where both variables are in the neighborhood of their expectations, and more is added for increasing b 's. For $a > 1$, the probability mass close to their joint expectation is not affected. The main difference is for joint values less than the expectation and to a minor degree above. Again the addition of mass in this region is increasing with the b 's. As the a 's increase the addition moves further away from the joint expectation. In none of the cases there are a major transfer of mass to the extreme tails in the covariate direction. Plots of $D(z)$ for some choices of (a, b) is given in Figure 7.

The identification from data of a model within our class may be done via empirical counterparts of the conditions $E k_i(Z_i) = 0$ and $E(Z_i k_i(Z_i))$ for $i = 1, 2$ that relates to $cov(Z_1, Z_2)$. The empirical equations for $i = 1, 2$

$$E_i(a, b) = \frac{1}{n} \sum_{j=1}^n \text{sign}(Z_{ij} - \zeta_i) \left(\frac{Z_{ij}}{\zeta_i} \right)^{1/2} v(Z_{ij})^{a-1} e^{-bv(Z_{ij})/2} = 0$$

may have several solutions in terms of a and b , and has to be supplied by some other moment restriction. Often one may have a rough idea of these parameters as a starting point to find a solution in the neighborhood.

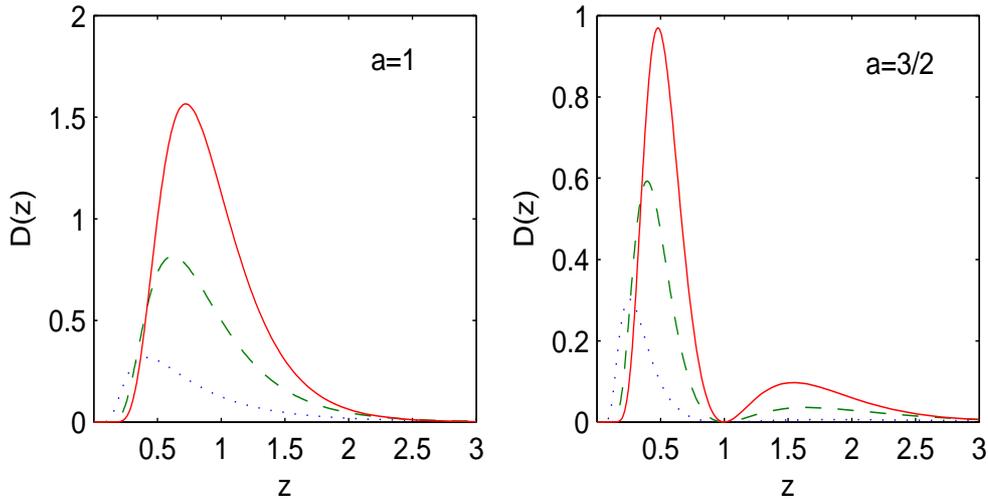


Figure 7: Examples of "mass transfer" functions $D(z)$

Example: Calibration

For our car maker equity return data, we looked for a solution with small a , and obtained the parameter fit given in Table 4. From this we see that it will make sense to choose a common structure by taking $a_i = 2$ and $b_i = 3$ for $i = 1, 2$, as compared to the choice of the model of Al-Hussain taking $a_i = 3/2$ and $b_i = 1$. This means that we have calibrated a model with $\theta = 2\pi \cdot \tau_1 = 3.28$.

	a_i	b_i	E_i
Ford	2.06	3.01	$-5.50 \cdot 10^{-6}$
GM	1.98	3.00	$-9.47 \cdot 10^{-6}$

Table 4: Parameter fit Ford - GM

The corresponding NIG-density may be obtained by straightforward integration. We write symbolically, noting that all unsubscripted quantities are essentially bivariate, and operations understood to be element-wise:

$$NIG_2[a, b](x; \mu, \beta, \delta, \gamma, \tau) = \int N_2(x; \mu + \beta z, z, \tau_2) \cdot IG_2[a, b](z; \delta/\gamma, \delta\gamma, \tau_1) dz$$

where $IG_2[a, b](z; \zeta, \phi, \tau_1)$ denotes the bivariate IG-distribution described above. A graph of a bivariate NIG density is given in Figure 8 for the case $(\mu, \beta, \delta, \gamma) = (0, 0, 1, 1)$ for both series and $(\tau_1, \tau_2) = (0.5, 0.5)$.

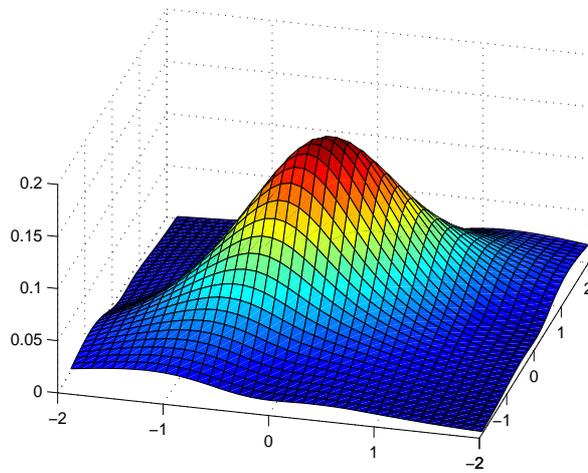


Figure 8: Bivariate NIG-density

A number of expressions for expectations, variances and covariances for the observable variables may be given in terms of marginal and covariate parameters, dependent on the parametrization used. In some applications it may be convenient to express a relationship in terms of $\kappa = E(Z_1 Z_2)^{1/2} / (E Z_1 \cdot E Z_2)^{1/2}$. Recalling the results obtained from the representation in section 1, we have for the symmetric case:

$$E(X_i) = \mu_i, \quad var(X_i) = \zeta_i, \quad cov(X_1, X_2) = (\zeta_1 \cdot \zeta_2)^{1/2} \cdot \tau_1 \cdot \kappa$$

The modifications needed for the asymmetric case are straightforward.

Although κ may be seen as an alternative measure of (Z_1, Z_2) -covariation, it varies for given τ_1 with the choice of (a, b) . To give an impression of this dependence, which affects the X-correlation ($\rho = \kappa \cdot \tau_2$ in the symmetric case), we provide Table 5 with the κ -factor for the cases $b = 1, 2, 3$ and some choices of a in steps of 0.5. As example is taken Z-correlation $\tau_1 = 0.5$ and identical univariate parameters $(\delta, \gamma) = (1, 1)$, i.e. $(\zeta, \phi) = (1, 1)$.

a	1.0	1.5	2.0	2.5	3.0	3.5	4.0
$b = 0$	0.9226	0.9108	0.9015	0.8941	0.8881	0.8831	0.8789
$b = 1$	0.9363	0.9276	0.9202	0.9139	0.9083	0.9034	0.8991
$b = 2$	0.9424	0.9356	0.9295	0.9241	0.9193	0.9149	0.9109
$b = 3$	0.9459	0.9403	0.9352	0.9306	0.9263	0.9224	0.9188

Table 5: κ -factor

Above we got some insight to how the choice of (a, b) affected the joint distribution of (Z_1, Z_2) . The question how it affects the distribution of (X_1, X_2) is a more subtle one. Clearly it affects the joint volatility in the way described above for (Z_1, Z_2) , but it also affects the joint skewness. These issues are probably best explored by computational examples, here limited to the symmetric case.

Example: Bivariate NIG and tail probabilities

It is of interest to compare tail probabilities of the defined NIG-distributions to the bivariate normal with the same expectation, variance and correlation. For this purpose we look at $Q(x) = P[(X_1 \geq x) \cap (X_2 \geq x)]$. Table 6 shows the cases $(n, b) = (1, 0), (1, 1), (1, 2), (1, 3)$ and $(n, b) = (2, 1)$ for both series (i.e. the half-integer cases corresponding to $a = 3/2$ and $a = 5/2$ respectively), and otherwise the parameters of the density in Figure 8 i.e. $(\mu, \beta, \delta, \gamma) = (0, 0, 1, 1)$ for both series and $(\tau_1, \tau_2) = (0.5, 0.5)$. The cases are indicated by appropriate subscripts and may be compared with the case

of bivariate normal $Q_0 = (x)$ with $\rho = 0.5$. To get the correct comparison, we should have chosen its X-correlation $\rho = \kappa \cdot 0.5$, with κ given in Table 5. This gives $\rho = 0.4554, 0.4638, 0.4678, 0.4702, 0.4570$ respectively for the five cases shown, and only slightly smaller binormal tail probabilities than the conservative ones given.

As expected, we see that the NIG-probabilities are somewhat higher in the tails than the binormal ones. In the context of extreme risk the probability of both variables being greater than 3.0 is typically about one in thousand for our NIG2, while it is one in ten thousand for the binormal. We also notice the differences between the NIG in the moderate tail, which disappears in the extreme tail.

x	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
$Q_{10}(x)$	0.3338	0.1285	0.0416	0.0139	0.0051	0.0021	0.0009	0.0004	0.0001
$Q_{11}(x)$	0.3338	0.1300	0.0449	0.0169	0.0068	0.0028	0.0012	0.0005	0.0002
$Q_{12}(x)$	0.3338	0.1306	0.0462	0.0182	0.0075	0.0030	0.0012	0.0005	0.0002
$Q_{13}(x)$	0.3338	0.1308	0.0469	0.0189	0.0079	0.0031	0.0012	0.0005	0.0002
$Q_{21}(x)$	0.3338	0.1290	0.0425	0.0146	0.0056	0.0024	0.0011	0.0005	0.0003
$Q_0(x)$	0.3333	0.1633	0.0625	0.0183	0.0040	0.0007	0.0001	0.0000	0.0000

Table 6: Joint tail probabilities

The model specifications above may also be used to illustrate the tail of the return (or loss) distribution for a portfolio of returns. Let us consider the equal weighted case $Y = \frac{1}{2}X_1 + \frac{1}{2}X_2$ and let $R(y) = P(Y \geq y)$. Table 7 displays the tail probabilities $R(y)$ for the same parameter choices as above, where the cases $(n, b) = (1, 0), (1, 1), (1, 2), (1, 3)$ and $(n, b) = (2, 1)$ are indicated by appropriate subscripts, and where we may compare with the corresponding binormal probabilities $R_0(y)$ for $\rho = 0.5$. In the context of extreme risk the probability of portfolio loss being greater than 3.0 is about four in thousand for our NIG2, while just two in ten thousand for the binormal with correlation

$\rho = 0.5$. Again the correct individual correlations for each separate case lead to minor differences.

y	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
$R_{10}(y)$	0.5000	0.2454	0.0987	0.0380	0.0155	0.0070	0.0038	0.0023	0.0015
$R_{11}(y)$	0.5000	0.2431	0.0975	0.0388	0.0169	0.0080	0.0043	0.0025	0.0016
$R_{12}(y)$	0.5000	0.2420	0.0967	0.0392	0.0175	0.0083	0.0043	0.0025	0.0016
$R_{13}(y)$	0.5000	0.2415	0.0966	0.0398	0.0182	0.0088	0.0044	0.0025	0.0016
$R_{21}(y)$	0.5000	0.2450	0.0985	0.0380	0.0157	0.0073	0.0040	0.0025	0.0016
$R_0(y)$	0.5000	0.2819	0.1241	0.0416	0.0105	0.0019	0.0002	0.0000	0.0000

Table 7: Tail probabilities: Portfolio

5 Further research

We have presented two possible approaches that may adapt to our general idea of using the latent Z 's as basis for covariate estimation and computation. We have explored the second approach more thoroughly than the first, mainly because it looks more promising, and some reasons are given below. Further research may include the following:

- Studies to provide a better basis for recommendation of model choice for various types of financial data, mainly by further study of the joint tail properties of each model, and providing tools for discriminating among models from data.
- Provide (approximate) expressions for conditional expectations and variances for specific models.
- Provide more sound theory behind the use of "estimated" Z -series.

- Study consequences of possible inconsistencies by pairwise modelling in the true multivariate context.

Important criteria for a viable risk management approach are of course

- the handling joint extreme tail events, and
- the handling of portfolios.

Our second approach has the advantage of providing expressions for the joint density and pairs of variables are easily simulated. Concerning the handling of portfolios, the moment generating function may be useful. By double expectation and conditioning on $Z = (Z_1, Z_2)$ it follows that the moment generating function of $X = (X_1, X_2)$ is (with $\tau_2 = \rho(W_1, W_2)$):

$$\begin{aligned} M_{(X_1, X_2)}(t_1, t_2) &= Ee^{t_1 X_1 + t_2 X_2} \\ &= e^{\mu_1 t_1 + \mu_2 t_2} E[e^{(\beta_1 t_1 + \frac{1}{2} t_1^2) Z_1 + (\beta_2 t_2 + \frac{1}{2} t_2^2) Z_2 + \tau_2 t_1 t_2 (Z_1 Z_2)^{1/2}}] \end{aligned}$$

Then the moment generating function of any linear combination $Y = w_1 X_1 + w_2 X_2$ is

$$M_Y(t) = Ee^{tY} = M_{(X_1, X_2)}(w_1 t, w_2 t)$$

The question is therefore how easy it is to evaluate bivariate integrals of exponentials with terms linear in z_1, z_2 and $(z_1 z_2)^{1/2}$.

In finance this can be applied to the return on a portfolio of correlated NIG-returns in several ways. Tail probabilities can be computed by numerical inversion as an alternative to bivariate integration of the density. A pragmatic alternative may be to approximate the distribution of Y by an appropriate NIG-distribution using the approach of Erickson et.al (2004). The general expression of the moment generating function in terms of the

Z 's may also be used for identification, estimation and/or "nonparametric" numerical inversion, in the sense that we do not have to be specific about the underlying bivariate IG. In the case of exponential utility, the formula can be used directly to rank portfolios, by appropriate choice of negative t . An advantage of the above formula is that it extends easily beyond the bivariate case.

References

- Aas, K., Haff, I.H. and X.K. Dimakos, 2005/2006, "Risk estimation using the multivariate normal inverse Gaussian distribution", *Journal of Risk, Vol. 8 Issue 2*, 39-60.
- Abate, J. and W. Whitt, 2006, "A Unified Framework for Numerically Inverting Laplace Transforms. *INFORMS Journal on Computing, vol.18 , No.4*, 408-421.
- Al-Hussaini, E.K. and N.S. Abd-El-Hakim, 1981, "Bivariate inverse Gaussian distribution", *Ann. Inst. Statist. Math.* 33, Part A, 57-66.
- Andersson, J., 2001, "On the Normal Inverse Gaussian Stochastic Volatility Model", *Journal of Business & Economic Statistics, vol. 19, Issue 1*, 44-54.
- Barndorff-Nielsen, O.E., Blæsild, P., and V. Seshadri, 1992, "Multivariate distributions with generalized inverse Gaussian marginals and associated Poisson mixtures", *The Canadian Journal of Statistics*, 20,109-120.
- Barndorff-Nielsen, O.E., 1997, "Normal Inverse Gaussian distributions and stochastic volatility modelling", *Scandinavian Journal of Statistics*, 24,1-13.
- Barndorff-Nielsen, O.E., 1998, "Processes of normal inverse Gaussian type". *Finance & Stochastics, Vol. 2 Issue 1*, 41-69

- Cherubini U., Luciano E. and W. Vecchiato, 2004, *Copula methods in Finance*, Wiley, New York.
- Eberlein, E. and U. Keller, 1995, "Hyperbolic distributions in finance", *Bernoulli*, 1, 281-299.
- Eriksson, A., Forsberg, L. and E. Ghysels, 2004, "Approximating the probability distribution of functions of random variables: A new approach", CIRANO Scientific Series, Montreal.
- Genest, C. and L. Rivest, 1993, "Statistical inference procedures for bivariate Archimedian copulas", *Journal of the American Statistical Association*, 88, 1034-1043.
- Johnson, N.L., Kotz, S. and N. Balakrishnan, 1994, *Continuous Univariate Distributions, vol.1*, Wiley, New York.
- 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. Lillestøl, 2004, "Bayesian estimation of NIG models via Markov chain Monte Carlo methods", *Applied Stochastic Models in Business and Industry*, 20, 323-338.
- Kocherlakota, S, 1986, "The bivariate inverse Gaussian distributions: an introduction", *Comm. Statist. Theor. Meth.* 15(4), 1081-1112.
- Lillestøl, J., 2000, "Risk analysis and the NIG distribution", *Journal of Risk*, 2, 41-56.
- Marshall, A.W. and I. Olkin, 1988, "Families of multivariate distributions", *Journal of the American Statistical Association*, 83, 834-841.

- McLeish, D.L. and C. G. Small, 1988, "The Theory and Applications of Statistical Inference Functions", *Lecture Notes in Statistics vol. 44*, Springer, New York.
- McNeil, A.J., Frey, R. and P. Embrechts, 2005, *Quantitative Risk Management*, Princeton University Press, Princeton.
- Michael, J.R., Schuhany, W.R. and R.W. Haas, 1976, "Generating random variables using transformations with multiple roots", *The American Statistician*, 30, 88-89.
- Nelsen, R.B., 1999, *An Introduction to Copulas*, Springer, New York.
- Rydberg, T.H., 2000, "Realistic statistical modelling of financial data", *International Statistical Review*, 68, 233-258.
- Seshadri, V., 1993, *The Inverse Gaussian Distribution*, Oxford Science Publications, Oxford.