The Generalised Hyperbolic distribution*

Some practical issues

Javier Mencía
Bank of Spain
<javier.mencia@bde.es>

Enrique Sentana
CEMFI
<sentana@cemfi.es>

March 2009

*Address for correspondence: Casado del Alisal 5, E-28014 Madrid, Spain, tel: +34 91 429 05 51, fax: +34 91 429 1056.

1 The density function

Consider an N-dimensional random vector $\mathbf{u}$, which can be expressed in terms of the following Location-Scale Mixture of Normals (*LSMN*):

$$\mathbf{u} = \boldsymbol{\alpha} + \xi^{-1}\boldsymbol{\Upsilon}\boldsymbol{\beta} + \xi^{-1/2}\boldsymbol{\Upsilon}^{1/2}\mathbf{r}, \tag{1}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are N-dimensional vectors, $\boldsymbol{\Upsilon}$ is a positive definite matrix of order N, $\mathbf{r} \sim N(\mathbf{0}, \mathbf{I}_N)$, and ξ is an independent positive mixing variable. If the mixing variable follows a Generalised Inverse Gaussian distribution (*GIG*), then the distribution of $\mathbf{u}$ will be the Generalised Hyperbolic distribution (*GH*) introduced by Barndorff-Nielsen (1977). More explicitly, if $\xi \sim GIG(-\nu, \gamma, \delta)$ then the density of the $N \times 1$ *GH* random vector $\mathbf{u}$ will be given by

$$\begin{aligned} f_{GH}(\mathbf{u}) &= \frac{\left(\frac{\gamma}{\delta}\right)^{\nu}}{(2\pi)^{\frac{N}{2}}\left[\boldsymbol{\beta}'\boldsymbol{\Upsilon}\boldsymbol{\beta}+\gamma^2\right]^{\nu-\frac{N}{2}}|\boldsymbol{\Upsilon}|^{\frac{1}{2}}K_{\nu}(\delta\gamma)}\left\{\sqrt{\boldsymbol{\beta}'\boldsymbol{\Upsilon}\boldsymbol{\beta}+\gamma^2}\,\delta q\left[\delta^{-1}(\mathbf{u}-\boldsymbol{\alpha})\right]\right\}^{\nu-\frac{N}{2}} \\ &\quad \times K_{\nu-\frac{N}{2}}\left\{\sqrt{\boldsymbol{\beta}'\boldsymbol{\Upsilon}\boldsymbol{\beta}+\gamma^2}\,\delta q\left[\delta^{-1}(\mathbf{u}-\boldsymbol{\alpha})\right]\right\}\exp\left[\boldsymbol{\beta}'(\mathbf{u}-\boldsymbol{\alpha})\right], \end{aligned}$$

where $-\infty < \nu < \infty$, $\gamma > 0$, $q\left[\delta^{-1}(\mathbf{u}-\boldsymbol{\alpha})\right] = \sqrt{1+\delta^{-2}(\mathbf{u}-\boldsymbol{\alpha})'\boldsymbol{\Upsilon}^{-1}(\mathbf{u}-\boldsymbol{\alpha})}$ and $K_{\nu}(\cdot)$ is the modified Bessel function of the third kind (see Abramowitz and Stegun, 1965, p. 374, as well as Section 4).

Given that δ and $\boldsymbol{\Upsilon}$ are not separately identified, Barndorff-Nielsen and Shephard (2001) set the determinant of $\boldsymbol{\Upsilon}$ equal to 1. However, it is more convenient to set $\delta = 1$ instead in order to reparametrise the *GH* distribution so that it has mean vector $\mathbf{0}$ and covariance matrix $\mathbf{I}_N$. Hence, if $\xi \sim GIG(-\nu, \gamma, 1)$, then $\boldsymbol{\tau} = (\nu, \gamma)'$, $\pi_1(\boldsymbol{\tau}) = R_{\nu}(\gamma)/\gamma$, and $c_v(\boldsymbol{\tau}) = \sqrt{D_{\nu+1}(\gamma) - 1}$, where $R_{\nu}(\gamma) = K_{\nu+1}(\gamma)/K_{\nu}(\gamma)$ and $D_{\nu+1}(\gamma) = K_{\nu+2}(\gamma)K_{\nu}(\gamma)/K^2_{\nu+1}(\gamma)$. It is then straightforward to use Proposition 1 in Mencía and Sentana (2009) to obtain a standardised *GH* distribution. Specifically, we set $\boldsymbol{\alpha} = -c(\boldsymbol{\beta}, \nu, \gamma)\boldsymbol{\beta}$ and

$$\boldsymbol{\Upsilon} = \frac{\gamma}{R_{\nu}(\gamma)}\left[\mathbf{I}_N + \frac{c(\boldsymbol{\beta},\nu,\gamma)-1}{\boldsymbol{\beta}'\boldsymbol{\beta}}\boldsymbol{\beta}\boldsymbol{\beta}'\right], \tag{2}$$

where

$$c(\boldsymbol{\beta},\nu,\gamma) = \frac{-1+\sqrt{1+4[D_{\nu+1}(\gamma)-1]\boldsymbol{\beta}'\boldsymbol{\beta}}}{2[D_{\nu+1}(\gamma)-1]\boldsymbol{\beta}'\boldsymbol{\beta}}. \tag{3}$$

Thus, the distribution of $\boldsymbol{\varepsilon}_t^*$ depends on two shape parameters, ν and γ, and a vector of N skewness parameters, denoted by $\boldsymbol{\beta}$.

One of the most attractive properties of the *GH* distribution is that it contains as particular cases several of the most important multivariate distributions already used in the literature. The best known examples are:

• **Normal**, which can be achieved in three different ways: (i) when $\nu \to -\infty$ or (ii) $\nu \to +\infty$, regardless of the values of γ and $\boldsymbol{\beta}$; and (iii) when $\gamma \to \infty$ irrespective of the values of ν and $\boldsymbol{\beta}$.

• **Symmetric Student** t, obtained when $-\infty < \nu < -2$, $\gamma = 0$ and $\boldsymbol{\beta} = \mathbf{0}$.

• **Asymmetric Student** t, which is like its symmetric counterpart except that the vector $\boldsymbol{\beta}$ of skewness parameters is no longer zero.

• **Asymmetric Normal-Gamma**, which is obtained when $\gamma = 0$ and $0 < \nu < \infty$ (see Madan and Milne, 1991).

• **Normal Inverse Gaussian**, for $\nu = -.5$ (see Aas, Dimakos, and Haff, 2005).

• **Hyperbolic**, for $\nu = 1$ (see Chen, Härdle, and Jeong, 2008)

• **Asymmetric Laplace**, for $\nu = 1$ and $\gamma = 0$ (see Cajigas and Urga, 2007).

2 The score function

Let $\mathbf{y}_t$ be a vector of N observed variables. To accommodate flexible specifications, we assume the following conditionally heteroskedastic dynamic regression model:

$$\left.\begin{array}{c} \mathbf{y}_t = \boldsymbol{\mu}_t(\boldsymbol{\theta}) + \boldsymbol{\Sigma}_t^{\frac{1}{2}}(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^*, \\ \boldsymbol{\mu}_t(\boldsymbol{\theta}) = \boldsymbol{\mu}\left(I_{t-1};\boldsymbol{\theta}\right), \\ \boldsymbol{\Sigma}_t(\boldsymbol{\theta}) = \boldsymbol{\Sigma}\left(I_{t-1};\boldsymbol{\theta}\right), \end{array}\right\} \tag{4}$$

where $\boldsymbol{\mu}()$ and $vech\left[\boldsymbol{\Sigma}()\right]$ are N and $N(N+1)/2$-dimensional vectors of functions known up to the $p \times 1$ vector of true parameter values, $\boldsymbol{\theta}_0$, I_{t-1} denotes the information set available at $t-1$, which contains past values of $\mathbf{y}_t$ and possibly other variables, $\boldsymbol{\Sigma}_t^{1/2}(\boldsymbol{\theta})$ is some $N \times N$ "square root" matrix such that $\boldsymbol{\Sigma}_t^{1/2}(\boldsymbol{\theta})\boldsymbol{\Sigma}_t^{1/2\prime}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}_t(\boldsymbol{\theta})$, and $\boldsymbol{\varepsilon}_t^*$ is a standardised *GH* vector martingale difference sequence satisfying $E(\boldsymbol{\varepsilon}_t^*|I_{t-1};\boldsymbol{\theta}_0) = \mathbf{0}$ and $V(\boldsymbol{\varepsilon}_t^*|I_{t-1};\boldsymbol{\theta}_0) = \mathbf{I}_N$. As a consequence, $E(\mathbf{y}_t|I_{t-1};\boldsymbol{\theta}_0) = \boldsymbol{\mu}_t(\boldsymbol{\theta}_0)$ and $V(\mathbf{y}_t|I_{t-1};\boldsymbol{\theta}_0) = \boldsymbol{\Sigma}_t(\boldsymbol{\theta}_0)$.

Importantly, given that $\boldsymbol{\varepsilon}_t^*$ is not generally observable, the choice of "square root" matrix is not irrelevant except in univariate *GH* models, or in multivariate *GH* models in which either $\boldsymbol{\Sigma}_t(\boldsymbol{\theta})$ is time-invariant or $\boldsymbol{\varepsilon}_t^*$ is spherical (i.e. $\boldsymbol{\beta} = \mathbf{0}$). But, if we parametrise $\boldsymbol{\beta}$ as a function of past information and a new vector of parameters $\mathbf{b}$ in the following way:

$$\boldsymbol{\beta}_t(\boldsymbol{\theta},\mathbf{b}) = \boldsymbol{\Sigma}_t^{\frac{1}{2}\prime}(\boldsymbol{\theta})\mathbf{b}, \tag{5}$$

then it is straightforward to see that the resulting distribution of $\mathbf{y}_t$ conditional on I_{t-1} will not depend on the choice of $\boldsymbol{\Sigma}_t^{\frac{1}{2}}(\boldsymbol{\theta})$. Finally, it is analytically convenient to replace ν and γ by η and ψ, where $\eta = -.5\nu^{-1}$ and $\psi = (1+\gamma)^{-1}$, although we continue to use ν and γ in some equations for notational simplicity.

If the mean vector and covariance matrix specifications in (4) were constant, it would be potentially advantageous to use the EM algorithm for estimation purposes. In general dynamic models, though, the EM is not as useful because it typically requires numerical maximisation procedures at each M step. However, the EM principle can still be useful to derive the score function of the *GH* distribution. In this context, the procedure that we follow is divided in two parts. In the first step, we derive $l\left(\mathbf{y}_t|\xi_t, I_{t-1};\boldsymbol{\phi}\right)$ and $l\left(\xi_t|I_{t-1};\boldsymbol{\phi}\right)$ with respect to $\boldsymbol{\phi}$. Then, in the second step, we take the expected value of these derivatives given $I_T = \{\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_T\}$ and the parameter values.

Conditional on ξ_t, $\mathbf{y}_t$ is the following multivariate normal:

$$\mathbf{y}_t|\xi_t, I_{t-1} \sim N\left[\boldsymbol{\mu}_t(\boldsymbol{\theta}) + \boldsymbol{\Sigma}_t(\boldsymbol{\theta})c_t(\boldsymbol{\phi})\mathbf{b}\left[\frac{\gamma}{R_\nu(\gamma)}\frac{1}{\xi_t}-1\right], \frac{\gamma}{R_\nu(\gamma)}\frac{1}{\xi_t}\boldsymbol{\Sigma}_t^*(\boldsymbol{\phi})\right],$$

where $c_t(\boldsymbol{\phi}) = c[\boldsymbol{\Sigma}_t^{\frac{1}{2}\prime}(\boldsymbol{\theta})\mathbf{b}, \nu, \gamma]$ and

$$\boldsymbol{\Sigma}_t^*(\boldsymbol{\phi}) = \boldsymbol{\Sigma}_t(\theta) + \frac{c_t(\boldsymbol{\phi})-1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})$$

If we define $\mathbf{p}_t = \mathbf{y}_t - \boldsymbol{\mu}_t(\boldsymbol{\theta}) + c_t(\boldsymbol{\phi})\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}$, then we have the following log-density

$$\begin{aligned} l\left(\mathbf{y}_t|\xi_t, I_{t-1};\boldsymbol{\phi}\right) &= \frac{N}{2}\log\left[\frac{\xi_t R_\nu(\gamma)}{2\pi\gamma}\right] - \frac{1}{2}\log|\boldsymbol{\Sigma}_t^*(\boldsymbol{\phi})| - \frac{\xi_t}{2}\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\mathbf{p}_t \\ &\quad +\mathbf{b}'\mathbf{p}_t - \frac{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}{2\xi_t}\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}. \end{aligned}$$

Similarly, ξ_t is distributed as a *GIG* with parameters $\xi_t|I_{t-1} \sim GIG(-\nu, \gamma, 1)$, with a log-likelihood given by

$$l\left(\xi_t|I_{t-1};\boldsymbol{\phi}\right) = \nu\log\gamma - \log 2 - \log K_\nu(\gamma) - (\nu+1)\log\xi_t - \frac{1}{2}\left(\xi_t + \gamma^2\frac{1}{\xi_t}\right).$$

In order to determine the distribution of ξ_t given all the observable information I_T, we can exploit the serial independence of ξ_t given $I_{t-1};\boldsymbol{\phi}$ to show that

$$f\left(\xi_t|I_T;\boldsymbol{\phi}\right) = \frac{f\left(\mathbf{y}_t,\xi_t|I_{t-1};\boldsymbol{\phi}\right)}{f\left(\mathbf{y}_t|I_{t-1};\boldsymbol{\phi}\right)} \propto\ f\left(\mathbf{y}_t|\xi_t, I_{t-1};\boldsymbol{\phi}\right)f\left(\xi_t|I_{t-1};\boldsymbol{\phi}\right)$$

$$\propto \xi_t^{\frac{N}{2}-\nu-1} \times \exp\left\{\frac{-1}{2}\left[\left(\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\mathbf{p}_t + 1\right)\xi_t + \left(\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b} + \gamma^2\right)\frac{1}{\xi_t}\right]\right\},$$

which implies that

$$\xi_t|I_T;\phi \sim GIG\left(\frac{N}{2}-\nu, \sqrt{\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2}, \sqrt{\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\mathbf{p}_t+1}\right).$$

From here, we can use (9) and (10) to obtain the required moments. Specifically,

$$E(\xi_t|I_T;\boldsymbol{\phi}) = \frac{\sqrt{\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2}}{\sqrt{\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t+1}}$$
$$\times R_{\frac{N}{2}-\nu}\left[\sqrt{\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2}\sqrt{\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t+1}\right],$$

$$E\left(\left.\frac{1}{\xi_t}\right| I_T;\boldsymbol{\phi}\right) = \frac{\sqrt{\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t+1}}{\sqrt{\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2}}$$
$$\times\frac{1}{R_{\frac{N}{2}-\nu-1}\left[\sqrt{\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2}\sqrt{\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t+1}\right]},$$

$$E(\log\xi_t|\,I_T;\boldsymbol{\phi}) = \log\left(\sqrt{\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2}\right)-\log\left(\sqrt{\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t+1}\right)$$
$$+\frac{\partial}{\partial x}\log K_x\left[\sqrt{\frac{\gamma c_t(\boldsymbol{\phi})}{R_\nu(\gamma)}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}+\gamma^2}\sqrt{\frac{R_\nu(\gamma)}{\gamma}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}\mathbf{p}_t+1}\right]\Bigg|_{x=\frac{N}{2}-\nu}.$$

If we put all the pieces together, we will finally have that

$$\frac{\partial l(\mathbf{y}_t|\,I_{t-1};\boldsymbol{\phi})}{\partial\boldsymbol{\theta}'} = -\frac{1}{2}vec'[\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})]\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial\boldsymbol{\theta}'}-f(I_T,\boldsymbol{\phi})\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\frac{\partial\mathbf{p}_t}{\partial\boldsymbol{\theta}'}$$
$$-\frac{1}{2}\frac{c_t(\boldsymbol{\phi})-1}{c_t(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\sqrt{1+4(D_{\nu+1}(\gamma)-1)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}vec'(\mathbf{b}\mathbf{b}')\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial\boldsymbol{\theta}'}+\mathbf{b}'\frac{\partial\mathbf{p}_t}{\partial\boldsymbol{\theta}'}$$
$$+\frac{1}{2}f(I_T,\boldsymbol{\phi})[\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\otimes\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})]\frac{\partial vec[\boldsymbol{\Sigma}_t^{*}(\boldsymbol{\phi})]}{\partial\boldsymbol{\theta}'}$$
$$-\frac{1}{2}\frac{g(I_T,\boldsymbol{\phi})}{\sqrt{1+4(D_{\nu+1}(\gamma)-1)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}vec'(\mathbf{b}\mathbf{b}')\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial\boldsymbol{\theta}'},$$

$$\frac{\partial l(\mathbf{y}_t|\,I_{t-1};\boldsymbol{\phi})}{\partial\mathbf{b}} = -\frac{c_t(\boldsymbol{\phi})-1}{c_t(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\sqrt{1+4(D_{\nu+1}(\gamma)-1)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})$$
$$-f(I_T,\boldsymbol{\phi})c_t(\boldsymbol{\phi})\mathbf{p}_t'+\boldsymbol{\varepsilon}_t'+f(I_T,\boldsymbol{\phi})\frac{c_t(\boldsymbol{\phi})-1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}(\mathbf{b}'\mathbf{p}_t)$$
$$\times\left\{\frac{[c_t(\boldsymbol{\phi})-1](\mathbf{b}'\mathbf{p}_t)}{c_t^2(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\sqrt{1+4(D_{\nu+1}(\gamma)-1)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\right.$$
$$\left.+\frac{\mathbf{p}_t'}{c_t(\boldsymbol{\phi})}-\frac{1}{\sqrt{1+4(D_{\nu+1}(\gamma)-1)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\right\}$$
$$+\frac{[2-g(I_T,\boldsymbol{\phi})]}{\sqrt{1+4(D_{\nu+1}(\gamma)-1)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta}),$$

$$
\begin{aligned}
\frac{\partial l\left(\mathbf{y}_t|\,I_{t-1};\boldsymbol{\phi}\right)}{\partial\eta} &= \frac{N}{2}\frac{\partial\log R_\nu\left(\gamma\right)}{\partial\eta}+\left(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}-\frac{1}{2c_t(\boldsymbol{\phi})}\right)\frac{\partial c_t(\boldsymbol{\phi})}{\partial\eta}+\frac{\log\left(\gamma\right)}{2\eta^2}\\
&-\frac{\partial\log K_\nu\left(\gamma\right)}{\partial\eta}-\frac{1}{2\eta^2}E\left[\log\xi_t|Y_T;\boldsymbol{\phi}\right]-\frac{f\left(I_T,\boldsymbol{\phi}\right)}{2}\left\{\frac{\partial\log R_\nu\left(\gamma\right)}{\partial\eta}\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\mathbf{p}_t\right.\\
&\left.+\frac{\partial c_t(\boldsymbol{\phi})}{\partial\eta}\left[\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}-\frac{\left(\mathbf{b}'\boldsymbol{\varepsilon}_t\right)^2}{c_t^2(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\right]\right\}\\
&-\frac{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}{2}g\left(I_T,\boldsymbol{\phi}\right)\left\{\frac{\partial c_t(\boldsymbol{\phi})}{\partial\eta}-c_t(\boldsymbol{\phi})\frac{\partial\log R_\nu\left(\gamma\right)}{\partial\eta}\right\},
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial l\left(\mathbf{y}_t|\,I_{t-1};\boldsymbol{\phi}\right)}{\partial\psi} &= \frac{N}{2}\frac{\partial\log R_\nu\left(\gamma\right)}{\partial\psi}+\frac{N}{2\psi\left(1-\psi\right)}+\left(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}-\frac{1}{2c_t(\boldsymbol{\phi})}\right)\frac{\partial c_t(\boldsymbol{\phi})}{\partial\psi}\\
&+\frac{1}{2\eta\psi\left(1-\psi\right)}-\frac{\partial\log K_\nu\left(\gamma\right)}{\partial\psi}-\frac{f\left(I_T,\boldsymbol{\phi}\right)}{2}\left\{\left[\frac{\partial\log R_\nu\left(\gamma\right)}{\partial\psi}+\frac{1}{\psi\left(1-\psi\right)}\right]\mathbf{p}_t'\boldsymbol{\Sigma}_t^{*-1}(\boldsymbol{\phi})\mathbf{p}_t\right.\\
&\left.+\frac{\partial c_t(\boldsymbol{\phi})}{\partial\psi}\left[\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}-\frac{\left(\mathbf{b}'\boldsymbol{\varepsilon}_t\right)^2}{c_t^2(\boldsymbol{\phi})\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\right]\right\}\\
&-\frac{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}{2}g\left(I_T,\boldsymbol{\phi}\right)\left\{-\frac{c_t(\boldsymbol{\phi})}{\psi\left(1-\psi\right)}+\frac{\partial c_t(\boldsymbol{\phi})}{\partial\psi}-c_t(\boldsymbol{\phi})\frac{\partial\log R_\nu\left(\gamma\right)}{\partial\psi}\right\}+g\left(I_T,\boldsymbol{\phi}\right)\frac{R_\nu\left(\gamma\right)}{\psi^2},
\end{aligned}
$$

where

$$
\begin{aligned}
f\left(I_T,\boldsymbol{\phi}\right) &= \gamma^{-1}R_\nu\left(\gamma\right)E\left(\xi_t|I_T;\boldsymbol{\phi}\right),\\
g\left(I_T,\boldsymbol{\phi}\right) &= \gamma R_\nu^{-1}\left(\gamma\right)E\left(\xi_t^{-1}|I_T;\boldsymbol{\phi}\right),
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial vec[\boldsymbol{\Sigma}_t^*(\boldsymbol{\phi})]}{\partial\boldsymbol{\theta}'} &= \frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial\boldsymbol{\theta}'}+\frac{c_t(\boldsymbol{\phi})-1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\left\{\left[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\mathbf{b}'\otimes I_N\right]+\left[I_N\otimes\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\mathbf{b}'\right]\right\}\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial\boldsymbol{\theta}'}\\
&+\frac{c_t(\boldsymbol{\phi})-1}{\left[\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\right]^2}\left\{\frac{1}{\sqrt{1+4\left(D_{\nu+1}\left(\gamma\right)-1\right)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}-1\right\}\\
&\times vec\left[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\right]vec'\left(\mathbf{b}\mathbf{b}'\right)\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial\boldsymbol{\theta}'},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial\mathbf{p}_t}{\partial\boldsymbol{\theta}'} &= -\frac{\partial\boldsymbol{\mu}_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}+c_t(\boldsymbol{\phi})\left[\mathbf{b}'\otimes I_N\right]\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial\boldsymbol{\theta}'}\\
&+\frac{c_t(\boldsymbol{\phi})-1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\frac{1}{\sqrt{1+4\left(D_{\nu+1}\left(\gamma\right)-1\right)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}vec'\left(\mathbf{b}\mathbf{b}'\right)\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial\boldsymbol{\theta}'},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial c_t(\boldsymbol{\phi})}{\partial\left(\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}\right)} &= \frac{c_t(\boldsymbol{\phi})-1}{\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}\frac{1}{\sqrt{1+4\left(D_{\nu+1}\left(\gamma\right)-1\right)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}},\\
\frac{\partial c_t(\boldsymbol{\phi})}{\partial\eta} &= \frac{c_t(\boldsymbol{\phi})-1}{\left[D_{\nu+1}\left(\gamma\right)-1\right]\sqrt{1+4\left(D_{\nu+1}\left(\gamma\right)-1\right)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\frac{\partial D_{\nu+1}\left(\gamma\right)}{\partial\eta},
\end{aligned}
$$

and

$$
\frac{\partial c_t(\boldsymbol{\phi})}{\partial\psi} = \frac{c_t(\boldsymbol{\phi})-1}{\left[D_{\nu+1}\left(\gamma\right)-1\right]\sqrt{1+4\left(D_{\nu+1}\left(\gamma\right)-1\right)\mathbf{b}'\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\mathbf{b}}}\frac{\partial D_{\nu+1}\left(\gamma\right)}{\partial\psi}.
$$

3 Skewness and kurtosis of GH distributions

We can tediously show that

$$E\left[vec\left(\boldsymbol{\varepsilon}^{*}\boldsymbol{\varepsilon}^{*\prime}\right)\boldsymbol{\varepsilon}^{*\prime}\right]=E\left[\left(\boldsymbol{\varepsilon}^{*}\otimes\boldsymbol{\varepsilon}^{*}\right)\boldsymbol{\varepsilon}^{*\prime}\right]$$
$$=c^{3}(\boldsymbol{\beta},\nu,\gamma)\left[\frac{K_{\nu+3}\left(\gamma\right)K_{\nu}^{2}\left(\gamma\right)}{K_{\nu+1}^{3}\left(\gamma\right)}-3D_{\nu+1}\left(\gamma\right)+2\right]vec\left(\boldsymbol{\beta}\boldsymbol{\beta}^{\prime}\right)\boldsymbol{\beta}^{\prime}$$
$$+c(\boldsymbol{\beta},\nu,\gamma)\left[D_{\nu+1}\left(\gamma\right)\text{-}1\right]\left(\mathbf{K}_{NN}+\mathbf{I}_{N^{2}}\right)\left(\boldsymbol{\beta}\otimes\mathbf{A}\right)\mathbf{A}^{\prime}+c(\boldsymbol{\beta},\nu,\gamma)\left[D_{\nu+1}\left(\gamma\right)\text{-}1\right]vec\left(\mathbf{A}\mathbf{A}^{\prime}\right)\boldsymbol{\beta}^{\prime},$$

and

$$E\left[vec\left(\boldsymbol{\varepsilon}^{*}\boldsymbol{\varepsilon}^{*\prime}\right)vec^{\prime}\left(\boldsymbol{\varepsilon}^{*}\boldsymbol{\varepsilon}^{*\prime}\right)\right]=E\left[\boldsymbol{\varepsilon}^{*}\boldsymbol{\varepsilon}^{*\prime}\otimes\boldsymbol{\varepsilon}^{*}\boldsymbol{\varepsilon}^{*\prime}\right]$$
$$=c^{4}(\boldsymbol{\beta},\nu,\gamma)\left[\frac{K_{\nu+4}\left(\gamma\right)K_{\nu}^{3}\left(\gamma\right)}{K_{\nu+1}^{4}\left(\gamma\right)}-4\frac{K_{\nu+3}\left(\gamma\right)K_{\nu}^{2}\left(\gamma\right)}{K_{\nu+1}^{3}\left(\gamma\right)}+6D_{\nu+1}\left(\gamma\right)-3\right]vec\left(\boldsymbol{\beta}\boldsymbol{\beta}^{\prime}\right)vec^{\prime}\left(\boldsymbol{\beta}\boldsymbol{\beta}^{\prime}\right)$$
$$+c^{2}(\boldsymbol{\beta},\nu,\gamma)\left[\frac{K_{\nu+3}\left(\gamma\right)K_{\nu}^{2}\left(\gamma\right)}{K_{\nu+1}^{3}\left(\gamma\right)}-2D_{\nu+1}\left(\gamma\right)+1\right]$$
$$\times\left\{vec\left(\boldsymbol{\beta}\boldsymbol{\beta}^{\prime}\right)vec^{\prime}\left(\mathbf{A}\mathbf{A}^{\prime}\right)+vec\left(\mathbf{A}\mathbf{A}^{\prime}\right)vec^{\prime}\left(\boldsymbol{\beta}\boldsymbol{\beta}^{\prime}\right)+\left(\mathbf{K}_{NN}+\mathbf{I}_{N^{2}}\right)\left[\boldsymbol{\beta}\boldsymbol{\beta}^{\prime}\otimes\mathbf{A}\mathbf{A}^{\prime}\right]\left(\mathbf{K}_{NN}+\mathbf{I}_{N^{2}}\right)\right\}$$
$$+D_{\nu+1}\left(\gamma\right)\left\{\left[\mathbf{A}\mathbf{A}^{\prime}\otimes\mathbf{A}\mathbf{A}^{\prime}\right]\left(\mathbf{K}_{NN}+\mathbf{I}_{N^{2}}\right)+vec\left(\mathbf{A}\mathbf{A}^{\prime}\right)vec^{\prime}\left(\mathbf{A}\mathbf{A}^{\prime}\right)\right\},$$

where

$$\mathbf{A}=\left[\mathbf{I}_{N}+\frac{c(\boldsymbol{\beta},\nu,\gamma)-1}{\boldsymbol{\beta}^{\prime}\boldsymbol{\beta}}\boldsymbol{\beta}\boldsymbol{\beta}^{\prime}\right]^{\frac{1}{2}},$$

and $\mathbf{K}_{NN}$ is the commutation matrix (see Magnus and Neudecker, 1988). In this respect, note that Mardia's (1970) coefficient of multivariate excess kurtosis will be -1 plus the trace of the fourth moment above divided by $N(N+2)$.

Under symmetry, the distribution of the standardised residuals $\boldsymbol{\varepsilon}^{*}$ is clearly elliptical, as it can be written as $\boldsymbol{\varepsilon}^{*}=\sqrt{\zeta/\xi}\sqrt{\gamma/R_{\nu}\left(\gamma\right)}\mathbf{u}$, where $\zeta\sim\chi_{N}^{2}$ and $\xi^{-1}\sim GIG\left(\nu,1,\gamma\right)$. This is confirmed by the fact that the third moment becomes 0, while

$$E\left[\boldsymbol{\varepsilon}^{*}\boldsymbol{\varepsilon}^{*\prime}\otimes\boldsymbol{\varepsilon}^{*}\boldsymbol{\varepsilon}^{*\prime}\right]=D_{\nu+1}\left(\gamma\right)\left\{\left[\mathbf{I}_{N}\otimes\mathbf{I}_{N}\right]\left(K_{NN}+\mathbf{I}_{N^{2}}\right)+vec\left(\mathbf{I}_{N}\right)vec^{\prime}\left(\mathbf{I}_{N}\right)\right\}.$$

In the symmetric case, therefore, the coefficient of multivariate excess kurtosis is simply $D_{\nu+1}\left(\gamma\right)$-1, which is always non-negative, but monotonically decreasing in γ and $|\nu|$.

4 Modified Bessel function of the third kind

The modified Bessel function of the third kind with order ν, which we denote as $K_{\nu}\left(\cdot\right)$, is closely related to the modified Bessel function of the first kind $I_{\nu}\left(\cdot\right)$, as

$$K_{\nu}\left(x\right)=\frac{\pi}{2}\frac{I_{-\nu}\left(x\right)-I_{\nu}\left(x\right)}{\sin\left(\pi\nu\right)}. \tag{6}$$

Some basic properties of $K_\nu(\cdot)$, taken from Abramowitz and Stegun (1965), are $K_\nu(x) = K_{-\nu}(x)$, $K_{\nu+1}(x) = 2\nu x^{-1}K_\nu(x) + K_{\nu-1}(x)$, and $\partial K_\nu(x)/\partial x = -\nu x^{-1}K_\nu(x) - K_{\nu-1}(x)$. For small values of the argument x, and ν fixed, it holds that

$$K_\nu(x) \simeq \frac{1}{2}\Gamma(\nu)\left(\frac{1}{2}x\right)^{-\nu}.$$

Similarly, for ν fixed, $|x|$ large and $m = 4\nu^2$, the following asymptotic expansion is valid

$$K_\nu(x) \simeq \sqrt{\frac{\pi}{2x}}e^{-x}\left\{1+\frac{m\text{-}1}{8x}+\frac{(m\text{-}1)(m\text{-}9)}{2!(8x)^2}+\frac{(m\text{-}1)(m\text{-}9)(m\text{-}25)}{3!(8x)^3}+\cdots\right\}. \tag{7}$$

Finally, for large values of x and ν we have that

$$K_\nu(x) \simeq \sqrt{\frac{\pi}{2\nu}}\frac{\exp(-\nu l^{-1})}{l^{-2}}\left[\frac{(x/\nu)}{1+l^{-1}}\right]^{-\nu}\left[1\text{-}\frac{3l\text{-}5l^3}{24\nu}+\frac{81l^2\text{-}462l^4\text{+}385l^6}{1152\nu^2}+\cdots\right], \tag{8}$$

where $\nu > 0$ and $l = \left[1 + (x/\nu)^2\right]^{-\frac{1}{2}}$. Although the existing literature does not discuss how to obtain numerically reliable derivatives of $K_\nu(x)$ with respect to its order, our experience suggests the following conclusions:

- For $\nu \leq 10$ and $|x| > 12$, the derivative of (7) with respect to ν gives a better approximation than the direct derivative of $K_\nu(x)$, which is in fact very unstable.
- For $\nu > 10$, the derivative of (8) with respect to ν works better than the direct derivative of $K_\nu(x)$.
- Otherwise, the direct derivative of the original function works well.

We can express such a derivative as a function of $I_\nu(x)$ by using (6) as:

$$\frac{\partial K_\nu(x)}{\partial \nu} = \frac{\pi}{2\sin(\nu\pi)}\left[\frac{\partial I_{-\nu}(x)}{\partial \nu} - \frac{\partial I_\nu(x)}{\partial \nu}\right] - \pi\cot(\nu\pi)K_\nu(x)$$

However, this formula becomes numerically unstable when ν is near any non-negative integer $n = 0, 1, 2, \cdots$ due to the sine that appears in the denominator. In our experience, it is much better to use the following Taylor expansion for small $|\nu - n|$:

$$\begin{aligned}\frac{\partial K_\nu(x)}{\partial \nu} &= \left.\frac{\partial K_\nu(x)}{\partial \nu}\right|_{\nu=n} + \left.\frac{\partial^2 K_\nu(x)}{\partial \nu^2}\right|_{\nu=n}(\nu - n)\\ &+ \left.\frac{\partial^3 K_\nu(x)}{\partial \nu^3}\right|_{\nu=n}(\nu - n)^2 + \left.\frac{\partial^4 K_\nu(x)}{\partial \nu^4}\right|_{\nu=n}(\nu - n)^3,\end{aligned}$$

where for *integer* ν:

$$\frac{\partial K_\nu(x)}{\partial \nu} = \frac{1}{4\cos(\pi n)}\left[\frac{\partial^2 I_{-\nu}(x)}{\partial \nu^2} - \frac{\partial^2 I_\nu(x)}{\partial \nu^2}\right] + \pi^2\left[I_{-\nu}(x) - I_\nu(x)\right],$$

$$\frac{\partial^2 K_\nu(x)}{\partial \nu^2} = \frac{1}{6\cos(\pi n)}\left[\frac{\partial^3 I_{-\nu}(x)}{\partial \nu^3} - \frac{\partial^3 I_\nu(x)}{\partial \nu^3}\right] + \frac{\pi^2}{3\cos(\pi n)}\left[\frac{\partial I_{-\nu}(x)}{\partial \nu} - \frac{\partial I_\nu(x)}{\partial \nu}\right] - \frac{\pi^2}{3}K_n(x),$$

$$\frac{\partial^3 K_\nu(x)}{\partial \nu^3} = \frac{1}{8\cos(\pi n)}\left\{\left[\frac{\partial^4 I_{-\nu}(x)}{\partial \nu^4} - \frac{\partial^4 I_\nu(x)}{\partial \nu^4}\right]\right.$$
$$\left. -4\pi^2\left[\frac{\partial^2 I_{-\nu}(x)}{\partial \nu^2} - \frac{\partial^2 I_\nu(x)}{\partial \nu^2}\right] - 12\pi^4\left[I_{-\nu}(x) - I_\nu(x)\right]\right\} + 3\pi^2\frac{\partial K_n(x)}{\partial \nu},$$

and

$$\frac{\partial^4}{\partial \nu^4}K_\nu(x) = \frac{1}{8\cos(\pi n)}\left\{\frac{3}{2}\left[\frac{\partial^5 I_{-\nu}(x)}{\partial \nu^5} - \frac{\partial^5 I_\nu(x)}{\partial \nu^5}\right]\right.$$
$$\left. -10\pi^2\left[\frac{\partial^3 I_{-\nu}(x)}{\partial \nu^3} - \frac{\partial^3 I_\nu(x)}{\partial \nu^3}\right] - 4\pi^4\left[\frac{\partial I_{-\nu}(x)}{\partial \nu} - \frac{\partial I_\nu(x)}{\partial \nu}\right]\right\} + 6\pi^2\frac{\partial^2 K_n(x)}{\partial \nu^2} - \pi^4 K_n(x).$$

Let $\psi^{(i)}(\cdot)$ denote the polygamma function (see Abramowitz and Stegun, 1965). The first five derivatives of $I_\nu(x)$ for any real ν are as follows:

$$\frac{\partial I_\nu(x)}{\partial \nu} = I_\nu(x)\log\left(\frac{x}{2}\right) - \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{\infty}\frac{Q_1(\nu+k+1)}{k!}\left(\frac{1}{4}x^2\right)^k,$$

where

$$Q_1(z) = \begin{cases} \psi(z)/\Gamma(z) \ \text{ if } z > 0 \\ \pi^{-1}\Gamma(1-z)\left[\psi(1-z)\sin(\pi z) - \pi\cos(\pi z)\right] \ \text{ if } z \le 0 \end{cases}$$

$$\frac{\partial^2 I_\nu(x)}{\partial \nu^2} = 2\log\left(\frac{x}{2}\right)\frac{\partial I_\nu(x)}{\partial \nu} - I_\nu(x)\left[\log\left(\frac{x}{2}\right)\right]^2 - \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{\infty}\frac{Q_2(\nu+k+1)}{k!}\left(\frac{1}{4}x^2\right)^k,$$

where

$$Q_2(z) = \begin{cases} \left[\psi'(z) - \psi^2(z)\right]/\Gamma(z) \ \text{ if } z > 0 \\ \pi^{-1}\Gamma(1-z)\left[\pi^2 - \psi'(1-z) - \left[\psi(1-z)\right]^2\right]\sin(\pi z) \\ +2\Gamma(1-z)\psi(1-z)\cos(\pi z) \ \text{ if } z \le 0 \end{cases}$$

$$\begin{aligned}\frac{\partial^3 I_\nu(x)}{\partial \nu^3} &= 3\log\left(\frac{x}{2}\right)\frac{\partial^2 I_\nu(x)}{\partial \nu^2} - 3\left[\log\left(\frac{x}{2}\right)\right]^2\frac{\partial I_\nu(x)}{\partial \nu} + \left[\log\left(\frac{x}{2}\right)\right]^3 I_\nu(x) \\ &\quad - \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{\infty}\frac{Q_3(\nu+k+1)}{k!}\left(\frac{1}{4}x^2\right)^k,\end{aligned}$$

where

$$Q_3(z) = \begin{cases} \left[\psi^3(z) - 3\psi(z)\psi'(z) + \psi''(z)\right]/\Gamma(z) \ \text{ if } z > 0 \\ \pi^{-1}\Gamma(1-z)\left\{\psi^3(1-z) - 3\psi(1-z)\left[\pi^2 - \psi'(1-z)\right] + \psi''(1-z)\right\}\sin(\pi z) \\ +\Gamma(1-z)\left\{\pi^2 - 3\left[\psi^2(1-z) + \psi'(1-z)\right]\right\}\cos(\pi z) \ \text{ if } z \le 0 \end{cases}$$

$$\begin{aligned}\frac{\partial^4 I_\nu(x)}{\partial \nu^4} &= 4\log\left(\frac{x}{2}\right)\frac{\partial^3 I_\nu(x)}{\partial \nu^3} - 6\left[\log\left(\frac{x}{2}\right)\right]^2\frac{\partial^2 I_\nu(x)}{\partial \nu^2} + 4\left[\log\left(\frac{x}{2}\right)\right]^3\frac{\partial I_\nu(x)}{\partial \nu} \\ &\quad - \left[\log\left(\frac{x}{2}\right)\right]^4 I_\nu(x) - \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{\infty}\frac{Q_4(\nu+k+1)}{k!}\left(\frac{1}{4}x^2\right)^k,\end{aligned}$$

where

$$Q_4(z) = \begin{cases} \left[-\psi^4(z) + 6\psi^2(z)\,\psi'(z) - 4\psi(z)\,\psi''(z) - 3\left[\psi'(z)\right]^2 + \psi'''(z)\right]/\Gamma(z) \text{ if } z > 0 \\ \pi^{-1}\Gamma(1-z)\left\{-\psi^4(1-z) + 6\pi^2\psi^2(1-z) - 6\psi^2(1-z)\,\psi'(1-z)\right. \\ -4\psi(1-z)\,\psi''(1-z) - 3\left[\psi'(1-z)\right]^2 + 6\pi^2\psi'(1-z) \\ \left.-\psi'''(1-z) - \pi^4\right\}\sin(\pi z) + \Gamma(1-z)\,4\psi^3(1-z) - 4\pi^2\psi(1-z) \\ +12\psi(1-z)\,\psi'(1-z) + 4\psi''(1-z)\cos(\pi z) \text{ if } z \le 0 \end{cases}$$

and finally,

$$\frac{\partial^5 I_\nu(x)}{\partial\nu^5} = 5\log\left(\frac{x}{2}\right)\frac{\partial^4 I_\nu(x)}{\partial\nu^4} - 10\left[\log\left(\frac{x}{2}\right)\right]^2\frac{\partial^3 I_\nu(x)}{\partial\nu^3} + 10\left[\log\left(\frac{x}{2}\right)\right]^3\frac{\partial^2 I_\nu(x)}{\partial\nu^2}$$
$$-5\left[\log\left(\frac{x}{2}\right)\right]^4\frac{\partial I_\nu(x)}{\partial\nu} + \left[\log\left(\frac{x}{2}\right)\right]^5 I_\nu(x) - \left(\frac{x}{2}\right)^\nu\sum_{k=0}^{\infty}\frac{Q_5(\nu+k+1)}{k!}\left(\frac{1}{4}x^2\right)^k,$$

where

$$Q_5(z) = \begin{cases} \left\{\psi^5(z) - 10\psi^3(z)\,\psi'(z) + 10\psi^2(z)\,\psi''(z) + 15\psi(z)\left[\psi'(z)\right]^2\right. \\ \left.-5\psi(z)\,\psi'''(z) - 10\psi'(z)\,\psi''(z) + \psi^{(iv)}(z)\right\}/\Gamma(z) \text{ if } z > 0 \\ \pi^{-1}\Gamma(1-z)\,f_a(z)\sin(\pi z) + \Gamma(1-z)\,f_b(z)\cos(\pi z) \text{ if } z \le 0 \end{cases}$$

with

$$f_a(z) = \psi^5(1-z) - 10\pi^2\psi^3(1-z) + 10\psi^3(1-z)\,\psi'(1-z) + 10\psi^2(1-z)\,\psi''(1-z)$$
$$+15\psi(1-z)\left[\psi'(1-z)\right]^2 + 5\psi(1-z)\,\psi'''(1-z) + 5\pi^4\psi(1-z)$$
$$-30\pi^2\psi(1-z)\,\psi'(1-z) + 10\psi'(1-z)\,\psi''(1-z) - 10\pi^2\psi''(1-z) + \psi^{(iv)}(1-z),$$

and

$$f_b(z) = -5\psi^4(1-z) + 10\pi^2\psi^2(1-z) - 30\psi^2(1-z)\,\psi'(1-z)$$
$$-20\psi(1-z)\,\psi''(1-z) - 15\left[\psi'(1-z)\right]^2 + 10\pi^2\psi'(1-z) - 5\psi'''(1-z) - \pi^4.$$

5 Moments of the GIG distribution

If $X \sim GIG(\nu, \delta, \gamma)$, its density function will be

$$\frac{(\gamma/\delta)^\nu}{2K_\nu(\delta\gamma)}x^{\nu-1}\exp\left[-\frac{1}{2}\left(\frac{\delta^2}{x} + \gamma^2 x\right)\right],$$

where $K_\nu(\cdot)$ is the modified Bessel function of the third kind and $\delta, \gamma \ge 0$, $\nu \in \mathbb{R}$, $x > 0$. Two important properties of this distribution are $X^{-1} \sim GIG(-\nu, \gamma, \delta)$ and $(\gamma/\delta)X \sim GIG\left(\nu, \sqrt{\gamma\delta}, \sqrt{\gamma\delta}\right)$. For our purposes, the most useful moments of X when $\delta\gamma > 0$ are

$$E\left(X^k\right) = \left(\frac{\delta}{\gamma}\right)^k\frac{K_{\nu+k}(\delta\gamma)}{K_\nu(\delta\gamma)} \tag{9}$$

$$E(\log X) = \log\left(\frac{\delta}{\gamma}\right) + \frac{\partial}{\partial\nu}K_\nu(\delta\gamma). \tag{10}$$

The GIG nests some well-known important distributions, such as the gamma ($\nu > 0$, $\delta = 0$), the reciprocal gamma ($\nu < 0$, $\gamma = 0$) or the inverse Gaussian ($\nu = -1/2$). Importantly, all the moments of this distribution are finite, except in the reciprocal gamma case, in which (9) becomes infinite for $k \geq |\nu|$. A complete discussion on this distribution can be found in Jørgensen (1982), who also presents several useful Gaussian approximations based on the following limits:

$$\sqrt{\delta\gamma}[(\gamma x/\delta) - 1] \overset{\delta\gamma \to \infty}{\to} N(0,1)$$

$$\sqrt{\delta\gamma}\log(\gamma x/\delta) \overset{\delta\gamma \to \infty}{\to} N(0,1)$$

$$\frac{\gamma^2}{2\sqrt{\nu}}\left[x - \frac{2\nu}{\gamma^2}\right] \overset{\nu \to +\infty}{\to} N(0,1)$$

$$\frac{-2\nu^{3/2}}{\delta^2}\left[x + \frac{\delta^2}{2\nu}\right] \overset{\nu \to -\infty}{\to} N(0,1)$$

References

Aas, K., X. Dimakos, and I. Haff (2005). Risk estimation using the multivariate normal inverse gaussian distribution. *Journal of Risk 8*, 39–60.

Abramowitz, M. and A. Stegun (1965). *Handbook of mathematical functions*. New York: Dover Publications.

Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proc. R. Soc. 353*, 401–419.

Barndorff-Nielsen, O. and N. Shephard (2001). Normal modified stable processes. *Theory of Probability and Mathematical Statistics 65*, 1–19.

Cajigas, J. and G. Urga (2007). A risk management analysis using the AGDCC model with asymmetric multivariate Laplace distribution of innovations. mimeo Cass Business School.

Chen, Y., W. Härdle, and S. Jeong (2008). Nonparametric risk management with Generalized Hyperbolic distributions. *Journal of the American Statistical Association 103*, 910–923.

Jørgensen, B. (1982). *Statistical properties of the generalized inverse Gaussian distribution*. New York: Springer-Verlag.

Madan, D. B. and F. Milne (1991). Option pricing with V.G. martingale components. *Mathematical Finance 1*, 39–55.

Magnus, J. R. and H. Neudecker (1988). *Matrix differential calculus with applications in statistics and econometrics*. Chichester: John Wiley and Sons.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika 57*, 519–530.

Mencía, J. and E. Sentana (2009). Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation. mimeo.