The generalized hyperbolic (GH) distribution is a continuous probability distribution superclass which includes several classes of distributions, e.g. the Student’s t-distribution. Introduced by Barndorff-Nielsen during his studies of aeolian processes, it possesses the same closure properties as the Gaussian distribution, making it relatively easy to handle analytically. Furthermore, it allows to include both asymmetric and heavy tails, i.e. tails that decline like a power law, which makes it suitable for modeling the dynamics in financial markets. In its multivariate version, the parameters are: (1) $\lambda$: real number; (2) $\chi$: real number; (3) $\psi$: real number; (4) $\boldsymbol{\mu}$: location vector; (5) $\Sigma$: dispersion matrix; (6) $\boldsymbol{\gamma}$: asymmetry vector. The parameters $\boldsymbol{\mu}$ and $\Sigma$ are not, in general, related to respectively the mean vector and the covariance matrix of the GH distribution because of the normal variance-mean mixture explained below (@QRM).

**\*\*Mathematical Definition:\*\***

$\boldsymbol{X} \sim GH(\boldsymbol{x}; \lambda, \chi, \psi, \boldsymbol{\mu}, \Sigma, \boldsymbol{\gamma})$ is a normal variance-mean mixture distribution, meaning:

$${\boldsymbol{X}} = \boldsymbol{\mu} + W \boldsymbol{\gamma} + \sqrt{W} A \boldsymbol{Z}$$

Where $W$ is a random variable that follows a generalized inverse Gaussian distribution $f\_W \left( w;\lambda,\chi,\psi \right)$, $A$ is a real-valued matrix subject to $AA^{\top} = \Sigma$, and $\boldsymbol{Z} \sim (\boldsymbol{0}, I)$.

Probability density function:

\begin{align\*} f\_{\boldsymbol{X}}(\boldsymbol{x}; \lambda, \chi, \psi, \boldsymbol{\mu}, \Sigma, \boldsymbol{\gamma}) &= \int \frac{w^{\frac{-d}{2}} e^{(\boldsymbol{x} - \boldsymbol{\mu})^\top \Sigma^{-1} \boldsymbol{\gamma}}}{(2 \pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} e^{\left(-\frac{(\boldsymbol{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu})}{2w} - w \frac{\boldsymbol{\gamma}^\top \Sigma^{-1} \boldsymbol{\gamma}}{2}\right)} f\_W \left( w;\lambda,\chi,\psi \right) \mathrm{d}w \\ &= c \frac{K\_{\lambda-\frac{d}{2}}\left(\sqrt{ \left(\chi + (\boldsymbol{x} - \boldsymbol{\mu})^{\top} \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \right) \left(\psi + \boldsymbol{\gamma}^\top \Sigma^{-1} \boldsymbol{\gamma} \right)}\right) \mathrm{e}^{(\boldsymbol{x} - \boldsymbol{\mu})^\top \Sigma^{-1} \boldsymbol{\gamma}}}{\left(\sqrt{ \left(\chi+(\boldsymbol{x} - \boldsymbol{\mu})^{\top} \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \right) \left(\psi+\boldsymbol{\gamma}^{\top}\Sigma^{-1}\boldsymbol{\gamma} \right)}\right)^{\frac{d}{2}-\lambda}}\end{align\*}

Where $c$ is the normalizing constant

$$c=\frac{(\chi\psi)^{-\frac{\lambda}{2}}\psi^{\lambda}(\psi+\boldsymbol{\gamma}^{\top}\Sigma^{-1}\boldsymbol{\gamma})^{\frac{d}{2}-\lambda}}{(2\pi)^\frac{d}{2}|\Sigma|^{\frac{1}{2}}K\_{\lambda}(\sqrt{\chi\psi})}$$

$d$ is the number of dimensions (i.e. variables), and $K\_i(\cdot)$ is the modified Bessel function of the third kind with index $i$.

Expected value:

$$\mathrm{E}(\boldsymbol{X}) \stackrel{(1)}{=} \boldsymbol{\mu} + \mathrm{E}(W) \boldsymbol{\gamma} \stackrel{(2)}{=} \boldsymbol{\mu} + \sqrt{\frac{\chi}{\psi}} \frac{K\_{\lambda + 1} \left(\sqrt{\chi \psi} \right)}{ K\_{\lambda} \left(\sqrt{\chi \psi} \right)} \boldsymbol{\gamma}$$

Covariance:

$$\mathrm{Cov}(\boldsymbol{X}) \stackrel{(1)}{=} \mathrm{E}(W) \Sigma + \mathrm{Var}(W) \boldsymbol{\gamma} \boldsymbol{\gamma}^{\top} \stackrel{(2)}{=} \sqrt{\frac{\chi}{\psi}} \frac{K\_{\lambda + 1} \left(\sqrt{\chi \psi} \right)}{ K\_{\lambda} \left(\sqrt{\chi \psi} \right)} \Sigma + \frac{\chi}{\psi} \left[\frac{K\_{\lambda + 2} \left(\sqrt{\chi \psi} \right)}{ K\_{\lambda} \left(\sqrt{\chi \psi} \right)} - \left(\frac{K\_{\lambda + 1} \left(\sqrt{\chi \psi} \right)}{K\_{\lambda}\left(\sqrt{\chi \psi} \right)} \right)^{2} \right] \boldsymbol{\gamma} \boldsymbol{\gamma}^{\top}$$

From the first equality (1) in each of the above two formulas it can be inferred why in general $\boldsymbol{\mu}$ and $\Sigma$ are not related to respectively the mean vector and the covariance matrix of $\boldsymbol{X}$. For $\boldsymbol{\mu}$ to be the mean vector, the joint distribution has to be symmetric, i.e. $\boldsymbol{\gamma} = \boldsymbol{0}$; additionally, for $\Sigma$ to correspond to the covariance matrix, $\mathrm{E}(W) = 1$ has to hold. The last step (2) from the general to the specific form of the moment formulas is valid only if the first two moments of $W$ can be calculated analytically, i.e. if $\chi$ and $\psi$ are strictly positive. It should be noted that, all else being equal, for $\boldsymbol{\gamma} = \boldsymbol{0}$ the GH distribution is symmetric because the normal variance-mean mixture above simplifies to a normal mean mixture $\boldsymbol{X} = \boldsymbol{\mu} + \sqrt{W} A \boldsymbol{Z}$. It is the mixture of variances that enables the GH distribution to have tails with different weights.