



Value at Risk Using Hyperbolic Distributions

Christian Bauer

This article deals with the Value at Risk concept as it is used in practice. We show that, like the Gaussian distribution, elliptical distributions lend themselves to simple practical computations. All necessary computations are detailed for the symmetric hyperbolic distributions. A test on real stock market and exchange rate data shows the new distributions fit the data better and outperform equivalent estimators used in RiskMetrics™.¹ © 2000 Elsevier Science Inc.

Keywords: Hyperbolic distribution; Elliptical distributions; Value at risk

JEL classification: C5; G1

I. Introduction

The modeling of financial time series generally proceeds in one of two main directions. The first is to model the process itself (this is common in the univariate case), and the second is to model the distribution of the price change (common in the more complicated multivariate case). It is well known that the real distribution of the percentage price change is not Gaussian for it has fatter tails and a thinner waist than the Gaussian distribution. For the univariate case, some approaches such as ARCH-type processes in Bollerslev et al. (1992) or Levy processes based on hyperbolic distributions in Eberlein and Keller (1995) take this into account. However, many multivariate applications, like the widely used value-at-risk concept (VaR) measuring the risk of a certain portfolio, are still based on normally distributed random variables. These distributions are considered the only ones simple enough to allow rapid computations.²

To remedy this problem, we show that a more general class of distributions, namely the elliptical distributions, can be used for calculating the VaR with little additional effort, but with much better fits to the data. For practical reasons, one additionally needs simply

Universitätsstr. 30, Gebäude NW II, D-95447 Bayreuth, Germany.

Address correspondence to: Christian Bauer, Universitätsstr. 30, Gebäude NW II, D-95447 Bayreuth, Germany.

¹ The risk management system RiskMetrics™ by J. P. Morgan is probably the most widely spread implementation of the VaR concept. It is therefore used as reference.

² Longerstae et al. (1996).

computable marginal distributions, efficient simulation algorithms, and good estimation procedures. One class of multivariate distributions with these properties is the class of symmetric hyperbolic distributions to be introduced in Section III. All necessary calculation algorithms, formulas, and generators for random variables are provided in the same section and the Appendix.

II. The Value-at-Risk Concept

Let $PV(r)$ denote the present value of a given portfolio at price r of the underlying assets. $\Delta PV_r(x) \stackrel{\text{def}}{=} PV(r \cdot (\mathbf{1} + x)) - PV(r)$ is the change in the value of the portfolio, if the asset price moves 100x%.³ The value-at-risk of level p (usually $p = 5\%$ or $p = 1\%$) is defined as the infimum value, such that

$$P(\Delta PV_r(x) \leq -\text{VaR}) \leq p. \quad (1)$$

This implies that the probability of losing more than VaR is p . So the negative VaR is simply the p -quantile of the distribution of ΔPV and does, despite its name, not represent the total risk structure of the portfolio. The distribution of ΔPV is usually far from normal. Portfolios with the same VaR of level p may have various conditional expected losses, if the value falls below the VaR. Presenting a VaR of a second level, for example, 5% and 1%, or categorizing ΔPV to flat, semi-heavy or heavy tailed would allow a better understanding of the risk structure.

The critical point, however, is the stochastic nature of the move x . Present risk management systems often suppose x to be normally distributed and estimate the parameters from the past data. Once the distribution in the model is defined, one has three distribution dependent ways, beside historic simulation, to estimate the VaR:

- the structured Monte–Carlo–Simulation,
- the δ -method for linear portfolios and
- the $\delta\Gamma$ -method for portfolios containing nonlinear assets like options.

The structured Monte–Carlo method estimates the distribution of $\Delta PV_r(x)$ by simulating a large number of random variables x and calculating the resulting change of value of the portfolio. One therefore needs efficient generators for the distribution of x .

In linear portfolios, for example, portfolios containing no derivatives, the value of the portfolio is just $a\mathbf{1}^T$, if the components of the vector a refer to the value of a particular asset in the portfolio. This means $a = \alpha \cdot r$, where α is the vector containing the number of shares of each asset. This obviously yields $\Delta PV_r(x) = ax^T$. This simple relation implies an easy formula for the VaR, if x is elliptically distributed.

The distribution F of a random vector X is called elliptic, if there exists a nonsingular matrix Γ , so that the distribution of $X\Gamma$ is rotation-symmetric. For elliptic distributions, there is a representation of the density f of F , which contains X only as $(X - \mu)C^{-1}(X - \mu)^T$, where μ is the parameter of location and $C = \lambda\Gamma\Gamma^T$ a positive definite, symmetric matrix. C is called the scale matrix of F , because, λ appropriately chosen, any change of scale⁴ in X will result only in an equivalent change in μ and C . Let h be an eigenvector

³ x is the vector of the percentual price chages of each asset and $\mathbf{1}$ is a vector containing 1 in every component. As product we use the usual vector product.

⁴ To be more precise any regular, affine transformation $X \rightarrow XA + b$ will change μ to $\mu A + b$ and C to $A^T C A$. All other parameters of f remain unchanged.

of C with eigenvalue λ_h and $\langle h \rangle$ the one-dimensional subspace spanned by h . F_h is the one-dimensional projection of F on the line $\langle h \rangle$ normed by the square root of λ_h . F_h is independent of the choice of h and called the one-dimensional distribution of F . Let $c(p)$ denote the p -quantile of F_h .

If H is a tangent hyperplane to the ellipsoid $(x - \mu)C^{-1}(x - \mu)^T = c(p)^2$ in \mathbb{R}^d and H_p the half-space cut off by H , which is not containing the ellipsoid, then $P(x \in H_p) = p$. This is obvious by the definition of $c(p)$, if H is normal to an eigenvector h of C . The other cases can easily be derived using the “elliptical symmetry” of F .

It is well known that the minimum of a linear function ax^T under the condition $xC^{-1}x^T = c^2$, C positive definite, is attained at the point $x_0 = -(c/\sqrt{aCa^T})aC$ with value $-c\sqrt{aCa^T}$. Combining these two facts, one gets the following:⁵

Lemma 1: *The Value-at-Risk at level p of a linear portfolio a with an elliptic distribution F of the percentage change x of the asset prices is*

$$\text{VaR} = -c(p)\sqrt{aCa^T} + a\mu^T, \quad (2)$$

where C is the scale matrix of F , μ the parameter of location and $c(p)$ the p -quantile of the one-dimensional distribution of F .

This result holds only for positive definite matrices. However, this formula for the computation of the VaR in practice is also used with nondefinite estimates of covariance matrices.

The $\delta\Gamma$ -method is the Taylor approximation of degree 2 of the value function of the portfolio. $\Delta PV_r(x) \approx \delta x + 0.5x\Gamma x^T + \theta\tau$, with $\delta = \partial PV_r/\partial x(x)$, $\Gamma = \partial^2 PV_r/\partial x^2(x)$ and $\theta\tau$ is the time-dependent value to maturity. There are various (not completely satisfying) methods to estimate the VaR of this approximation, for example, fitting Johnson curves in RiskMetrics™, Longerstaeay et al. (1996) or nonlinear minimum problems in Haaf (1996), if the price change x is normally distributed. All of these can be extended to symmetric hyperbolic distributions.⁶

Two further practical requirements for the model used shall now be specified. Firstly, working with many portfolios necessitates a very fast calculation of the parameters of the estimated distribution. So far this is done by estimating the parameters for the distribution of a huge variety of assets (general distribution) in overnight computations. The user then obtains this time-intensive result and is able to calculate thereof the parameters of the marginal distribution of the assets used in a certain portfolio very quickly. Using the Gaussian distribution model, this just means taking the relevant part of the mean vector and the covariance matrix. We will show that getting the parameters for the marginal distribution of a symmetric hyperbolic distribution is equally simple. Secondly, the distribution of the price change of assets seems to be nonconstant over time.⁷ So the commonly used estimators, like the exponentially weighted estimator in RiskMetrics™, Longerstaeay et al. (1996), use rather little of the past data and/or weigh recent data more than past data.

⁵ See Bauer (1998) for an explicit proof.

⁶ The approximation is a linear combination of terms x_i and $x_i x_j$. To estimate its moments one only needs to estimate the moments of x . These can be calculated from the parameter estimates given in Section IV.

⁷ This is indicated by structures as e.g. volatility clustering.

III. The Symmetric Hyperbolic Distribution

In 1977, Barndorff-Nielsen (1977) introduced the family of generalized hyperbolic functions. The d -dimensional density function H_d has six parameters: δ and Δ for multivariate scales, μ for location, β for skewness, and α and λ mainly change the tails.⁸ We have

$$H_d(\alpha, \beta, \delta, \mu, \Delta, \lambda)(x) = \frac{\kappa^\lambda \alpha^{(1/2)d-\lambda}}{(2\pi)^{(1/2)d} \delta^\lambda K_\lambda(\delta\kappa)} \frac{K_{\lambda-(1/2)d}(\alpha[\delta^2 + (x - \mu)\Delta^{-1}(x - \mu)^T]^{1/2})}{[\delta^2 + (x - \mu)\Delta^{-1}(x - \mu)^T]^{1/2((1/2)d-\lambda)}} e^{\beta(x-\mu)^T}, \quad (3)$$

with $\kappa = \sqrt{\alpha^2 - \beta\Delta\beta^T}$, Δ a symmetric, positive definite $d \times d$ -matrix with determinant 1 and K_ν the modified Bessel function of the third kind.⁹ It is a well known theorem of Blæsild, that the family of generalized hyperbolic distributions is closed under regular affine transformations and with respect to formation of marginal and conditional distributions.¹⁰ Blæsild explicitly describes the parameters of marginal, conditional or affine transformed hyperbolic distributions. A direct result of this theorem are the following lemma 2 and 3:¹¹

Lemma 2: *The parameter β is 0, iff the corresponding parameter $\hat{\beta}$ of any one dimensional marginal distribution is 0, i.e. the generalized hyperbolic distribution is elliptic, iff all one dimensional marginal distributions are symmetric.*

The one dimensional marginal distributions are normal hyperbolic distributions, iff the generalizing parameter λ is 1.

A generalized hyperbolic distribution with parameters $\beta = 0$ and $\lambda = 1$ is therefore called a symmetric hyperbolic distribution. We denote the density E and replace the parameter α with $\zeta \stackrel{\text{def}}{=} \alpha\delta$:

$$E_d(\zeta, \delta, \mu, \Delta)(x) = \frac{\zeta^{d-1}}{(2\pi)^{(1/2)d} \delta^d K_1(\zeta)} \frac{K_{1-(1/2)d}\left(\zeta \sqrt{1 + \frac{x - \mu}{\delta} \Delta^{-1} \frac{x - \mu^T}{\delta}}\right)}{\left(\zeta \sqrt{1 + \frac{x - \mu}{\delta} \Delta^{-1} \frac{x - \mu^T}{\delta}}\right)^{(1/2)d-1}}. \quad (4)$$

The scale matrix of E is $\delta^2\Delta$.

We can specialise Blæsild's theorem to the case of symmetric hyperbolic distributions. Lemma 3 then states that the subfamily of symmetric hyperbolic distributions is closed with respect to forming marginal distributions and regular affine transformations.

Lemma 3: *Let X be a d -dimensional random vector distributed according to the symmetric hyperbolic distribution $E_d(\zeta, \delta, \mu, \Delta)(X)$. Given a partition (X_1, X_2) of X , with*

⁸ α , δ and λ are real, μ and β vector parameters. Using $\lambda = 1/2(d + 1)$ one gets the usual multivariate hyperbolic distribution with density

$$\frac{\kappa^{1/2(d+1)}}{(2\pi)^{1/2(d-1)} 2\alpha\delta^{1/2(d+1)} K_{1/2(d+1)}(\delta\kappa)} e^{\alpha \sqrt{\delta^2 + (x-\mu)\Delta^{-1}(x-\mu)^T + \beta(x-\mu)^T}}$$

⁹ Abramowitz and Stegun (1972) provide a description of the properties, approximation and computation algorithms of the Bessel functions. The modified Bessel functions of the third kind are also known as MacDonald functions.

¹⁰ See Blæsild (1981).

¹¹ See Bauer (1998) for an explicit proof of the theorem and the corollaries with the used notation.

r the dimension of X_1 , and analogously defined partitions (β_1, β_2) and (μ_1, μ_2) . Let Δ have block representation

$$\Delta = \begin{pmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{pmatrix},$$

so that Δ_{11} is a $r \times r$ -matrix. Then the following holds:

1. The distribution of X_1 is a r -dimensional symmetric hyperbolic distribution $E_r(\hat{\xi}, \hat{\delta}, \hat{\mu}, \hat{\Delta})$ with parameters $\hat{\xi} = \zeta$, $\hat{\delta} = \delta|\Delta_{11}|^{1/2r}$, $\hat{\mu} = \mu_1$ and $\hat{\Delta} = |\Delta_{11}|^{1/r}\Delta_{11}$.
2. Let $Y = XA + B$ be a regular affine transformation of X and $\|A\|$ denote the absolute value of the determinant of A . Then Y is distributed according to the symmetric hyperbolic distribution $E_d(\tilde{\xi}, \tilde{\delta}, \tilde{\mu}, \tilde{\Delta})$ with $\tilde{\xi} = \zeta$, $\tilde{\delta} = \delta\|A\|^{1/d}$, $\tilde{\mu} = \mu A + B$ and $\tilde{\Delta} = \|A\|^{2/d}A^T\Delta A$.

We then need the moments of the distribution. The generalized hyperbolic distribution is an exponential family with respect to β . So its moments can be derived from the partial derivatives of the logarithm of the densities norming constant.¹² We get

$$\mathbf{E}(X - \mu) = 0, \tag{5}$$

$$\mathbf{E}((X - \mathbf{E}(X))^2) = \text{Cov}(X) = \frac{K_2(\zeta)}{\zeta K_1(\zeta)} \delta^2 \Delta, \tag{6}$$

$$\mathbf{E}((X - \mathbf{E}(X))^3)_{i,j,k} = 0 \quad \text{and} \tag{7}$$

$$\mathbf{E}((X - \mathbf{E}(X))^4)_{i,j,k,l} = \frac{K_3(\zeta)}{\zeta^2 K_1(\zeta)} \delta^4 (\Delta_{i,i}\Delta_{j,k} + \Delta_{j,l}\Delta_{i,k} + \Delta_{k,l}\Delta_{j,i}). \tag{8}$$

The interpretation of the 4 parameters is: μ is a parameter of location, Δ defines only the shape of the covariance matrix and δ varies only its size, while ζ is independent of transformations of the parameters of scale and location and determines the relation between the tails and the waist of the distribution. However, a change in ζ will result in a change of the size of the covariance matrix.

There also exists an efficient generator for generalized hyperbolic random variables. Atkinson (1982) suggested a two or three envelope rejection algorithm. It uses the fact that the generalised hyperbolic distributions are mixed distributions. If σ^2 is a random variable from an generalized inverse Gaussian distribution with parameters λ , δ^2 and κ^2 , Φ the Cholesky decomposition of Δ and Y a d -dimensional standard normal random vector, then $X = \sigma Y \Phi^T + \mu + \sigma^2 \beta \Delta$ is distributed with respect to $H_d(\alpha, \beta, \delta, \mu, \Delta, \lambda)$.

The algorithm for creating random variables from an generalized inverse Gaussian distribution with parameters λ , δ^2 and κ^2 simplifies, if we restrict the parameters to $\lambda = 1$ and $\beta = 0$.¹³ Using small values of ζ , as those estimated from the data used, it takes an

¹² The i^{th} derivative of $-\ln(\kappa^\lambda \alpha^{1/2(d-\lambda)} / (2\pi)^{(1/2)d} \delta^\lambda K_\lambda(\delta\kappa))$ with respect to β is the i^{th} cumulant of $H_d(\alpha, \beta, \delta, \mu, \Delta, \lambda)$. With some calculation we get

$$\mathbf{E}(X - \mu) = \frac{\partial}{\partial \beta} \ln \left(\frac{(2\pi)^{(1/2)d} \delta^\lambda K_\lambda(\delta\kappa)}{\kappa^\lambda \alpha^{(1/2)d-\lambda}} \right) = \frac{\delta K_{\lambda+1}(\delta\kappa)}{\kappa K_\lambda(\delta\kappa)} \beta \Delta,$$

which simplifies to 0 for $\beta = 0$. The higher order moments can be derived in the same way.

¹³ A description is given in the Appendix.

ISP-macro¹⁴ less than 1 min to generate 10 000 generalized inverse Gaussian random variables.

The efficiency of the algorithm relies only on ζ and the algorithm creating normal and uniformly distributed random variables. Creating high dimensional hyperbolically distributed random variables and a Monte-Carlo evaluation of a portfolio with many assets one faces no more problems than if normal random vectors were used.

IV. Data and Estimation: A Comparison of the Gaussian and the Hyperbolic Model

Estimating the parameters of a symmetric hyperbolic distribution is not as simple as it is for a normal distribution. The Bessel functions make it impossible to obtain a closed formula for a maximum likelihood estimator. Numerical approximations to ML estimators of all parameters—the dimension of the parameter space is $\frac{1}{2}(d + 1)(d + 2)$ —face enormous numerical problems, which we could not solve with justifiable efforts. However, the first two moments indicate some simple estimators for μ and Δ . According to Fang and Zhang (1990) the ML-estimates for location and scale of an elliptical distribution are the mean and the empirical covariance matrix multiplied with real number depending on the distribution. So we can choose various estimators for mean and variance, when norming the covariance matrix to determinant 1. This leaves two parameters, δ and ζ , to be estimated. We suggest a two-step procedure. First calculate estimates $\tilde{\mu}$ and $\tilde{\Delta}$ from the data $\{x_i\}_{i=1..n}$, then estimate δ and ζ from the transformed data $\omega_i = (x_i - \tilde{\mu})\tilde{\Delta}^{-1}(x_i - \tilde{\mu})^T$. The simultaneous estimation of δ and ζ still needs care because the likelihood function of δ and ζ is very flat, that is, various values of δ and ζ have almost the same likelihood. The independence of the parameters makes a flat-tailed distribution with a big scale look like a fat-tailed distribution with small scale over the set of data. The likelihood function appears to have several local minima. We solved the problem using a two-step procedure. First, we approximated the global maximum of the likelihood function with a lattice. We used an iterative procedure evaluating the likelihood function on a lattice and put a tighter lattice around the minimum in the next step. In a second part, we used the first approximation as a starting point for a Newton procedure.¹⁵

We use daily data of German stocks (Bayer, Deutsche Bank, and VW) and international indices (DAX, Dow Jones, and Nikkei) from the years 1987 to 1997 to test our model. The total length of the data set is 2240. In Figure 1 a quantile-quantile-plot of the DAX-data against the estimated normal distribution and the estimated hyperbolic distribution shows the flat tails of the Gaussian distribution.¹⁶

For a more detailed examination of the fit of one-dimensional hyperbolic distributions to financial data, we refer to the papers of Eberlein and Keller (1995).

As remarked in Section II, the distribution of the price change seems to be non-constant over time. We therefore use the same estimators as J. P. Morgan RiskMetrics™: an exponential weighted estimator for the covariance matrix $\tilde{C}_t = \text{const}/n \sum_{i=0}^n \lambda^i (x_{t-i} - \tilde{\mu})(x_{t-i} - \tilde{\mu})^T$. Because we want to estimate Δ with determinant 1, we do not care for the norming constant. The optimal value, according to J. P. Morgan, for λ is 0.94 and for n is 74.¹⁷ J. P. Morgan assumes $\tilde{\mu} \equiv 0$, and we adapt this to keep the models comparable.

¹⁴ Interactive Scientific Processor is a statistical programme from Datavision AG.

¹⁵ See Bauer (1998) for the C++ source code.

¹⁶ We estimated the multivariate parameters and calculated the marginal distribution.

Table 1. Results of VaR Estimation of Stock Data Using Hyperbolic and Gaussian Models

Asset	$p = 5\%$		$p = 1\%$	
	Hyperbolic	Gaussian	Hyperbolic	Gaussian
DJII	4.71	5.90	1.07	1.83
DAX	4.17	4.87	1.36	2.31
Nikkei	5.15	6.70	2.13	3.14
Bayer	4.95	5.79	1.21	2.31
Dt. Bank	4.47	5.21	1.31	2.60
VW	4.02	5.60	1.21	2.07

We began with the estimation in 1987, and counted how often in the following 2 weeks (10 data points) the estimated VaR for any single asset price was exceeded. We then shifted the data used for estimation 10 days into the future, and went on. In all, the estimation was repeated 216 times and 2160 data points were compared to the resulted VaR-predictions. Table 1 shows the percentage of the exceedings for the hyperbolic and the normal estimates.

Obviously the 5% Value at Risk is estimated quite well by both models, the hyperbolic tends to be a little too high, the Gaussian model to be too low. However, the 1% VaR is estimated correctly only by the hyperbolic model. The flat tails of the normal distributions underestimate the VaR to a high extent

We also tested the model in the same way on data of exchange rates (US-Dollar, Yen and French Franc) against the German Mark over the same period with similar results shown in Table 2.

A comparison of a kernel density estimate of US-Dollar data to the estimated Gaussian and the hyperbolic distribution shown in Figure 2 makes the problem of the flat tails obvious.¹⁹

The symmetric hyperbolic model has another advantage. It contains the Gaussian model as a limiting case. If δ and ζ tend to infinity in such a way that

$$\lim_{\delta, \zeta \rightarrow \infty} \frac{\delta^2}{\zeta} = \sqrt[4]{c^2},$$

Table 2. Results of VaR Estimation of Exchange Rate Data Using Hyperbolic and Gaussian Models

Asset	$p = 5\%$		$p = 1\%$	
	Hyperbolic	Gaussian	Hyperbolic	Gaussian
Jap. Yen	4.90	6.03	1.31	2.75
US-Dollar	4.66	5.50	1.21	2.51
Fr. Franc ¹⁸	5.19	6.32	2.18	3.62

¹⁷ So the weight of the last data that enters the estimate is approximately 1%.

¹⁸ The French Franc probably should be evaluated separately, because it was tied to the German Mark rather strictly in the EMS compared to US-Dollar and Yen.

¹⁹ For data $\{x_i\}_{i=1, \dots, n}$ we estimate $f(X) = 1/n \sum_{i=1}^n K_\rho(x_i - X)$ using a Picard type kernel $K_\rho(x) = (1/\rho)e^{\rho|x|}$ with $\rho = 1000$.

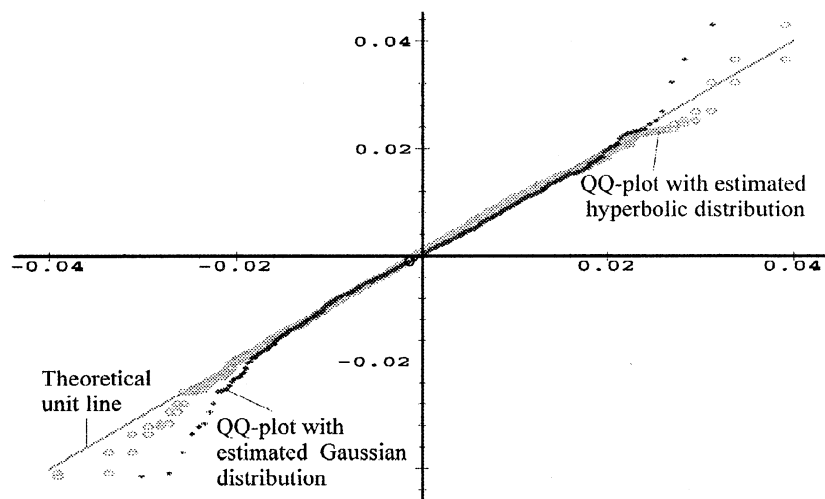


Figure 1. QQ-plot of DAX data against estimated normal ($\hat{\sigma} = 0.00917$) and hyperbolic distribution ($\hat{\zeta} = 2.16$ and $\hat{\delta} = 0.101$).

then the hyperbolic distribution tends to a Gaussian distribution with covariance matrix $C = \sqrt[4]{c^2} \Delta$. Estimating hyperbolic parameters of two-dimensional, simulated normal data resulted in ζ -estimates of 300–500. ζ -estimates for any of the data lay between 2 and 5, indicating that the data are far from normal.

We see that the VaR at level 5% is obtained quite well by both models. This may be due to the fit of the covariance estimator, that is, the choice of $\lambda = 0.94$ and $n = 74$. The VaR on lower levels is underestimated by the Gaussian model. To be able to judge the risk of a portfolio one needs more than just the VaR on only one level, for example, a second VaR, a category like heavy-tailed or flat-tailed or even the whole risk structure. Using the VaR estimated on Gaussian random variables will result in a misrepresentation of the real risk.

V. Conclusions

The symmetric hyperbolic distributions fit financial data better than the Gaussian distribution. Calculating the VaR on the basis of symmetric hyperbolic random variables, one gets more precise results for all levels p simultaneously. Using the normal distribution the estimators may only be adapted to a single level ($p = 5\%$ in the example) and yield biased results elsewhere. This does not allow a sufficient representation of the portfolio risk structure. The necessary calculations using hyperbolic distributions are nearly as easy and fast as in the Gaussian case.

Appendix: The Loglikelihood Function for ζ and δ

We present here the loglikelihood function l for ζ and δ after having estimated μ and Δ from a set of data $x = (x_i)_{i=1, \dots, n}$. We write $\omega_i = (x_i - \mu) \Delta^{-1} (x_i - \mu)^T \in \mathbb{R}$. In practice we use the estimates $\tilde{\omega}_i$ derived from the estimates $\tilde{\mu}$ and $\tilde{\Delta}$. Another abbreviation is $Q_\nu(x) \stackrel{\text{def}}{=} K_{\nu+1}(x)/K_\nu(x)$.

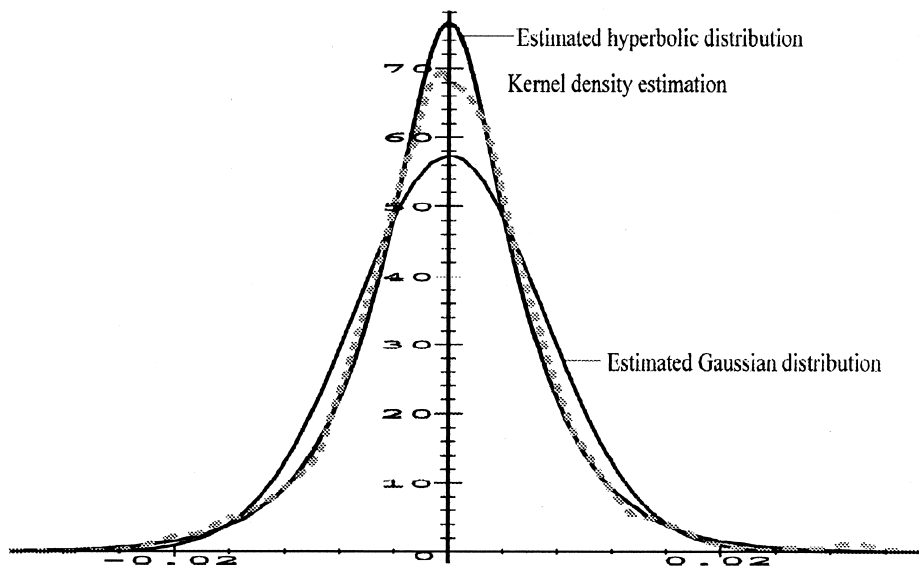


Figure 2. Plot of the kernel density estimation, the estimated hyperbolic ($\hat{\zeta} = 2.16$ and $\hat{\delta} = 0.101$) and the estimated Gaussian density ($\hat{\sigma} = 0.00696$) for the US-Dollar data.

$$\begin{aligned}
 l(x) &\stackrel{\text{def}}{=} -\ln \prod_{i=1}^n \frac{\zeta^{d-1}}{(2\pi)^{(1/2)d} \delta^d K_1(\zeta)} \frac{K_{1-(1/2)d} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right)}{\left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right)^{(1/2)d-1}} \\
 &= c - n(d \ln \zeta - d \ln \delta - \ln (K_1(\zeta)\zeta)) - \sum_{i=1}^n \ln \frac{K_{1-(1/2)d} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right)}{\left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right)^{(1/2)d-1}}. \quad (9)
 \end{aligned}$$

The derivatives—necessary for the Newton method to obtain the ML-estimate—with respect to ζ and δ are:

$$\begin{aligned}
 \frac{\partial l}{\partial \zeta} &= -n \left(\frac{d}{\zeta} + Q_1(\zeta) \right) + \sum_{i=1}^n Q_{(1/2)d-1} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right) \sqrt{1 + \frac{\omega_i}{\delta^2}}, \\
 \frac{\partial l}{\partial \delta} &= n \frac{d}{\delta} + \sum_{i=1}^n Q_{(1/2)d-1} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right) \zeta \frac{1}{2 \sqrt{1 + \frac{\omega_i}{\delta^2}}} \omega_i (-2\delta^3) \\
 &= n \frac{d}{\delta} - \sum_{i=1}^n Q_{(1/2)d-1} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right) \zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \frac{\omega_i}{\delta^3 + \omega_i \delta},
 \end{aligned}$$

$$\frac{\partial^2 l}{\partial \zeta^2} = -n \left(\frac{-d}{\zeta^2} + Q^2_1(\zeta) + \frac{1}{\zeta} Q_1(\zeta) - 1 \right) + \sum_{i=1}^n \left(1 + \frac{\omega_i}{\delta^2} \right) \left(Q^2_{(1/2)d-1} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right) - \frac{d-1}{\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}}} Q_{(1/2)d-1} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right) - 1 \right),$$

$$\begin{aligned} \frac{\partial^2 l}{\partial \delta^2} &= -n \frac{d}{\delta^2} - \sum_{i=1}^n \left(Q_{(1/2)d-1} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right) \zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \frac{-\omega_i(3\delta^2 + \omega_i)}{(\delta^3 + \omega_i\delta)^2} \right. \\ &\quad \left. + \left(Q^2_{(1/2)d-1} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right) - \frac{d-2}{\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}}} Q_{(1/2)d-1} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right) - 1 \right) \right. \\ &\quad \left. \zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \zeta \frac{-2\omega_i\delta^3}{2\sqrt{1 + \frac{\omega_i}{\delta^2}}} \frac{\omega_i}{\delta^3 + \omega_i\delta} \right) \\ &= -n \frac{d}{\delta^2} + \sum_{i=1}^n \frac{\zeta^2 \omega_i^2}{\delta^4(\delta^2 + \omega_i)} \left(Q^2_{(1/2)d-1} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right) - \left(\frac{d-2}{\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}}} - \frac{3\delta^2 + \omega_i}{\omega_i \zeta \sqrt{1 + \frac{\omega_i}{\delta^2}}} \right) \right. \\ &\quad \left. \cdot Q_{(1/2)d-1} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right) - 1 \right) \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 l}{\partial \delta \partial \zeta} &= - \sum_{i=1}^n \left(Q^2_{(1/2)d-1} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right) - \frac{d-2}{\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}}} Q_{(1/2)d-1} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right) - 1 \right) \\ &\quad \zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \sqrt{1 + \frac{\omega_i}{\delta^2}} \frac{\omega_i}{\delta^3 + \omega_i\delta} = - \sum_{i=1}^n \frac{\omega_i \zeta}{\delta^3} \left(Q^2_{(1/2)d-1} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right) \right. \\ &\quad \left. - \frac{d-2}{\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}}} Q_{(1/2)d-1} \left(\zeta \sqrt{1 + \frac{\omega_i}{\delta^2}} \right) - 1 \right). \end{aligned}$$

The ML-estimation routine was written in C++. With values $\{\omega_i\}_{i=1 \dots 1000}$ as input it takes roughly 3 minutes to perform the estimation on a Pentium 133.

An Algorithm for Creating Random Variables from a Generalized Inverse Gaussian Distribution with Parameters 1 , δ^2 , and ζ^2

This algorithm is based on Atkinson (1982). To maximize the efficiency of the algorithm one first needs to find two numbers $s_1 \geq 0$ minimizing

$$f_1 \stackrel{\text{def}}{=} \frac{1}{s_1} \left(\exp\left(\frac{s_1 \delta}{\zeta}\right) - 1 \right) \exp(-\delta \sqrt{\zeta^2 + 2s_1})$$

and $s_2 \in [-\zeta^2, 0]$ minimizing

$$f_2 \stackrel{\text{def}}{=} \frac{1}{s_2} \exp\left(\frac{s_2 \delta}{\zeta}\right) \exp(-\delta \sqrt{\zeta^2 + 2s_2}).$$

The efficiency of the algorithm is reciprocal to $f_1 + f_2$. Due to Atkinson a numerical approximation of the minima is sufficient for an acceptable performance. However in most cases even constants like $s_1 = \zeta^2/2$ and $s_2 = -\zeta^2/2$ are acceptable.

begin

create independent random variables u and v uniformly distributed on $(0, 1)$

$$\text{if } u \leq \frac{f_1}{f_1 + f_2}$$

then

$$x = \frac{1}{s_1} \ln(1 + us_1(f_1 + f_2) \exp(\delta \sqrt{\zeta^2 + 2s_1}))$$

$$\text{if } \ln(v) \leq -\frac{\zeta^2 + (\zeta^2 + 2s_1)x^2}{2x} + \delta \sqrt{\zeta^2 + 2s_1}$$

then end

else goto begin

fi

else

$$x = \frac{1}{s_2} \ln(-s_2(1 - u)(f_1 + f_2) \exp(\delta \sqrt{\zeta^2 + 2s_2}))$$

$$\text{if } \ln(v) \leq -\frac{\zeta^2 + (\zeta^2 + 2s_2)x^2}{2x} + \delta \sqrt{\zeta^2 + 2s_2}$$

then end

else goto begin

fi

fi

Then x is distributed according to the generalized inverse Gaussian distribution with parameters 1, δ^2 and ζ^2 . So if Y is a d -dimensional normal random vector, μ a location vector and Φ the Cholesky decomposition of Δ , then the random vector $X = \sqrt{x}Y\Phi^T + \mu$ is distributed according to $E_d(\zeta, \delta, \mu, \Delta)$.

The author would like to thank Prof. Dr. P. J. Huber, Dr. E. Tabakis, and two unknown referees for valuable comments.

References

- Atkinson, A. C. 1982. The simulation of generalized inverse Gaussian and hyperbolic random variables. *SIAM Journal of Scientific and Statistical Computations* 3:502–515.
- Bagnold, R. A., and Barndorff-Nielsen, O. E. 1980. The pattern of natural size distribution. *Sedimentology* 27:199–207.
- Barndorff-Nielsen, O. E. 1977. Exponentially decreasing distributions for the logarithm of particle size. *Proc. R. Soc. London A*, 353:401–419.
- Barndorff-Nielsen, O. E., Blæsild, P., Jensen, J. L., and Sørensen, M. 1985. The fascination of sand. In *A Celebration of Statistics* (A. C. Atkinson and S. E. Fienberg, eds.) New York: Springer, pp. 57–87.
- Barndorff-Nielsen, O. E., Jensen, J. L., and Sørensen, M. 1989. Wind shear and hyperbolic distributions. *Meteorology*, 49:417–431.
- Bauer, C. 1998. Value-at-risk bei hyperbolischer Verteilung. Master's thesis, Bayreuth.
- Bibby, B. M., and Sørensen, M. 1997. A hyperbolic diffusion model for stock prices. *Finance and Stochastics*, 1(1):25–41.
- Blæsild, P. 1981. The two-dimensional hyperbolic distribution and related distributions, with an application to Johannsen's bean data. *Biometrika* 68(1):251–263.
- Blæsild, P. 1990a. *Hyperbolic Distributions: Cumulants, Skewness and Kurtosis*, volume 209 of *Research Reports*. Department of Theoretical Statistics, Institute of Mathematics University of Aarhus, Aarhus.
- Blæsild, P. 1990b. *The Shape Cone of the d-Dimensional Hyperbolic Distribution*, volume 208 of *Research Reports*. Department of Theoretical Statistics, Institute of Mathematics University of Aarhus, Aarhus.
- Blæsild, P., and Sørensen, M. K. (1992). 'hyp'—a Computer Program for Analyzing Data by Means of the Hyperbolic Distribution, volume 248 of *Research Reports*. Department of Theoretical Statistics, Institute of Mathematics University of Aarhus, Aarhus.
- Bollerslev, T., Chou, R. Y., and Kroner, K. F. 1992. ARCH modeling in finance. *Journal of Econometrics*, 52:5–59.
- Bühler, W., Korn, O., and Schmidt, A. 1998. Ermittlung von Eigenkapitalanforderungen mit Internen Modellen. *Der Betriebswirt*, 58(1):64–85.
- Eberlein, E., and Keller, U. 1995. Hyperbolic distributions in finance. *Bernoulli: Official Journal of the Bernoulli Society of Mathematical Statistics and Probability*, 1(3):281–299.
- Fang, K.-T., and Zhang, Y.-T. 1990. *Generalized Multivariate Analysis*. Beijing: Science Press.
- Haaf, H. 1996. Mathematische Beschreibung der δ - Γ -Methode. Unpublished working paper.
- Longerstaey, J., and Zangari, P. 1995. *Five Questions about RiskMetrics™*. New York: J. P. Morgan, <http://www.jpmorgan.com/RiskManagement/RiskMetrics>.
- Longerstaey, J., Zangari, P., Finger, C. C., and Howard, S. 1996. *RiskMetrics™—Technical*

Dokument. New York: J. P. Morgan, 4th edition. <http://www.jpmorgan.com/RiskManagement/RiskMetrics/RiskMetrics.html>.

Poppensieker, T. 1997. *Strategisches Risikomanagement in deutschen Großbanken*. Deutscher Universitäts Verlag, Wiesbaden.

Stangier, W. 1994. *Effiziente Schätzung der Wahrscheinlichkeitsdichte durch Kerne*, volume 24 of *Arbeiten zur Angewandten Statistik*. Physika Verlag, Würzburg.