RISK MANAGEMENT WITH GENERALIZED HYPERBOLIC DISTRIBUTIONS

Wenbo Hu
Bell Trading
Chicago, IL, USA
email: wenbohu@yahoo.com

Alec Kercheval
Department of Mathematics
Florida State University
Tallahassee, FL, USA
email: kercheval@math.fsu.edu

ABSTRACT
We examine certain Generalized Hyperbolic (GH) distributions for modeling equity returns, compared to usual Normal distributions. We describe these GH distributions and some of their properties, and test them against six years of daily S&P500 index prices. We estimate Value-at-Risk from calibrated distributions, and show that the Normal distribution leads to VaR estimates that significantly underestimate the realized empirical values, while the GH distributions do not. Of several GH distribution families considered, the most successful is the skewed-t distribution.

KEY WORDS
Risk, VaR, Generalized Hyperbolic distributions, skewed-t.

1 Introduction

Financial risk management requires an understanding of the range of possible uncertain future returns. Quantitatively this relies on the use of probability distributions to model these uncertain return outcomes. It has been traditional, mostly for reasons of technical convenience, to use Normal distributions for this purpose, calibrating the parameters (means, covariances) to available data. However, we now know two things: (1) the Normal distribution is not a very good model for asset returns, especially in the tails, and (2) understanding of other probability distributions has progressed to the point where they can be practically used to model returns.

The so-called "stylized facts" of observed equity returns enjoy general agreement these days. Among them are:

- actual return distributions appear fat-tailed (compared to Normal), and skewed
- volatility appears time-varying and clustered
- returns are serially uncorrelated, but squared returns are serially correlated.

It is no longer necessary to ignore these facts in practical risk modeling applications. In this paper we describe the use of Generalized Hyperbolic (GH) distributions for equity risk management. These distributions were introduced in [1] in other contexts, and in [2] in a financial context. We will especially focus on a specific subfamily of GH known as the skewed-t distributions, generalizations of the usual t distributions, and championed by McNeil, et. al. in [3]. We argue that the multivariate skewed-t distribution is preferable to the Normal in equity risk management applications. More details of this analysis can be found in [4] and [5]. Also see [6], [7], [8].

A risk model also requires, in addition to the choice of distribution family, a way to quantify the level of risk. The Markowitz approach to portfolio management has been to use standard deviation of the returns distribution as the risk measure. Other choices, such as Value-at-Risk (VaR), or Expected Shortfall (ES, also called Conditional VaR), have been studied extensively since the advent of the concept of a coherent risk measure in [9]. However, the choice of risk measure is less important for portfolio management than the choice of distribution family. This is due in part to a result of [10] showing that, for elliptic distributions, the portfolios on the efficient frontier do not depend on the choice of risk measure. See also [8].

For portfolio management, the practical utility of non-normal distributions like GH requires two things:

1. There must be a fast algorithm for calibrating the parameters to data, and
2. the distribution family must be closed under linear combinations – the Portfolio Property.

The first requirement is satisfied for the GH family because of the EM algorithm; see [4] for details. The second requirement is also satisfied for GH – see below.

In this paper we examine the case for GH with equity index returns. Further work will examine the portfolio optimization problem.

2 Mean-Variance Mixture Distributions

The Generalized Hyperbolic distributions are part of a larger family with nice properties called the Normal Mean-Variance Mixture distributions.

Definition 2.1 Normal Mean-Variance Mixture. *The d-dimensional random variable $\mathbf{X}$ is said to have a multivari-*

ate normal mean-variance mixture distribution if

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + W\boldsymbol{\gamma} + \sqrt{W}A\mathbf{Z}, \quad where \tag{1}$$

1. $\mathbf{Z} \sim N_k(\mathbf{0}, \mathbf{I_k})$*, the standard* k*-dimensional Normal distribution,*
2. $W \geq 0$ *is a positive, scalar-valued r.v. which is independent of* $\mathbf{Z}$*,*
3. $A \in \mathbb{R}^{d\times k}$ *is a matrix, and*
4. $\boldsymbol{\mu}$ *and* $\boldsymbol{\gamma}$ *are parameter vectors in* $\mathbb{R}^d$*.*

From the definition, we can see that

$$\mathbf{X} \mid W \sim N_d(\boldsymbol{\mu} + W\boldsymbol{\gamma}, W\Sigma), \tag{2}$$

where $\Sigma = AA'$. We easily obtain the following moment formulas from the mixture definition:

$$E(\mathbf{X}) = \boldsymbol{\mu} + E(W)\boldsymbol{\gamma}, \tag{3}$$

$$COV(\mathbf{X}) = E(W)\Sigma + var(W)\boldsymbol{\gamma}\boldsymbol{\gamma}', \tag{4}$$

when the mixture variable W has finite variance $var(W)$. The mixture variable W can be interpreted as a shock that changes the volatility and mean of the normal distribution.

If the mixture variable W is generalized inverse gaussian (GIG) distributed (see below), then $\mathbf{X}$ is said to have a generalized hyperbolic distribution (GH). The GIG distribution has three real parameters, λ, χ, ψ, and we write $W \sim N^-(\lambda, \chi, \psi)$ when W is GIG.

Definition 2.2 Generalized Inverse Gaussian distribution (GIG**).** *The random variable* X *is said to have a generalized inverse gaussian (GIG) distribution if its probability density function is* $h(x; \lambda, \chi, \psi) =$

$$\frac{\chi^{-\lambda}(\sqrt{\chi\psi})^{\lambda}}{2K_\lambda(\sqrt{\chi\psi})} x^{\lambda-1} exp\left(-\frac{1}{2}(\chi x^{-1} + \psi x)\right), \tag{5}$$

for $x > 0$*, and where* $\chi, \psi > 0$*, and* K_λ *is a modified Bessel function of the third kind with index* λ*.*

Hence the multivariate generalized hyperbolic distribution depends on three real paramters λ, χ, ψ, two parameter vectors $\boldsymbol{\mu}$ (the location parameter) and $\boldsymbol{\gamma}$ (the skewness parameter) in $\mathbb{R}^d$, and a $d \times d$ positive semidefinite matrix Σ. We write

$$\mathbf{X} \sim GH_d(\lambda, \chi, \psi, \boldsymbol{\mu}, \boldsymbol{\gamma}, \Sigma).$$

If $\boldsymbol{\gamma} = \mathbf{0}$, then $\mathbf{X}$ is said to have a symmetric generalized hyperbolic distribution and it is in that case elliptical.

2.1 Some Special Cases

Hyperbolic distributions:

If $\lambda = 1$, we get the multivariate generalized hyperbolic distribution whose univariate margins are one-dimensional hyperbolic distributions. If $\lambda = (d+1)/2$, we get the d-dimensional hyperbolic distribution. However, its marginal distributions are no longer hyperbolic.

The one dimensional hyperbolic distribution is widely used in the modelling of univariate financial data, for example in [2].

Normal Inverse Gaussian distributions (NIG):

If $\lambda = -1/2$, then the distribution is known as normal inverse gaussian (NIG). NIG is also commonly used in the modelling of univariate financial returns.

Variance Gamma distribution (VG):

If $\lambda > 0$ and $\chi = 0$, then we get a limiting case known as the variance gamma distribution.

Skewed t Distribution:

If $\lambda = -\nu/2, \chi = \nu$ and $\psi = 0$, we get a limiting case which is called the skewed-t distribution by [11], because it generalizes the usual Student t distribution, obtained from the skewed-t by setting the skewness parameter $\boldsymbol{\gamma} = 0$. This distribution is denoted $SkewedT_d(\nu, \boldsymbol{\mu}, \Sigma, \boldsymbol{\gamma})$.

The Student t distribution is widely used in modelling univariate financial data since we can model the heaviness of the tail by controlling the degree of freedom ν. It can be used in the modelling of multivariate financial data too since the EM algorithm can be used to calibrate it (see [8] and [4]).

It is also widely used to model dependence by creating a Student t copula from the Student t distribution. The Student t copula is popular in the modelling of financial correlations since it is upper tail and lower tail dependent and is easy to calibrate. However, the Student t copula is symmetric and exchangeable, which are potential disadvantages. E.g. financial events appear to crash together more often than boom together, so that the upper and lower tail dependence need not be equal. This, in part, motivates the use of the skewed-t distribution. See [3].

2.2 The Portfolio Property

A great advantage of the generalized hyperbolic distributions with this parametrization is they are closed under linear transformation (see [8]).

Theorem 2.3 Linear Transformations of Generalized Hyperbolic Distributions. *If* $\boldsymbol{X} \sim GH_d(\lambda, \chi, \psi, \boldsymbol{\mu}, \Sigma, \boldsymbol{\gamma})$ *and* $\boldsymbol{Y} = B\boldsymbol{X} + \boldsymbol{b}$ *where* $B \in \mathbb{R}^{k\times d}$ *and* $\boldsymbol{b} \in \mathbb{R}^k$*, then* $\boldsymbol{Y} \sim GH_k(\lambda, \chi, \psi, B\boldsymbol{\mu} + \boldsymbol{b}, B\Sigma B', B\boldsymbol{\gamma})$*.*

Here we state as special cases the versions for our primary interest, the skewed t distributions.

Theorem 2.4 Linear Transformations of Skewed t Distributions.

If $\boldsymbol{X} \sim SkewedT_d(\nu, \boldsymbol{\mu}, \Sigma, \boldsymbol{\gamma})$ *and* $\boldsymbol{Y} = B\boldsymbol{X} + \boldsymbol{b}$ *where* $B \in \mathbb{R}^{k\times d}$ *and* $\boldsymbol{b} \in \mathbb{R}^k$*, then* $\boldsymbol{Y} \sim SkewedT_k(\nu, B\boldsymbol{\mu} + \boldsymbol{b}, B\Sigma B', B\boldsymbol{\gamma})$.

Corollary 2.5 Portfolio Property. *If* $B = \boldsymbol{\omega}^T = (\omega_1, \cdots, \omega_d)$*, and* $\mathbf{b} = \mathbf{0}$*, then the portfolio* $y = \boldsymbol{\omega}^T \mathbf{X}$ *is a one dimensional skewed-t distribution, and*

$$y \sim SkewedT_1(\nu, \boldsymbol{\omega}^T\boldsymbol{\mu}, \boldsymbol{\omega}^T\Sigma\boldsymbol{\omega}, \boldsymbol{\omega}^T\boldsymbol{\gamma}) \quad (6)$$

This corollary also shows that the marginal distributions are automatically obtained once we have calibrated the multivariate distributions, i.e., $X_i \sim SkewedT_1(\lambda, \chi, \psi, \mu_i, \Sigma_{ii}, \gamma_i)$.

3 One dimensional VaR: $GARCH$ and GH

Value-at-Risk has been the most widely used risk measure in the banking industry. In the remainder of this paper we look at historical data for the S&P500 index to examine the quality of out-of-sample VaR forecasts using a Normal model vs. other GH distributions.

3.1 Value at Risk (VaR)

Definition 3.1 Value at Risk (VaR). *VaR at confidence level $\alpha \in (0, 1)$ for loss L of a security or a portfolio is defined to be*

$$VaR_\alpha = inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\}, \quad (7)$$

where F is the distribution function of loss.

If the loss distribution function F is strictly increasing, then $VaR_\alpha = F^{-1}(\alpha)$. In practice, the confidence level ranges from 95 through 99.5%, though the Basel committee recommends 99%.

3.2 Data

We use 4,108 observations of adjusted daily close prices for the S&P500 index, from 4/18/1989 to 7/29/2005. The daily close prices are converted to daily negative log returns, and we wish to calibrate our model distributions via maximum likelihood.

This approach assumes the time series is approximately independent and identically distributed $(i.i.d)$, which is contrary to the "stylized facts" mentioned above.

Indeed, if we use most recent 1500 daily negative log return data of SP500 to plot the sample autocorrelation function (ACF), in Figure 1, we can see that the ACF of the negative log return series shows little evidence of serial correlation, while the ACF of the squared log return series does show evidence of serial dependence. A $GARCH$ model can be introduced to model the persistence in the volatility and filter the returns.

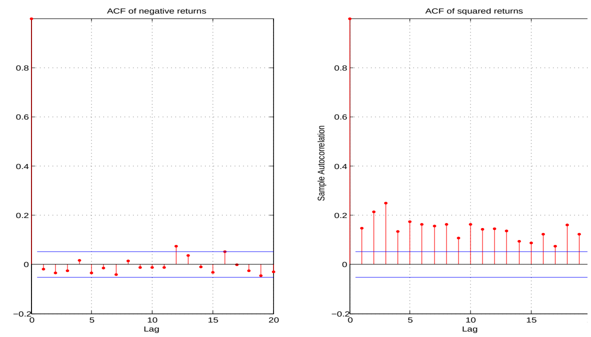

Figure 1. Correlograms for SP500 negative log return series

3.3 $GARCH$ Filter

Definition 3.2 $GARCH(1, 1)$ **process**. *Let Z_t be standard white noise $SWN(0, 1)$. The process (X_t) is a $GARCH(1, 1)$ if it is covariance stationary and satisfies the following equations,*

$$X_t = \sigma_t Z_t, \quad (8)$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

where $\alpha_0 > 0$,$\alpha_1 \geq 0$, and $1 > \beta_j \geq 0$. The innovation Z_t is independent of $(X_s)_{s<t}$.

McNeil et. al. in [3] argues that a $Garch(1, 1)$ model with student t innovations is enough to remove the dependence in return series, and sometimes normal innovations are enough too.

We create a filtered return series by subtracting the mean μ_0 of the raw series, and then calibrating the $Garch(1, 1)$ parameters above. The filtered return series is then defined to be

$$\hat{X}_t = \frac{X_t - \mu_0}{\sigma_t} \quad (9)$$

and should be approximately $i.i.d.$

From figure 2, we can see that the ACF of both filtered return series and squared filtered return series for SP500 show little evidence of serial correlation. This filtered series can then be used for calibrating parameters of various model distributions; $i.i.d.$ samples from these distributions can then be defiltered to give model distributions for the serially dependent returns.

A QQ-plot can be used to compare the empirical quantiles with quantiles of a designated distribution. We

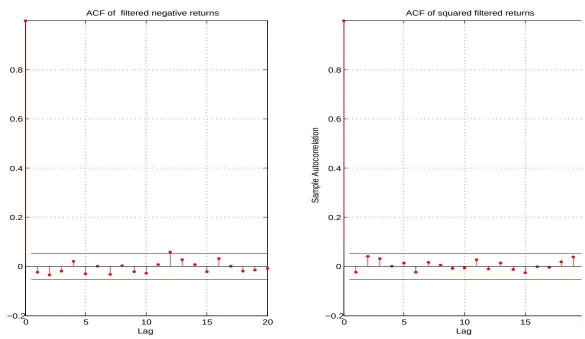

Figure 2. Correlograms for SP500 filtered negative log return series

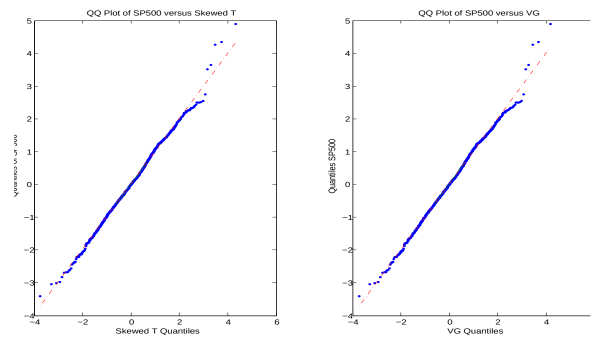

Figure 4. QQ-plot of S&P500 versus Skewed t and VG

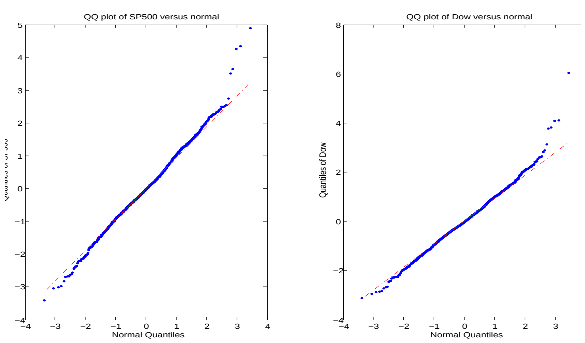

Figure 3. QQ-plot of S&P500 and Dow versus Normal

show the QQ-plot against a Gaussian distribution for filtered returns of both the S&P500 and Dow indices in Figure 3 – noting that both series have heavier tails than normal. In Figure 4 we can see that the skewed-t and VG distributions do much better in the tails.

3.4 Density estimation

The filtered data are now approximately $i.i.d.$ so that we can calibrate various generalized hyperbolic distributions using the EM algorithm (see [4]). We use the most recent 1500 observations to calibrate hyperbolic, NIG, skewed t, VG, student t, and Gaussian distributions. We report the calibrated parameters and the corresponding log likelihood in table 1 for the S&P500 index.

Model	λ or ν	χ	ψ	μ	σ^2	γ	Lh
Sk. t	13.88			-0.17	0.85	0.18	-20.7
VG	5.15		10.1	-0.20	0.97	0.23	-20.9
NIG	-0.5	4.0	6.78	-0.18	1.29	0.29	-20.9
Hyp.	1	103	0.1	-0.15	0.02	0.004	-21.8
t	13.95			0.03	0.86		-21.7
G.				0.04	1		-29.5

Table 1. Calibrated parameters of S&P500. Sk. t = skewed t, Hyp. = Hyperbolic, G. = Gaussian, Lh = log likelihood + 2100

From the viewpoint of maximum log likelihood, the skewed t has the largest log likelihood among all the distributions tested. NIG and hyperbolic are well known in the modelling of financial data; currently the skewed t is less common. In addition, the skewed t has the fewest number of parameters among the four generalized hyperbolic distributions so that it enjoys comparatively fast calibration[1].

[1] We use a laptop with centrino 1.3G, 1GB PC2700 memory. Software is Matlab R14. Under the same settings, skewed t needs 280 seconds, NIG needs 290 seconds, VG needs 390 seconds, while hyperbolic needs 630 seconds.

4 Backtesting of VaR

After we calibrate the filtered negative log return series to generalized hyperbolic distributions, we can calculate the α quantile, $z_\alpha = F^{-1}(\alpha)$, for the filtered negative log return series, where F is some distribution function. If we are standing at time t, we then de-filter the value at risk (VaR) of filtered returns to estimate the VaR of the negative log return at time $t+1$,

$$\hat{VaR}_\alpha(X_{t+1}|\mathcal{F}_t) = \hat{\sigma}_{t+1} z_\alpha + \mu_0, \tag{10}$$

where σ_{t+1} can be forecasted by using equation 8 and it is known at time t.

We have 4108 observations for both S&P and Dow index. Results for the Dow are similar, so we report here the S&P results. We use the most recent $N = 3000$ observations to backtest VAR violations. For each day, we use the previous 1000 observations to train the generalized hyperbolic distributions. We recalibrate the model at each day by initializing the parameters using the previous day's values, and recalibrate the model by re-initializing the parameters every 400 days to avoid overestimation.

At time t+1, where t is from 1000 from 4000, we use $(x_{t-1000+1}, \ldots, x_{t-1}, x_t)$ to calibrate the generalized hyperbolic distributions and estimate $VAR_\alpha(X_{t+1}|\mathcal{F}_t)$ for $\alpha = 0.95$, $\alpha = 0.975$, $\alpha = 0.99$, and $\alpha = 0.995$. A violation occurs if $x_{t+1} > V\hat{A}R_\alpha(X_{t+1}|\mathcal{F}_t)$. The number of total violations during those N tests is denoted by n. The actual violation frequency is n/N, while the expected violation probability, q, should be 0.05, 0.025, 0.01 and 0.005 respectively.

To evaluate the VaR backtest, we use a likelihood ratio statistic due to [12]. The null hypothesis is that the expected violation probability is equal to q. Under the null hypothesis, the likelihood ratio, given by

$$\begin{aligned} &-2[(N-n)\log(1-q) + n\log(q)] \\ &+2[(N-n)\log(1-n/N) + n\log(n/N)], \end{aligned}$$

is asymptotically $\chi^2(1)$ distributed.

We list the results of S&P 500 VaR backtesting in table 2. We calculate the actual violation probability at level q, where the expected violation probability, q, is 0.05, 0.025, and 0.01 respectively, and its corresponding p-value[2] for the likelihood ratio test. We test the normal distribution and four generalized hyperbolic distributions. At all levels, the generalized hyperbolic distributions pass the test, but the Normal distribution fails at all levels below 0.05.[3]

Model	0.05	p	0.025	p	0.01	p
N.	0.048	0.56	0.031	**0.032**	0.018	**0.0001**
Sk. t	0.049	0.74	0.026	0.73	0.009	0.71
VG	0.046	0.31	0.026	0.82	0.010	0.86
NIG	0.047	0.40	0.025	0.91	0.009	0.71
H.	0.045	0.23	0.024	0.72	0.010	0.85

Table 2. VaR violation backtesting for S&P500. N. = Normal, Sk. t = Skewed t, H. = Hyperbolic.

5 Conclusion

Normal distributions tend to underestimate the risk of extreme events. Generalized hyperbolic distributions have semi-heavy tails so they can be good candidates for risk management. We have used a $GARCH$ model to filter the negative return series to get approximately $i.i.d.$ filtered negative returns and forecast volatility. After we get $i.i.d.$ filtered negative returns, we calibrate our generalized hyperbolic distributions and calculate the α quantile. Using the forecasted volatility and α quantile for filtered negative return series, we can restore the VaR_α for the unfiltered returns. In backtesting VaR using both generalized hyperbolic distributions and normal distributions, we find that the generalized hyperbolic distributions pass the VaR test, while the normal distribution fails.

The special case called the skewed t distribution is not yet commonly used. However, it has the fewest parameters among all (non-Normal) generalized hyperbolic distributions examined, and the fastest observed calibration speed. In addition, it has the largest log likelihood among all examined generalized hyperbolic distributions, including student t, and Normal distributions. Therefore, we feel it is a promising candidate for future risk management applications.

[2] We call CHIDIST(x,1) in Excel to calculate the p-value, where x is the value of likelihood ratio statistic.

[3] When the p value is less than 0.05, we reject the null hypothesis.

6 References

[1] O.E. Barndorff-Nielsen, Exponentially decreasing distributions for the logarithm of the particle size, *Proceedings of the Royal Society, Series A: Mathematical and Physical Sciences*, 350, 1977, 401–419.

[2] E. Eberlein and U. Keller, Hyperbolic distributions in finance, *Bernoulli*, 1, 1995, 131–140.

[3] A. McNeil, R. Frey, P. Embrechts, *Quantitative risk management: concepts, techniques, and tools,* Princeton Univ. Press, 2005.

[4] W. Hu, Calibration of multivariate generalized hyperbolic distributions using the EM algorithm, with applications in risk management, portfolio optimization, and portfolio credit risk, *PhD dissertation*, Florida State University, 2005.

[5] W. Hu and A. Kercheval, Portfolio optimization for skewed-t returns, 2007 working paper.

[6] H. Aas and H. Haff, The generalized hyperbolic skew Student's t-distribution, *Journal of Financial Econometrics*, 4, 2006.

[7] S. Keel, F. Herzog, H. Geering, and N. Mirjolet, Opti-

mal portfolios with skewed and heavy-tailed distributions, *Proc. Third IASTED International Conference on Financial Engineering and Applications*, Cambridge, MA, 2006.
[8] W. Hu and A. Kercheval, The skewed-t distribution for portfolio credit risk, *Advances in Econometrics*, 2007, to appear.
[9] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, Coherent measures of risk, *Mathematical Finance*, 9, 1998, 203–228.
[10] P. Embrechts, A. McNeil, and D. Straumann, Correlation and dependency in risk management: properties and pitfalls, in *Risk Management: Value at Risk and Beyond*, (M. Dempster and H. Moffat, eds.), Cambridge Univ. Press, 2001.
[11] S. Demarta and A. McNeil, The t copula and related copulas, *International Statistical Review*, 73, 2005, 111–129.
[12] P. Kupiec, Techniques for verifying the accuracy of risk measurement models, *Journal of Derivatives, 1995, 73–84.*