THE UNIVERSITY OF TEXAS AT SAN ANTONIO, COLLEGE OF BUSINESS

Working Paper SERIES

Date June 25, 2013

WP # 0050MSS-202-2013

Comparing VaR Approximation Methods Which Use the First Four Moments as Inputs

Donald Lien, Christopher Stroud, Keying Ye
University of Texas at San Antonio

UTSA COLLEGE OF BUSINESS KNOWLEDGE FOR A NEW WORLD

ONE UTSA CIRCLE
SAN ANTONIO, TEXAS 78249-0631
210 458-4317 | BUSINESS.UTSA.EDU

Comparing VaR Approximation Methods Which Use the First Four Moments as Inputs

Donald Lien, Christopher Stroud, Keying Ye
University of Texas at San Antonio

June, 2013

Abstract

This paper compares four methods used to approximate value at risk (VaR) from the first four moments of a probability distribution: Cornish-Fisher (1938), Edgeworth (1907), Gram-Charlier (1902), and Johnson distributions (1949). We apply a procedure described by Chernozhukov et al. (2010) called the increasing rearrangement to the Cornish-Fisher, Edgeworth, and Gram-Charlier methods. Using the increasing rearrangement yields a single VaR approximation for any possible combination of skewness and kurtosis, and facilitates comparison of all four methods across the entire skewness-kurtosis space. Simulation results suggest that with enough data, the Johnson family yields the most accurate approximation on average.

1 Introduction

The contribution of this paper to the VaR literature is to compare four methods used to approximate value at risk (VaR) from the first four moments of a probability distribution: Cornish-Fisher (1938), Edgeworth (1907), Gram-Charlier (1902), and Johnson distributions (1949). This paper restricts its focus to approximation methods which take as inputs the first four moments due to their intuitive appeal. Statistical experts and non-experts alike can readily grasp how the mean, variance, skewness, and kurtosis affect the shape of a distribution. Consequently, potential users of these methods have a direct translation of what the inputs mean in terms of evaluating and managing risk.

Simulation work done by Simonato (2011) compares Cornish-Fisher, Gram-Charlier, and Johnson distributions. As described by Simonato, a difficulty in making meaningful comparisons between the three methods is that the set of skewness and kurtosis values for which each method yields a single VaR approximation differs across all three methods. Consequently, his paper focuses on comparing the Gram-Charlier method with Johnson distributions and the Cornish-Fisher method with Johnson distributions by restricting each to their respective valid regions.

Simonato (2011) informally compares methods outside their valid regions by keeping only the most conservative VaR approximation for cases that result in multiple solutions. This paper applies a procedure called the increasing rearrangement described by Chernozhukov et al. (2010) to yield a single VaR approximation across the full skewness-kurtosis space. Consequently, meaningful comparisons can be made between all four methods without restricting the range of skewness or kurtosis.

The principle findings of this paper are the following. When using the first four population moments as inputs, or when estimating sample moments from a large sample, Johnson distributions yield the best VaR approximations in terms of mean-square-error. This is in agreement with the results of Simonato (2011). However, for smaller sample sizes, the other methods may provide better approximations than Johnson distributions. Furthermore, at the smallest sample size ($n = 20$), Johnson distributions yielded the worst VaR approximation in terms of mean-square-error.

The paper is organized as follows. Section 2 provides an extensive literature review. Section 3 provides technical details of the increasing rearrangement. Section 4 details how to approximate VaR using the Gram-Charlier, Edgeworth, and Cornish-Fisher methods. Section 5 explains how to approximate VaR when using an increasing rearrangement with the Gram-Charlier, Edgeworth, and Cornish-Fisher methods. Section 6 details Johnson distributions and how they are used to approximate VaR. Sections 7 and 8 detail the simulation setup and results, respectively. Section 9 concludes the paper.

2 Literature Review

Among the earliest attempts to relax the normality assumption in calculating VaR from historical returns are two papers by Zangari. Zangari (1996a) suggests using quickly computable analytical methods which adjust for the effects of higher order moments. One suggested method is to adjust the normal quantiles to accommodate skewness and kurtosis using a Cornish-Fisher expansion. A later suggestion by Zangari (1996b) is a moment-matching procedure by transforming a normal distribution to a member of the Johnson family.

Simulation work done by Pichler et al. (1999) compares five analytical methods used to approximate VaR: the delta-normal approach, moment-matching to a normal distribution, Cornish-Fisher with four moments, Cornish-Fisher with six moments, and moment-matching to a Johnson distribution. They conclude that the latter three methods greatly outperform both the delta-normal approach and moment-matching to a normal distribution. Among these three methods, they find the Cornish-Fisher method with six moments to be the most accurate. The accuracy of the Cornish-Fisher method with four moments is nearly indistinguishable to moment-matching with a distribution from the Johnson family.

Simulation work by Lien et al. (2013) compares three analytical methods used to approximate VaR: the Cornish-Fisher method with four moments, the Liu approximation with the first four L-moments, and the Sillitto approximation truncated to an order of 15. The authors find that the Sillitto approximation yields the smallest approximation errors among the three methods, and the Liu approximation the largest approximation errors. They attribute the better performance of Sillitto to using higher order L-moments. Furthermore, they acknowledge that Cornish-Fisher may have equal or better performance than Sillitto if expanded out to the same order, but expressing the Cornish-Fisher expansion out to such a high order is difficult in practice. The ease at which Sillitto can accommodate higher order moments makes it a desirable approximation method, particularly when interest lies in capturing the effects of high order moments.

Many sources document that the Gram-Charlier, Edgeworth, and Cornish-Fisher expansions yield invalid estimated probability functions for certain values of skewness and kurtosis. Maillard (2012) describes the set of skewness and kurtosis values for which the Cornish-Fisher expansion yields a monotonic estimate of the quantile function. Spiring (2011) describes both sets of skewness and kurtosis values for which the Gram-Charlier and Edgeworth expansions yield valid estimates of the pdf.

Chernozhukov et al. (2010) present a method called the increasing rearrangement which restores monotonicity to the Edgeworth, Cornish-Fisher, and other related asymptotic expansions when the skewness or kurtosis fall outside their valid regions. The authors prove that the rearranged function is at least as good as the originally estimated function in L_p norm, $p \in (1, \infty)$. If the originally estimated function is invalid, they go on to show that the increasing rearrangement leads to a strictly better estimate in L_p norm.

Amédée-Manesme et al. (2012) utilize the increasing rearrangement with a Cornish-Fisher expansion to calculate direct real estate VaR. They calculate a rolling VaR over time of real estate returns in the UK. The authors credit the increasing rearrangement procedure as making a rolling calculation possible, because it overcomes the problem of sample skewness or kurtosis falling outside the valid region.

Simonato (2011) describes in detail how to approximate VaR and expected shortfall using Gram-Charlier, Cornish-Fisher, and Johnson distributions. For technical details of these three methods we direct you to Simonato's paper. Also included is a simulation study designed to compare the accuracy of the three methods in approximating VaR and expected shortfall. Simonato concludes that in addition to having no skewness or kurtosis restrictions, Johnson distributions provide better VaR approximations on average than both the Gram-Charlier method and Cornish-Fisher method.

3 The Increasing Rearrangement

Chernozhukov et al. (2010) define the increasing rearrangement as follows: Let $f(x)$ be a measurable function mapping $[0,1] \to \mathcal{R}$, and let $F_f(y) = \int 1\{f(u) \leq y\}du$ denote the distribution function of $f(X)$ when $X \sim U(0,1)$. The function

$$f^*(x) = \inf\left\{y \in \mathcal{R} : \left[\int 1\{f(u) \leq y\}du\right] \geq x\right\}$$

is called the increasing rearrangement of the function f. The rearrangement operator transforms the function f into its quantile function f^*. The authors give the following working definition of the increasing rearrangement; given values of the function $f(x)$ evaluated at x in a fine enough mesh of equidistant points, sort the values in increasing order. We use this working definition to compute the VaR with Edgeworth, Gram-Charlier, and Cornish-Fisher.

Proposition 1 of Chernozhukov et al. (2010) states the following; Let f_0 be a weakly increasing measurable function that we want to approximate, and let $\widehat{f}$ be an initial approximation to f_0, then for any $p \in [1, \infty]$, the rearrangement of $\widehat{f}$, denoted $\widehat{f}^*$, weakly reduces the estimation error:

$$\left(\int_{\mathcal{X}} |\widehat{f}^*(x) - f_0(x)|^p dx\right)^{\frac{1}{p}} \leq \left(\int_{\mathcal{X}} |\widehat{f}(x) - f_0(x)|^p dx\right)^{\frac{1}{p}}.$$

Furthermore, Corollary 1 of Chernozhukov et al. (2010) states that if f_0 is strictly increasing over $\mathcal{X}$, and $\widehat{f}$ decreases over a subset of $\mathcal{X}$ with positive measure, then the improvement in L_p norm for $p \in (1, \infty)$ is strict.

4 Three Expansions

In keeping with the notation used by Simonato (2011), let r_h denote the continuously compounded returns of a portfolio over an h year horizon. The VaR represents the unique return such that over the next h year period, $P(r_h < VaR) = p$. The VaR is calculated as

$$VaR = \alpha_h + \sigma_h \times \phi^{-1}(p) \tag{1}$$

where α_h and σ_h denote the expected value and standard deviation of the returns, respectively, and $\phi^{-1}(p)$ is the quantile function of the standardized returns.

The cdf of the standardized returns distribution can be estimated from the skewness (κ_3) and kurtosis (κ_4) of the returns distribution using the Edgeworth expansion

$$\phi_E(k|\kappa_3, \kappa_4) = \phi_N - \frac{\kappa_3}{6}[f_N \times (k^2-1)] - \frac{(\kappa_4-3)}{24}[f_N \times k(k^2-3)] + \frac{(f_N \times \kappa_3^2)}{72}(k^5 - 10k^3 + 15k),$$

or the Gram-Charlier expansion,

$$\phi_{GC}(k|\kappa_3, \kappa_4) = \phi_N - \frac{\kappa_3}{6}[f_N \times (k^2-1)] - \frac{(\kappa_4-3)}{24}[f_N \times k(k^2-3)].$$

Here ϕ_N and f_N represent the standard normal cdf and pdf evaluated at k, respectively. The quantile $k = \phi^{-1}(p)$, needed in equation (1), can be found through numerical search by finding the value k such that $\phi(k|\kappa_3, \kappa_4) = p$. The Gram-Charlier expansion only differs from the Edgeworth expansion by truncation of the last term. When skewness is zero, the Gram-Charlier and Edgeworth expansions are equal.

The quantile function of the standardized return distribution can be estimated from the skewness and kurtosis of the returns distribution using the Cornish-Fisher expansion

$$\phi_{CF}^{-1}(p|\kappa_3, \kappa_4) = \phi_N^{-1} + \frac{\kappa_3}{6}\left[(\phi_N^{-1})^2 - 1\right] + \frac{(\kappa_4-3)}{24}\left[(\phi_N^{-1})^3 - 3\phi_N^{-1}\right] - \frac{\kappa_3^2}{36}\left[2(\phi_N^{-1})^3 - 5\phi_N^{-1}\right].$$

Here ϕ_N^{-1} represents the standard normal quantile function evaluated at p. The resultant quantile estimate $\phi_{CF}^{-1}(p)$ can be directly plugged into equation (1) to approximate VaR.

5 Computing VaR With an Increasing Rearrangement

We begin by describing how to approximate VaR when applying an increasing rearrangement to the Gram-Charlier or Edgeworth estimated cdfs. To simplify notation let $\phi(k|\kappa_3, \kappa_4)$ denote either $\phi_E(k|\kappa_3, \kappa_4)$ or $\phi_{GC}(k|\kappa_3, \kappa_4)$. If $\phi(k|\kappa_3, \kappa_4)$ is not monotone, then an increasing rearrangement is applied to obtain a monotone estimate $\phi^*(k|\kappa_3, \kappa_4)$.

Approximating VaR with a rearranged Edgeworth or Gram-Charlier cdf can be performed as follows. First, subdivide a large enough interval $[a, b] \subset \mathcal{X}$ of the standardized return support $\mathcal{X}$ into a fine grid. Second, evaluate $\phi(k|\kappa_3, \kappa_4)$ for all $k \in [a, b]$. Third, sort these values from smallest to largest to yield the monotone cdf estimate $\phi^*(k|\kappa_3, \kappa_4) = p$. Fourth, from the sorted values, the desired quantile is the value such that $\phi^{*-1}(p|\kappa_3, \kappa_4) = k$. The VaR of the returns is approximated by plugging $\phi^{*-1}(p)$ into equation (1).

Approximating VaR with a rearranged Cornish-Fisher quantile function can be performed as follows. First, subdivide the interval $p \in (0, 1)$ into a fine grid. Second, depending on the grid size, note down which position of the grid contains to the desired VaR probability p, say the i^{th} position. Third, evaluate $\phi_{CF}^{-1}(p|\kappa_3, \kappa_4)$ for all $p \in (0, 1)$. Fourth, sort these values from smallest to largest to yield the monotone quantile function estimate $\phi^{*-1}(p|\kappa_3, \kappa_4) = k$. After sorting, the value now occupying the i^{th} position $\phi^{*-1}(p|\kappa_3, \kappa_4)$ can be directly plugged into equation (1) to yield the VaR approximation.

6 Johnson Distributions

The Johnson (1949) system of distributions is made up of three translation functions. Letting Z denote a standard normal random variable, the system considers transformations of the form

$$Z = \gamma + \delta f\left(\frac{Y-\mu}{\sigma}\right),$$

where the function $f(.)$ is given by

$$f(u) = \begin{cases} \ln(u) & \text{Lognormal } (S_L) \\ \ln\left(\frac{u}{1-u}\right) & \text{Bounded } (S_B) \\ \sinh^{-1}(u) & \text{Unbounded } (S_U). \end{cases}$$

In the above definition, μ is a location parameter, σ is a scale parameter, γ is a shape parameter, and δ is another shape parameter. Together, the S_L, S_B, and S_U distributions can accommodate all possible combinations of skewness and kurtosis. Furthermore, each distribution covers a unique region on the skewness-kurtosis space.

Both the skewness and kurtosis are nonlinear functions of the shape parameters γ and δ, and therefore finding the Johnson distribution with desired values of skewness and kurtosis must be done numerically. The Fortran algorithm by Hill et al. (1976) numerically finds the Johnson distribution with the desired mean, standard deviation, skewness, and kurtosis. This algorithm has been ported into R by McLeod et al. (2012), and is freely available for use in their *JohnsonDistribution* package.

Simonato (2011) describes in detail how to approximate VaR using Johnson distributions. With a mean of zero, standard deviation of one, and desired skewness and kurtosis, first use the algorithm above to find the functional form $f(.)$ and estimates $(\widehat{\mu}, \widehat{\sigma}, \widehat{\gamma}, \widehat{\delta})$. With these values in hand, and letting ϕ_N^{-1} denote the standard normal quantile function evaluated at p, the quantile needed for equation (1) is given by

$$\phi_J^{-1}(p) = \widehat{\mu} + \widehat{\sigma} f^{-1}\left(\frac{\phi_N^{-1} - \widehat{\gamma}}{\widehat{\delta}}\right).$$

7 Simulation Setup

We simulate asset returns using Merton's (1976) jump-diffusion framework which specifies that asset prices evolve continuously but may occasionally have discontinuous jumps. Formally, the jump-diffusion framework is defined as the sum of a Brownian motion and a compound Poisson process, both assumed independent. Letting R_h denote the log-returns of an asset over h years, the process can be written as

$$R_h = \left[\alpha - \frac{1}{2}\sigma^2 - \lambda\left(e^{\alpha_J + \frac{1}{2}\sigma_J^2} - 1\right)\right] h + \sigma Z + I(N_h \geq 1) \sum_{j=1}^{N_h} Y_j \ .$$

In the equation above, Z is a normal random variable with mean zero and variance h, N_h is a Poisson(λh) random variable, and Y_j are *iid* normal random variables with mean α_J and variance σ_J^2. The model parameters have the following interpretation: h denotes time in years, α denotes the expected annual asset returns, σ^2 denotes the annual variance of returns conditional on no jumps, λ is the annual arrival rate of jumps, α_J is the mean jump magnitude, and σ_J^2 is the variance of the jump magnitude.

The Merton jump-diffusion framework gives rise to a Gaussian mixture distribution with Poisson mixture weights. Analytic forms of the pdf and the first four moments are provided in the Appendix.

To cover a wide set of skewness and kurtosis combinations we first simulated 20,000 parameter sets $(h, \alpha, \sigma, \lambda, \alpha_J, \sigma_J)$. Following Simonato (2011), each parameter value was independently drawn from a uniform distribution with limits given by: $h \in [1/250, 20/250]$, $\alpha \in [0.01, 0.1]$, $\sigma \in [0.1, 0.5]$, $\lambda \in [1, 5]$, $\alpha_J \in [-0.1, 0.1]$, and $\sigma_J \in [0.01, 0.1]$. These 20,000 sets of parameters were used for the remainder of the study.

One departure of our study from that of Simonato (2011) is the probability p at which VaR is approximated. Rather than randomly selecting the probability p from the uniform distribution with limits $p \in [0.001, 0.05]$, we approximated VaR at the four values $p \in \{0.001, 0.01, 0.05, 0.1\}$ in each run. It is our opinion that the probability p at which VaR is approximated is of specific interest, and predetermined by the user. Consequently, potential users of these methods may be interested in how they compare at specific values of p.

The first part of the simulation study investigates how well each method performs when actual population quantities are used as inputs into the four approximation methods. Using the 20,000 parameter sets $(h, \alpha, \sigma, \lambda, \alpha_J, \sigma_J)$, we calculate 20,000 sets of the following population quantities: mean, variance, skewness, and kurtosis $(\alpha_h, \sigma_h^2, \kappa_3, \kappa_4)$. These population quantities are then directly input into each of the four VaR approximation methods. Formulas expressing $(\alpha_h, \sigma_h^2, \kappa_3, \kappa_4)$ as functions of the parameters $(h, \alpha, \sigma, \lambda, \alpha_J, \sigma_J)$ are provided in the Appendix.

The second part of the simulation study investigates how well each method performs when the sample estimates $(\widehat{\alpha}_h, \widehat{\sigma}_h^2, \widehat{\kappa}_3, \widehat{\kappa}_4)$ are used as inputs into the four approximation methods, across a range of sample sizes. This phase involved simulating 5 independent samples of sizes $n \in \{20, 40, 60, 125, 250\}$ at each of the 20,000 parameter sets. Within each of the five samples, the sample mean, sample variance, sample skewness, and sample kurtosis $(\widehat{\alpha}_h, \widehat{\sigma}_h^2, \widehat{\kappa}_3, \widehat{\kappa}_4)$ are estimated with their typical sample estimators. These sample estimates are then used as the inputs into each of the four VaR approximation methods.

8 Simulation Results

Let VaR_B denote the true baseline VaR under Merton's jump-diffusion framework for parameter values $(h, \alpha, \sigma, \lambda, \alpha_J, \sigma_J)$, and let $d_i = (VaR_{Est} - VaR_B)_i^2$ denote the squared differences between the estimated VaR and true baseline VaR. We explain how to find VaR_B in the

Appendix. As used in Simonato (2011), we compare the performance of the four methods through their root-mean-square-errors ($rmse$) computed as:

$$rmse = \sqrt{\frac{1}{n}\sum_{i=1}^{n} d_i} = \sqrt{\bar{d}}.$$

In addition to $rmse$ we also provide the mean-square-error (mse) which allows for construction of approximate 99.7% Monte-Carlo error bounds. The mse is computed as:

$$mse = \frac{1}{n}\sum_{i=1}^{n} d_i = \bar{d},$$

and the associated error bounds are given by:

$$mse \pm \frac{3\hat{\sigma}_n}{\sqrt{n}}, \text{ where } \hat{\sigma}_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2}.$$

Tables 1-6 provide the results. Each table is organized as follows. The second and third columns provide the $rmse$ and mse, respectively. The fourth and fifth columns provide the lower and upper approximate 99.7% Monte-Carlo error bounds in estimating the mse, respectively. The sixth column of each table ranks the four methods from best (lowest mse) to worst (highest mse).

Ranks were assigned based on the Monte-Carlo intervals. When two intervals overlap, both methods are assigned the same rank. A small number of cases had to be assigned two ranks because one interval overlapped with one interval on the high end and another on the low end, but all three had no common overlap. These represent an ambiguous result.

Table 1 shows the results when the population quantities $(\alpha_h, \sigma_h^2, \kappa_3, \kappa_4)$ are used as inputs to approximate VaR. At all four probabilities, $p \in \{0.001, 0.01, 0.05, 0.1\}$, the Johnson class of distributions dominates the other three methods in terms of mse.

Table 2, Table 3, and Table 4 provide the small-sample results for sample sizes $n = 20$, $n = 40$ and $n = 60$, respectively. When compared to Table 1, some of the more surprising results are the following. For $n = 20$, Johnson distributions yield the worst VaR approximation at every probability, $p \in \{0.001, 0.01, 0.05, 0.1\}$. With the exception of Johnson distributions at $n = 20$, all four methods are indistinguishable in approximating VaR at $p \in \{0.1\}$. For $n = 20$, $n = 40$ and $n = 60$, the Gram-Charlier method yields the best VaR approximation at the lowest probabilities, $p \in \{0.001, 0.01\}$, and the worst approximation at $p \in \{0.05\}$. For $n = 20$, $n = 40$ and $n = 60$, the Cornish-Fisher method is a top performer for $p \in \{0.05\}$.

Table 5 provides the sample results for $n = 125$. With a sample of this size, Johnson distributions are as good as any other method at every probability $p \in \{0.001, 0.01, 0.05, 0.1\}$. The Gram-Charlier method also continues to be a top performer at the lowest probabilities, $p \in \{0.001, 0.01\}$. The Cornish-Fisher method no longer shows up as a top performer for $p \in \{0.05\}$.

Table 6 provides the sample results for $n = 250$. This is the largest sample size used in the study, and the results are nearly identical to the population results of Table 1. The ranks for $p \in \{0.001, 0.05, 0.1\}$ of Table 6 are identical to the corresponding ranks in Table 1. For $p \in \{0.01\}$ the ranks are also quite similar to Table 1, except for the inability to distinguish between the performance of the Gram-Charlier and Edgeworth methods in Table 6. The population results of Table 1 can be interpreted in practical terms as a very large sample where sample moments have converged to their population counterparts.

9 Conclusion

This study compares four methods used to approximate VaR from the first four moments of a distribution. Applying the increasing rearrangement to the Gram-Charlier, Edgeworth, and Cornish-Fisher methods yields a unique VaR approximation for any possible value of skewness and kurtosis. This in turn facilitates more meaningful comparison among the four methods, since they are all valid over the entire skewness-kurtosis space.

When using true population quantities, or with a large enough sample, the Johnson family yields better approximations of VaR on average at all four probability levels considered. This however does not in general hold for small samples, particularly when computing VaR at the smallest probability levels.

For $p \in \{0.001, 0.01\}$ Gram-Charlier outperformed all other methods for sample sizes $n \leq 60$, and performed equally as well as Johnson distributions at $n = 125$. For $p \in \{0.05\}$ Cornish-Fisher was at least as good as any other method for sample sizes $n \leq 60$.

References

[1] C. Amédée-Manesme and F. Barthélémy. Cornish-Fisher Expansion for Real Estate Value at Risk. Presented to 2012 European Real Estate Society Annual Conference in Edinburgh - PhD Session, 2012.

[2] C. Charlier. Über dir Darstellung willkürlicher Funktionen. *Ark. Mat. Ast. och Fys.*, 2(20):1–35, 1902.

[3] V. Chernozhukov, I. Fernández-Val, and A. Galichon. Rearranging Edgeworth–Cornish–Fisher Expansions. *Economic Theory*, 42(2):419–435, 2010.

[4] S. R. Das and R. K. Sundaram. Of Smiles and Smirks: A Term Structure Perspective. *Journal of Financial and Quantitative Analysis*, 34(2), 1999.

[5] F.Y. Edgeworth. On the Representation of Statistical Frequency by a Series. *Journal of the Royal Statistical Society*, 70(1):102–106, March 1907.

[6] I. Hill. Algorithm AS 100: Normal-Johnson and Johnson-Normal Transformations. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 25(2):190–192, 1976.

[7] I. Hill, R. Hill, and R. Holder. Algorithm AS 99: Fitting Johnson Curves by Moments. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 25(2):180–189, 1976.

[8] N. L. Johnson. Systems of Frequency Curves Generated by Methods of Translation. *Biometrika*, 36(1/2):149–176, 1949.

[9] D. Lien, X. Yang, and K. Ye. Alternative Approximations to Value-at-Risk: A Comparison. *Communications in Statistics: Simulation and Computation*, Forthcoming 2013.

[10] D. Maillard. A User's Guide to the Cornish Fisher Expansion. *Available at SSRN 1997178*, 2012.

[11] A. McLeod and L. King. *JohnsonDistribution, R package version 0.24*, 2012.

[12] R. C. Merton. Option Pricing when Underlying Stock Returns are Discontinuous. *Journal of Financial Economics*, 3(1):125–144, 1976.

[13] S. Pichler and K. Selitsch. A Comparison of Analytical VaR Methodologies for Portfolios that Include Options. *Model Risk: Concepts, Calibration, and Pricing, Risk Publications, London*, pages 253–265, 1999.

[14] J.-G. Simonato. The Performance of Johnson Distributions for Computing Value at Risk and Expected Shortfall. *The Journal of Derivatives*, 19(1):7–24, 2011.

[15] F. Spiring. The Refined Positive Definite and Unimodal Regions for the Gram-Charlier and Edgeworth Series Expansion. *Advances in Decision Sciences*, 2011.

[16] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013.

[17] P. Zangari. A VaR Methodology for Portfolios that Include Options. *RiskMetrics Monitor*, First Quarter:4–16, 1996.

[18] P. Zangari. How Accurate is the Delta-Gamma Methodology? *RiskMetrics Monitor*, Third Quarter:12–29, 1996.

Appendix

The Merton jump-diffusion framework gives rise to a mixture of normal distributions with mixing weights governed by a Poisson distribution. The pdf of the continuously compounded returns over h years is given by

$$f_{R_h}(r) = \sum_{n=0}^{\infty} \frac{e^{-\lambda h}(\lambda h)^n}{n!} \frac{1}{\sqrt{2\pi(\sigma^2 h + n\sigma_J^2)}} e^{\frac{-\left[r - \left(\left[\alpha - \frac{1}{2}\sigma^2 - \lambda\left(e^{\alpha_J + \frac{1}{2}\sigma_J^2} - 1\right)\right] h + n\alpha_J\right)\right]^2}{2(\sigma^2 h + n\sigma_J^2)}}.$$

Given a particular probability p, the baseline VaR_B is computed by numerically solving the following expression for q

$$\int_{-\infty}^{q} \sum_{n=0}^{\infty} \frac{e^{-\lambda h}(\lambda h)^n}{n!} \frac{1}{\sqrt{2\pi(\sigma^2 h + n\sigma_J^2)}} e^{\frac{-\left[r - \left(\left[\alpha - \frac{1}{2}\sigma^2 - \lambda\left(e^{\alpha_J + \frac{1}{2}\sigma_J^2} - 1\right)\right] h + n\alpha_J\right)\right]^2}{2(\sigma^2 h + n\sigma_J^2)}} dr = p.$$

For the range of parameter values considered in this paper, much of the density is contained in the first few terms of the infinite sum above. All computations used to compute VaR_B were done by expanding $f_{R_h}(r)$ out 100 terms. This is many more terms than needed for an accurate computation.

Das and Sundaram (1999) derive the following closed form formulas for the mean, variance, skewness, and kurtosis over h years:

$$\alpha_h = h\left[\alpha - \lambda\left(e^{\alpha_J + \frac{1}{2}\sigma_J^2} - 1\right) - \frac{1}{2}\sigma^2 + \lambda\alpha_J\right]$$

$$\sigma_h^2 = h(\sigma^2 + \lambda\sigma_J^2 + \lambda\alpha_J^2)$$

$$\kappa_3 = \frac{1}{\sqrt{h}}\left[\frac{\lambda(\alpha_J^3 + 3\alpha_J\sigma_J^2)}{(\sigma^2 + \lambda\sigma_J^2 + \lambda\alpha_J^2)^{3/2}}\right]$$

$$\kappa_4 = 3 + \frac{1}{h}\left[\frac{\lambda(\alpha_J^4 + 6\alpha_J^2\sigma_J^2 + 3\sigma_J^4)}{(\sigma^2 + \lambda\sigma_J^2 + \lambda\alpha_J^2)^2}\right]$$

Tables

Table 1: VaR Computed from Population Quantities

$p = 0.001$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0369	1.365×10^{-3}	1.296×10^{-3}	1.433×10^{-3}	3
$VaR_E - VaR_B$	0.0334	1.117×10^{-3}	1.064×10^{-3}	1.170×10^{-3}	2
$VaR_{CF} - VaR_B$	0.0531	2.815×10^{-3}	2.464×10^{-3}	3.167×10^{-3}	4
$VaR_J - VaR_B$	0.0134	1.789×10^{-4}	1.627×10^{-4}	1.950×10^{-4}	1
$p = 0.01$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0121	1.454×10^{-4}	1.352×10^{-4}	1.555×10^{-4}	3
$VaR_E - VaR_B$	0.0107	1.137×10^{-4}	1.074×10^{-4}	1.200×10^{-4}	2
$VaR_{CF} - VaR_B$	0.0172	2.966×10^{-4}	2.556×10^{-4}	3.376×10^{-4}	4
$VaR_J - VaR_B$	0.0063	4.022×10^{-5}	3.778×10^{-5}	4.266×10^{-5}	1
$p = 0.05$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0105	1.094×10^{-4}	1.034×10^{-4}	1.155×10^{-4}	4
$VaR_E - VaR_B$	0.0091	8.214×10^{-5}	7.677×10^{-5}	8.750×10^{-5}	3
$VaR_{CF} - VaR_B$	0.0058	3.356×10^{-5}	3.126×10^{-5}	3.587×10^{-5}	2
$VaR_J - VaR_B$	0.0035	1.231×10^{-5}	1.136×10^{-5}	1.326×10^{-5}	1
$p = 0.10$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0068	4.649×10^{-5}	4.342×10^{-5}	4.956×10^{-5}	3
$VaR_E - VaR_B$	0.0044	1.936×10^{-5}	1.812×10^{-5}	2.060×10^{-5}	2
$VaR_{CF} - VaR_B$	0.0047	2.245×10^{-5}	2.047×10^{-5}	2.444×10^{-5}	2
$VaR_J - VaR_B$	0.0027	7.477×10^{-6}	6.791×10^{-6}	8.163×10^{-6}	1

Table 2: VaR Computed from Sample Estimates ($n = 20$)

$p = 0.001$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0855	7.306×10^{-3}	7.067×10^{-3}	7.545×10^{-3}	1
$VaR_E - VaR_B$	0.0873	7.623×10^{-3}	7.370×10^{-3}	7.876×10^{-3}	1
$VaR_{CF} - VaR_B$	0.1000	9.994×10^{-3}	9.679×10^{-3}	1.031×10^{-2}	2
$VaR_J - VaR_B$	0.1045	1.093×10^{-2}	1.060×10^{-2}	1.126×10^{-2}	3
$p = 0.01$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0492	2.416×10^{-3}	2.316×10^{-3}	2.516×10^{-3}	1
$VaR_E - VaR_B$	0.0517	2.677×10^{-3}	2.556×10^{-3}	2.797×10^{-3}	2
$VaR_{CF} - VaR_B$	0.0526	2.765×10^{-3}	2.657×10^{-3}	2.874×10^{-3}	2
$VaR_J - VaR_B$	0.0578	3.343×10^{-3}	3.204×10^{-3}	3.482×10^{-3}	3
$p = 0.05$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0335	1.124×10^{-3}	1.066×10^{-3}	1.181×10^{-3}	2
$VaR_E - VaR_B$	0.0298	8.896×10^{-4}	8.480×10^{-4}	9.311×10^{-4}	1
$VaR_{CF} - VaR_B$	0.0308	9.510×10^{-4}	9.070×10^{-4}	9.950×10^{-4}	1
$VaR_J - VaR_B$	0.0331	1.097×10^{-3}	1.045×10^{-3}	1.148×10^{-3}	2
$p = 0.10$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0243	5.886×10^{-4}	5.622×10^{-4}	6.151×10^{-4}	1
$VaR_E - VaR_B$	0.0243	5.922×10^{-4}	5.665×10^{-4}	6.180×10^{-4}	1
$VaR_{CF} - VaR_B$	0.0251	6.307×10^{-4}	6.018×10^{-4}	6.597×10^{-4}	1,2
$VaR_J - VaR_B$	0.0257	6.585×10^{-4}	6.291×10^{-4}	6.878×10^{-4}	2

Table 3: VaR Computed from Sample Estimates ($n = 40$)

$p = 0.001$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0702	4.929×10^{-3}	4.765×10^{-3}	5.094×10^{-3}	1
$VaR_E - VaR_B$	0.0711	5.054×10^{-3}	4.885×10^{-3}	5.223×10^{-3}	1
$VaR_{CF} - VaR_B$	0.0842	7.097×10^{-3}	6.803×10^{-3}	7.390×10^{-3}	2
$VaR_J - VaR_B$	0.0824	6.783×10^{-3}	6.554×10^{-3}	7.012×10^{-3}	2
$p = 0.01$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0381	1.449×10^{-3}	1.387×10^{-3}	1.511×10^{-3}	1
$VaR_E - VaR_B$	0.0405	1.644×10^{-3}	1.572×10^{-3}	1.717×10^{-3}	2
$VaR_{CF} - VaR_B$	0.0415	1.724×10^{-3}	1.650×10^{-3}	1.797×10^{-3}	2,3
$VaR_J - VaR_B$	0.0424	1.797×10^{-3}	1.723×10^{-3}	1.872×10^{-3}	3
$p = 0.05$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0256	6.568×10^{-4}	6.198×10^{-4}	6.937×10^{-4}	2
$VaR_E - VaR_B$	0.0232	5.370×10^{-4}	5.114×10^{-4}	5.626×10^{-4}	1
$VaR_{CF} - VaR_B$	0.0227	5.159×10^{-4}	4.920×10^{-4}	5.397×10^{-4}	1
$VaR_J - VaR_B$	0.0230	5.279×10^{-4}	5.033×10^{-4}	5.526×10^{-4}	1
$p = 0.10$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0175	3.056×10^{-4}	2.920×10^{-4}	3.192×10^{-4}	1
$VaR_E - VaR_B$	0.0173	2.987×10^{-4}	2.848×10^{-4}	3.126×10^{-4}	1
$VaR_{CF} - VaR_B$	0.0179	3.215×10^{-4}	3.064×10^{-4}	3.366×10^{-4}	1
$VaR_J - VaR_B$	0.0178	3.157×10^{-4}	3.014×10^{-4}	3.300×10^{-4}	1

Table 4: VaR Computed from Sample Estimates ($n = 60$)

$p = 0.001$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0630	3.975×10^{-3}	3.838×10^{-3}	4.111×10^{-3}	1
$VaR_E - VaR_B$	0.0631	3.983×10^{-3}	3.846×10^{-3}	4.120×10^{-3}	1
$VaR_{CF} - VaR_B$	0.0759	5.764×10^{-3}	5.493×10^{-3}	6.034×10^{-3}	3
$VaR_J - VaR_B$	0.0705	4.973×10^{-3}	4.800×10^{-3}	5.146×10^{-3}	2
$p = 0.01$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0327	1.072×10^{-3}	1.027×10^{-3}	1.117×10^{-3}	1
$VaR_E - VaR_B$	0.0347	1.203×10^{-3}	1.152×10^{-3}	1.253×10^{-3}	2
$VaR_{CF} - VaR_B$	0.0359	1.291×10^{-3}	1.233×10^{-3}	1.349×10^{-3}	2
$VaR_J - VaR_B$	0.0349	1.219×10^{-3}	1.171×10^{-3}	1.267×10^{-3}	2
$p = 0.05$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0219	4.799×10^{-4}	4.544×10^{-4}	5.054×10^{-4}	3
$VaR_E - VaR_B$	0.0199	3.972×10^{-4}	3.777×10^{-4}	4.167×10^{-4}	2
$VaR_{CF} - VaR_B$	0.0190	3.596×10^{-4}	3.436×10^{-4}	3.755×10^{-4}	1
$VaR_J - VaR_B$	0.0187	3.496×10^{-3}	3.341×10^{-4}	3.651×10^{-4}	1
$p = 0.10$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0146	2.145×10^{-4}	2.051×10^{-4}	2.239×10^{-4}	1
$VaR_E - VaR_B$	0.0143	2.057×10^{-4}	1.962×10^{-4}	2.151×10^{-4}	1
$VaR_{CF} - VaR_B$	0.0148	2.181×10^{-4}	2.081×10^{-4}	2.280×10^{-4}	1
$VaR_J - VaR_B$	0.0145	2.108×10^{-4}	2.013×10^{-4}	2.203×10^{-4}	1

Table 5: VaR Computed from Sample Estimates ($n = 125$)

$p = 0.001$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0528	2.785×10^{-3}	2.687×10^{-3}	2.884×10^{-3}	1
$VaR_E - VaR_B$	0.0517	2.668×10^{-3}	2.576×10^{-3}	2.761×10^{-3}	1
$VaR_{CF} - VaR_B$	0.0670	4.494×10^{-3}	4.120×10^{-3}	4.868×10^{-3}	2
$VaR_J - VaR_B$	0.0534	2.853×10^{-3}	2.739×10^{-3}	2.968×10^{-3}	1
$p = 0.01$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0255	6.507×10^{-4}	6.221×10^{-4}	6.792×10^{-4}	1
$VaR_E - VaR_B$	0.0267	7.113×10^{-4}	6.802×10^{-4}	7.423×10^{-4}	2
$VaR_{CF} - VaR_B$	0.0291	8.471×10^{-4}	7.908×10^{-4}	9.033×10^{-4}	3
$VaR_J - VaR_B$	0.0252	6.368×10^{-4}	6.095×10^{-4}	6.641×10^{-4}	1
$p = 0.05$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0174	3.044×10^{-4}	2.880×10^{-4}	3.208×10^{-4}	4
$VaR_E - VaR_B$	0.0160	2.549×10^{-4}	2.411×10^{-4}	2.686×10^{-4}	3
$VaR_{CF} - VaR_B$	0.0142	2.005×10^{-4}	1.915×10^{-4}	2.096×10^{-4}	2
$VaR_J - VaR_B$	0.0132	1.751×10^{-3}	1.674×10^{-4}	1.828×10^{-4}	1
$p = 0.10$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0114	1.289×10^{-4}	1.232×10^{-4}	1.345×10^{-4}	3
$VaR_E - VaR_B$	0.0106	1.120×10^{-4}	1.069×10^{-4}	1.170×10^{-4}	1,2
$VaR_{CF} - VaR_B$	0.0108	1.171×10^{-4}	1.120×10^{-4}	1.223×10^{-4}	2
$VaR_J - VaR_B$	0.0103	1.058×10^{-4}	1.012×10^{-4}	1.103×10^{-4}	1

Table 6: VaR Computed from Sample Estimates ($n = 250$)

$p = 0.001$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0459	2.105×10^{-3}	2.026×10^{-3}	2.185×10^{-3}	3
$VaR_E - VaR_B$	0.0437	1.911×10^{-3}	1.843×10^{-3}	1.979×10^{-3}	2
$VaR_{CF} - VaR_B$	0.0614	3.766×10^{-3}	3.370×10^{-3}	4.162×10^{-3}	4
$VaR_J - VaR_B$	0.0400	1.604×10^{-3}	1.538×10^{-3}	1.670×10^{-3}	1
$p = 0.01$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0205	4.204×10^{-4}	4.017×10^{-4}	4.391×10^{-4}	2
$VaR_E - VaR_B$	0.0209	4.370×10^{-4}	4.180×10^{-4}	4.560×10^{-4}	2
$VaR_{CF} - VaR_B$	0.0242	5.847×10^{-4}	5.379×10^{-4}	6.316×10^{-4}	3
$VaR_J - VaR_B$	0.0184	3.383×10^{-4}	3.241×10^{-4}	3.525×10^{-4}	1
$p = 0.05$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0146	2.132×10^{-4}	2.013×10^{-4}	2.251×10^{-4}	4
$VaR_E - VaR_B$	0.0131	1.729×10^{-4}	1.627×10^{-4}	1.830×10^{-4}	3
$VaR_{CF} - VaR_B$	0.0109	1.183×10^{-4}	1.127×10^{-4}	1.239×10^{-4}	2
$VaR_J - VaR_B$	0.0096	9.168×10^{-5}	8.771×10^{-5}	9.565×10^{-5}	1
$p = 0.10$	$rmse$	mse	$mse - \frac{3\hat{\sigma}_n}{\sqrt{n}}$	$mse + \frac{3\hat{\sigma}_n}{\sqrt{n}}$	rank
$VaR_{GC} - VaR_B$	0.0093	8.634×10^{-5}	8.218×10^{-5}	9.050×10^{-5}	3
$VaR_E - VaR_B$	0.0081	6.509×10^{-5}	6.214×10^{-5}	6.804×10^{-5}	2
$VaR_{CF} - VaR_B$	0.0084	6.978×10^{-5}	6.633×10^{-5}	7.323×10^{-5}	2
$VaR_J - VaR_B$	0.0075	5.559×10^{-5}	5.319×10^{-5}	5.799×10^{-5}	1