Variability in Modified Estimators of VaR and ES

Rohit Arora

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Science in Computational Finance and Risk Management

University of Washington

2016

Committee:

Douglas Martin

Eric Zivot

Program Authorized to Offer Degree:

Department of Applied Mathematics

University of Washington

Abstract

Variability in Modified Estimators of VaR and ES

Rohit Arora

Chair of the Supervisory Committee:
Professor Douglas Martin
Department of Applied Mathematics

Modified Value-at-Risk (mVaR) and Modified Expected Shortfall (mES) are risk estimators that can be calculated without modelling the distribution of asset returns. These modifided estimators use skewness and kurtosis corrections to normal distribution parametric VaR and ES formulas to reduce bias in risk measurement for non-normal return distributions. However, the use of skewness and kurtosis estimators that are needed to implement mVaR and mES can lead to highly inflated mVaR and mES estimator standard errors. To assess the degree of inflation we derive formulas for the large sample standard errors of mVaR and mES using multivariate delta method. Finally, we assess the goodness of our analytical result with small sample Monte-Carlo at Normal and t-ditributions. Our results show that projected standard errors in small sample underestimate the true standard error. Also, the effect is worse for mES than mVaR.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I wish to thank my supervisory committee for their time and effort. I am fortunate to be advised by Dr. Douglas Martin. His constant belief in me let me transcend barriers and grow as a researcher during my time at the University of Washington. I would like to thank him for his guidance and insightful comments. I would also like to thank Dr. Steven Golbeck for his mentorship and many fruitful discussions outside the classroom. Although my parents do not understand the technicalities of my work, they have been constant source of encouragement. Lastly, I wish to thank my wife Sonali for her unconditional support and patience during the last two years. None of this would have been possible without her.

DEDICATION

To my dear wife, Sonali

Chapter 1

INTRODUCTION

Value-at-risk (VaR) and expected shortfall (ES) are popular measures for assessing downside risk. Estimating VaR and ES using the assumption of a normal distribution for asset returns does not take into account skewness and kurtosis of the returns distribution that arise in practice. A well-known solution to this problem is to adjust the normal distribution quantile using a Cornish-Fisher asymptotic expansion to account for skewness and kurtosis in the asset returns distribution. The resulting risk measure is known as modified VaR (mVaR). Subsequently the general approach of using asymptotic expansions was applied to normal distribution ES to obtain a modified ES (mES) that takes into account skewness and kurtosis. Thus estimators of mVaR or mES are used in those frequently occurring situations where the risk manager does not know the distribution of returns and wants protection against non-normality of returns.

However relatively little is known about the estimation risk of mVaR and mES. Prior research on estimation risk of mVaR and mES mainly utilized bootstrap and Monte Carlo methods to obtain estimator standard errors and confidence intervals. We take a different approach in this thesis and use the delta method to derive large sample analytic formulas for mVaR and mES variances that are valid for any distribution having finite central moments of order eight. A similar approach was taken by Ardia and Boudt (2015) to compute the standard errors of the modified Sharpe ratio. When the asymptotic biases are zero or negligibly small, these formulas provide risk managers with a method for calculating finite sample standard error approximations for mVaR and mES in the usual way one uses asymptotic estimator variances to obtain finite sample standard errors.

A major reason for concern about the variability of mVaR and mES estimators is that their variability depends not only on mean and standard deviation estimators but also on skewness and kurtosis estimators. The latter will result in additional variability of mVaR and mES relative to that of parametric VaR and ES, even in the case where the asset returns are normally distributed and normal VaR and normal ES is used. However, in the case of fat-tailed distributions, the skewness and kurtosis estimators used in mVaR and mES can have hugely inflated variability resulting in much larger variability for mVaR and mES estimators than for those parametric VaR and ES estimators. With respect to our derivation of formulas for the asymptotic variance of mVaR and mES we note the following. mVaR and mES estimators are non-linear functions of third and fourth central moments, which increases the complexity of deriving the asymptotic distribution of these estimators. An added complexity in the derivation is that mES turns out to be a function of mVaR.

The remainder of this thesis is organized as follows. Chapter 2 briefly discusses parametric VaR and ES for normal and t-distributions. Chapter 3 uses multivariate delta method in deriving the large sample variance formulas of mVaR and mES and briefly discusses the validity of the results. Chapter 4 presents the efficacy of our results in finite sample using projected standard errors. Chapter 5 provides concluding comments and points out some further research in progress.

Chapter 2

VAR AND ES

To set the stage for later defining mVaR and mES, we briefly review VaR and ES for normal and t-distributions.

2.1 Parametric VaR and ES for Normal and t-distributions

Throughout we will be evaluating VaR and ES risk measures for returns distributions F that are continuous and strictly increasing, e.g., normal distributions and t-distributions. The inverse F^{-1} exists for such distributions and the $\gamma-$quantile $q_\gamma(F)$ is given by

$$q_\gamma(F) = F^{-1}(\gamma), \quad \gamma \in (0,1). \tag{2.1}$$

For our discussion of risk measures we will take loss as a negative and so for a tail probability γ the value at risk coincides with the above quantile.

It will be convenient to define

$$VaR_\gamma^{0,1} = F_{0,1}^{-1}(\gamma) \tag{2.2}$$

where $F_{0,1}$ is a distribution with with mean zero and standard deviation one. Then since a quantile functional is location and scale equivariant the value-at-risk for a distribution, $F_{\mu,s}$ with location parameter μ and scale parameter s is:

$$VaR_\gamma = \mu + s \cdot VaR_\gamma^{0,1}. \tag{2.3}$$

The general definition of expected shortfall is

$$ES_\gamma(F) = E_F\left[r | \, r \le VaR_\gamma(F)\right] \tag{2.4}$$

and we define

$$ES_{\gamma}^{0,1} = ES_{\gamma}(F_{0,1}). \tag{2.5}$$

Noting that expected shortfall is also location and scale equivariant the expected shortfall for the distribution with mean μ and scale parameter s is:

$$ES_{\gamma} = \mu + s \cdot ES_{\gamma}^{0,1}. \tag{2.6}$$

In the case of a normal distribution and a t-distribution with degrees of freedom ν greater than one the location parameter μ is the mean. The scale parameter s is equal to the standard deviation σ for a normal distribution, and for a t-distribution with $\nu > 2$ the standard deviation exists and is given by $\sigma = \sqrt{\nu/(\nu-2)} \cdot s$.

Table **??** displays the formulas for VaR_{γ} and ES_{γ} for normal distributions $\mathcal{N}(\mu, \sigma^2)$ and a t-distributions $t_{\nu}(\mu, s)$ with location parameter μ and scale parameter s, whose probability density is

$$t_{\nu}(x; \mu, s) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu} \cdot \Gamma\left(\frac{\nu}{2}\right) \cdot s} \left(1 + \frac{1}{\nu}\left(\frac{x-\mu}{s}\right)^2\right)^{-\frac{(1+\nu)}{2}} \tag{2.7}$$

and we write the standard t-density as $t_{\nu}^{0}(x) = t_{\nu}(x; 0, 1)$. In the table z_{γ} is the $\gamma-$quantile of the standard normal distribution, $\phi(x)$ is the standard normal density function, $q_{\nu,\gamma}$ is the $\gamma-$quantile of the $t_{\nu}(0, 1)$ distribution, and $h_{\nu,\gamma}$ is defined as

$$h_{\nu,\gamma} = \sqrt{\frac{\nu}{\nu-2}} \cdot t_{\nu-2}^{0}\left(\sqrt{\frac{\nu-2}{\nu}} \cdot q_{\nu,\gamma}\right). \tag{2.8}$$

Parametric VaR and ES estimators are obtained by replacing the unknown μ and σ with sample based estimators.

Chapter 3

MODIFIED VAR AND ES

3.1 mVaR and mES Estimators

A modified form of normal distribution VaR based on a Cornish–Fisher expansion was proposed by Zangari (1996) to take into account the effect of return distributions non-normality by adding correction terms based on the coefficient of skewness γ_1 and coefficient of excess kurtosis κ where

$$\gamma_1 = \frac{\mu_3}{\sigma^3}, \qquad \kappa = \frac{\mu_4}{\sigma^4} - 3.$$

with μ_3 and μ_4 the third and fourth central moments. The Cornish-Fisher expansion transforms the normal distribution quantile function z_γ to the quantile g_γ given by

$$g_\gamma = z_\gamma + \frac{1}{6}\left(z_\gamma^2 - 1\right)\gamma_1 - \frac{1}{36}\left(2z_\gamma^3 - 5z_\gamma\right)\gamma_1^2 + \frac{1}{24}\left(z_\gamma^3 - 3z_\gamma\right)\kappa \tag{3.1}$$

and the resulting modified VaR (mVaR) is

$$mVaR_\gamma = \mu + \sigma g_\gamma. \tag{3.2}$$

Motivated by the increased popularity of expected shortfall as a measure of risk Boudt et al. (2008) derived a modified version expected shortfall (mES), that takes into account returns skewness and kurtosis. Their derivation uses Edgeworth and Cornish-Fisher expansions of the density and quantile functions to obtain the following modification of expected shortfall

$$mES_\gamma = \mu - \frac{\sigma}{\gamma}\left(\begin{array}{c}\phi(g_\gamma) + \frac{1}{6}\left[\mathrm{I}^3 - 3\cdot\mathrm{I}\right]\gamma_1 + \frac{1}{72}\left[\mathrm{I}^6 - 15{\cdot}\mathrm{I}^4 + 45\cdot\mathrm{I}^2 - 15\phi(g_\gamma)\right]\gamma_1^2 \\ +\frac{1}{24}\left[\mathrm{I}^4 - 6\cdot\mathrm{I}^2 + 3\phi(g_\gamma)\right]\kappa\end{array}\right) \quad (3.3)$$

where $\mathrm{I}^k,\ k = 2,3,4,6$ are complicated sums of ratios of products involving $\phi(g_\gamma)$, $\Phi(g_\gamma)$ and powers of the modified standard normal quantile g_γ. The expressions for the I^k are provided in the Appendix where we show that the above expression for mES_γ can be simplified to the following form similar in style to the combined (3.1) and (3.2):

$$mES_\gamma = \mu - \frac{\sigma}{\gamma}\phi(g_\gamma)\left[1 + \frac{1}{6}g_\gamma^3\gamma_1 + \frac{1}{72}\left(g_\gamma^6 - 9g_\gamma^4 + 9g_\gamma^2 + 3\right)\gamma_1^2 + \frac{1}{24}\left(g_\gamma^4 - 2g_\gamma^2 - 1\right)\kappa\right]. \quad (3.4)$$

Estimators $\widehat{mVaR}_\gamma$ and $\widehat{mES}\gamma$ of $mVaR_\gamma$ and mES_γ respectively are obtained in the way the estimators are usually implemented in practice, namely by replacing μ and σ by the sample mean and sample standard deviation (the normal distribution MLE's) and replacing the coefficient of skewness γ_1 and excess kurtosis κ by their usual sample based estimators. Of course the latter are not MLE's.

3.2 Large Sample Variance of mVaR and mES Estimators

The general form of mVaR and mES can be written as

$$\begin{aligned}\rho &= \mu + C_0\sigma + C_1\sigma\gamma_1 + C_2\sigma\gamma_1^2 + C_3\sigma\kappa \\ &= \mathbf{x}'\mathbf{z} \qquad (3.5)\end{aligned}$$

where

$$\mathrm{x} = \left(\begin{array}{ccccc} 1, & C_0, & C_1, & C_2, & C_3\end{array}\right)^\top \quad (3.6)$$

is determined by (3.2) for mVaR and by (3.4) for mES, and where

$$\begin{aligned}\mathbf{z} &= \left(\mu,\ \sigma,\ \sigma\gamma_1,\ \sigma\gamma_1^2,\ \sigma\kappa\right)^\top \qquad (3.7)\\ &= \left(\begin{array}{ccccc}\mu, & \sqrt{\sigma^2}, & \frac{\mu_3}{\sigma^2}, & \frac{\mu_3^2}{\sqrt{(\sigma^2)^5}}, & \left(\frac{\mu_4}{\sqrt{(\sigma^2)^3}} - 3\sqrt{\sigma^2}\right)\end{array}\right). \qquad (3.8)\end{aligned}$$

The risk measure estimators have the form

$$\hat{\rho} = \mathbf{x}^\top \hat{\mathbf{z}} \tag{3.9}$$

where $\hat{\mathbf{z}}$ is obtained from either (3.7) or (3.8) by replacing the unknown parameters by their sample moment based estimators.

It is convenient to obtain the asymptotic variance of $\hat{\rho}$ via the delta rule by using $\hat{\mathbf{z}}$ based on (3.8), since then we can make use of asymptotic joint normality of the first four central moments that we describe shortly. Note that (3.5) and (3.8) give mVaR and mES in the general form

$$\rho = g(\mu, \sigma^2, \mu_3, \mu_4) \tag{3.10}$$

and so the mVaR and ES estimators have the general form

$$\hat{\rho} = g(\hat{\mu}, \hat{\sigma}^2, \hat{\mu}_3, \hat{\mu}_4) \tag{3.11}$$

where g is determined by the value $\mathbf{x}$ according to (3.2) and (3.4) for the case of mVaR and mES respectively. Let $\boldsymbol{\theta} = \begin{pmatrix} \mu & \sigma^2 & \mu_3 & \mu_4 \end{pmatrix}'$ and correspondingly $\rho = g(\boldsymbol{\theta})$. Then according to the delta rule with $(\hat{\mu}, \hat{\sigma}^2, \hat{\mu}_3, \hat{\mu}_4)$ a consistent and asymptotically normal estimator having asymptotic covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, $\hat{\rho}$ is a consistent and asymptotically normal estimator of ρ with asymptotic variance

$$V_{\hat{\rho}_\infty} = \nabla^\top_{g(\boldsymbol{\theta})} \boldsymbol{\Sigma} \nabla_{g(\boldsymbol{\theta})} \tag{3.12}$$

where

$$\begin{aligned} \nabla_{g(\boldsymbol{\theta})} &= \nabla_{\mathbf{x}(\boldsymbol{\theta})^\top \mathbf{z}(\boldsymbol{\theta})} && (3.13) \\ &= \frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\boldsymbol{\theta}} \mathbf{x} + \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\boldsymbol{\theta}} \mathbf{z} && (3.14) \end{aligned}$$

Straightforward calculation shows that

$$\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\boldsymbol{\theta}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2\sigma} & -\frac{\mu_3}{\sigma^4} & \frac{-5\mu_3^2}{2\sigma^7} & \frac{-3\mu_4}{2\sigma^5} - \frac{3}{2\sigma} \\ 0 & 0 & \frac{1}{\sigma^2} & \frac{2\mu_3}{\sigma^5} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sigma^3} \end{bmatrix} \tag{3.15}$$

For mES

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\boldsymbol{\theta}} = \frac{\mathrm{d}g_\gamma}{\mathrm{d}\boldsymbol{\theta}}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}g_\gamma} \tag{3.16}$$

where

$$\frac{\mathrm{d}g_\gamma}{\mathrm{d}\boldsymbol{\theta}} = \begin{bmatrix} 0, & \frac{\mu_3^2 z\left(2z^2-5\right)}{12\sigma^8} - \frac{\mu_3\left(z^2-1\right)}{4\sigma^5} - \frac{\mu_4 z\left(z^2-3\right)}{12\sigma^6}, & \frac{z^2-1}{6\sigma^3} - \frac{\mu_3 z\left(2z^2-5\right)}{18\sigma^6}, & \frac{z\left(z^2-3\right)}{24\sigma^4} \end{bmatrix}^\top$$

and

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}g_\gamma} = \frac{\sigma}{\gamma}\phi\left(g_\gamma\right)\begin{bmatrix} 0, & g_\gamma, & \frac{1}{6}\left(g_\gamma^4 - 3g_\gamma^2\right), & \frac{1}{72}\left(g_\gamma^7 - 15g_\gamma^5 + 45g_\gamma^3 - 15g_\gamma\right), & \frac{1}{24}\left(g_\gamma^5 - 6g_\gamma^3 + 3g_\gamma\right) \end{bmatrix}$$

.However for mVaR there is no dependence on g_γ and consequently $\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\boldsymbol{\theta}} = 0$. Thus mVaR estimator has the following simpler representation expression:

$$V_{\hat{\rho}_\infty} = \mathbf{x}^\top \left(\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\boldsymbol{\theta}}\right)^\top \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \left(\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\boldsymbol{\theta}}\right)\mathbf{x}. \tag{3.17}$$

We now just need the asymptotic covariance matrix $\boldsymbol{\Sigma}$ of the estimators of the first four central moments. It is shown in Appendix A.2 that the lower triangular part of that symmetric matrix is

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \begin{bmatrix} \sigma^2 & & & \\ \mu_3 & \mu_4 - \sigma^4 & & \\ \mu_4 - 3\sigma^4 & \mu_5 - 4\sigma^2\mu_3 & \mu_6 - \mu_3^2 - 6\sigma^2\mu_4 + 9\sigma^6 & \\ \mu_5 - 4\mu_3\sigma^2 & \mu_6 - \sigma^2\mu_4 - 4\mu_3^2 & \mu_7 - 5\mu_3\mu_4 - 3\sigma^2\mu_5 + 12\sigma^4\mu_3 & \mu_8 - \mu_4^2 - 8\mu_3\mu_5 + 16\sigma_2\mu_3^2 \end{bmatrix} \tag{3.18}$$

Finally we note that the asymptotic variance expression for mVaR (3.17) may be written in the form

$$V_{\hat{\rho}_\infty} = \mathbf{x}^\top \widetilde{\boldsymbol{\Sigma}}\,\mathbf{x}$$

where

$$\widetilde{\boldsymbol{\Sigma}} = \left(\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\boldsymbol{\theta}}\right)^\top \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \left(\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\boldsymbol{\theta}}\right)$$

is the asymptotic joint covariance matrix of estimators $\hat{\mathbf{z}} = \left(\hat{\mu}, \hat{\sigma},\ \hat{\sigma}\hat{\gamma}_1,\ \hat{\sigma}\hat{\gamma}_1^2,\ \hat{\sigma}\hat{\kappa}\right)^\top$

3.3 Validity of mVaR and mES estimators

First we note that the existence of asymptotic variance formula (3.17) requires the existence of the covariance matrix $\mathbf{\Sigma}$ given by (3.18). The latter depends on the existence of central moments of order eight. The existence of these moments requires that the t-distribution have greater than eight degrees of freedom. Thus for the remainder of this paper we only consider a t-distribution with at least nine degrees of freedom.

Skewness and excess kurtosis can have a large impact on the approximated quantile. In particular the Cornish Fisher expansion generates negative values for probability density function for certain skewness and kurtosis pairs. These pairs generate non-monotone quantile functions where the order of the quantiles of the distribution is not preserved and the risk measures will no longer be coherent. Valid pairs of skewness and kurtosis can be found using the formula

$$\frac{\gamma_1^2}{9} - 4\left(\frac{\kappa}{8} - \frac{\gamma_1^2}{6}\right)\left(1 - \frac{\kappa}{8} - \frac{5\gamma_1^2}{36}\right) \leq 0$$

where $\gamma_1$and κ are skewness and excess kurtosis respectively. Further details can be found in Maillard (2012). For normal distribution $\gamma_1 = 0$ and $\kappa = 0$ and the above relationship is always valid. For t-distributions$\gamma_1 = 0$ and $\kappa = \frac{6}{\nu-4}$, and so the above relationship is valid only for $\nu > \frac{19}{4}$. However, since we need $\nu > 8$ for our results to hold, all values of skewness and kurtosis for standardized-t distribution lie within the domain of validity for Cornish Fisher expansion.

Chapter 4

PROJECTED STANDARD ERRORS OF MODIFIED ESTIMATORS

Consider a modified estimator ρ that depends on the population central moments, $\rho = g(\mu, \sigma^2, \mu_3, \mu_4)$. The estimator is asymptotically normally distributed and its variance, σ^2_ρ, also depends on the population central moments

$$\sqrt{n}\left(\hat{\rho} - \rho\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2_\rho\right)$$

.To calculate the finite sample risk estimator for a sample of size n, we replace the population central moments with sample central moments $\hat{\rho} = g(\hat{\mu}, \hat{\sigma}^2, \hat{\mu}_3, \hat{\mu}_4)$. Similarly, to calumniate the finite sample variance of the estimator from its asymptotic variance formula, σ^2_ρ, we replace population central moments with sample central moments and calculate the projected standard error as, $\frac{\hat{\sigma}_\rho}{\sqrt{n}}$. In our Monte-Carlo study, we generate M replicates of size n from a standardized distribution. For each replicate, i, of size n we calculate, $\frac{\hat{\sigma}_{\rho,i}}{\sqrt{n}}$ and evaluate the projected standard error for $\hat{\rho}$ as

$$SE\left(\hat{\rho}\right) = \frac{1}{M}\sum_{i=1}^{M}\frac{\hat{\sigma}_{\rho,i}}{\sqrt{n}}$$

4.1 Normal Distribution Modified Estimators Projected Standard Errors

The obvious question is how well the projected S.E.'s represents the true finite-sample S.E's. The results in Figure 4.1 answer this question. For tail probability 2.5% mVaR the projected S.E.'s under-estimate the true S.E.'s by 5% at sample size 250, and by less than 5% for larger sample sizes and larger tail probabilities. The situation is worse for mES at sample size 250 and 2.5% tail probability where finite-sample S.E.'s are almost a 10% under-estimate, and are not much better at 5% tail probability. However, a reasonable way to improve finite-sample S.E.'s and associated confidence intervals for the modified estimators would be to inflate the projected S.E.'s by an appropriate amount revealed in the figures, e.g., by 5% for VaR and 10% for mES in the examples just cited

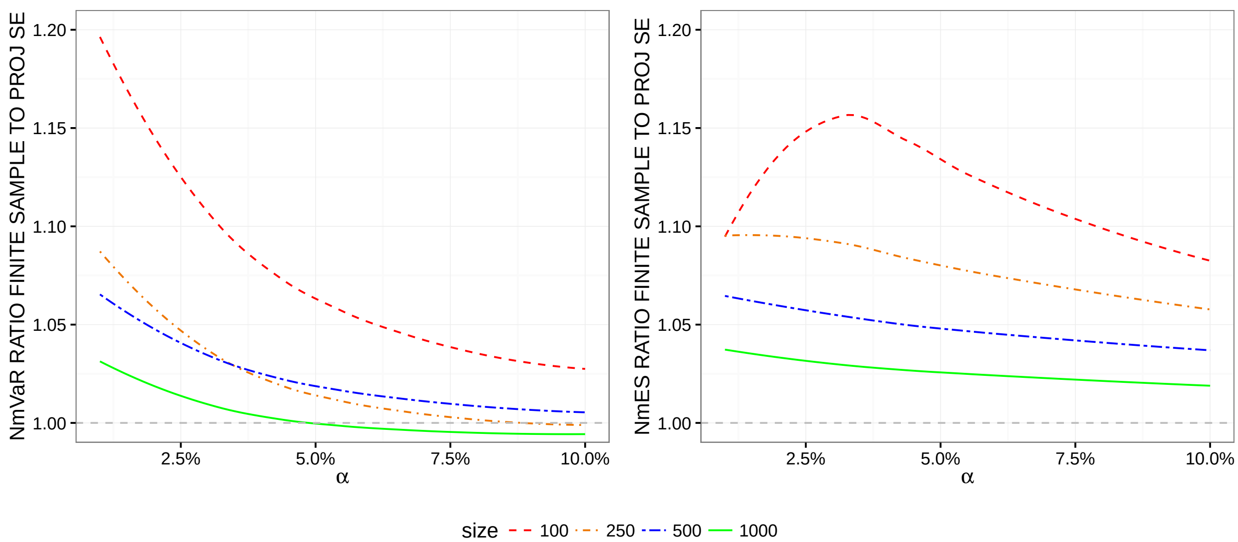

Figure 4.1: Normal Dist mVaR and mES Ratios of Finite-Sample to Projected S.E.'s

4.2 t-Distribution Modified Estimators Projected Standard Errors

Ratios of projected S.E.'s to true finite-sample S.E.'s for the modified estimators for t-distributions with unknown degrees of freedom are shown in Figure 4.2. If one is willing to tolerate an mVaR under-estimate of finite-sample S.E. as large as 20% across all degrees of freedom then one can use a 2.5% tail probability and a sample size of 500 or larger, and sample size 250 achieves almost the same result. But if one is willing to use a 5% mVaR tail probability then a sample of size 250 or larger will insure a S.E. under-estimate of not greater than 10%.

For mES at 2.5% tail probability, the projected S.E. under-estimation is much worse with a 9 degrees of freedom under-estimate of 90% at sample size 100 and about 68% at sample size 250. One possibility would be to adjust the projected S.E. by an inflation factor to match the Monte Carlo estimate of the true finite-sample S.E. The problem in doing so is the wide variability in the estimated finite-sample S.E. across different degrees of freedom.

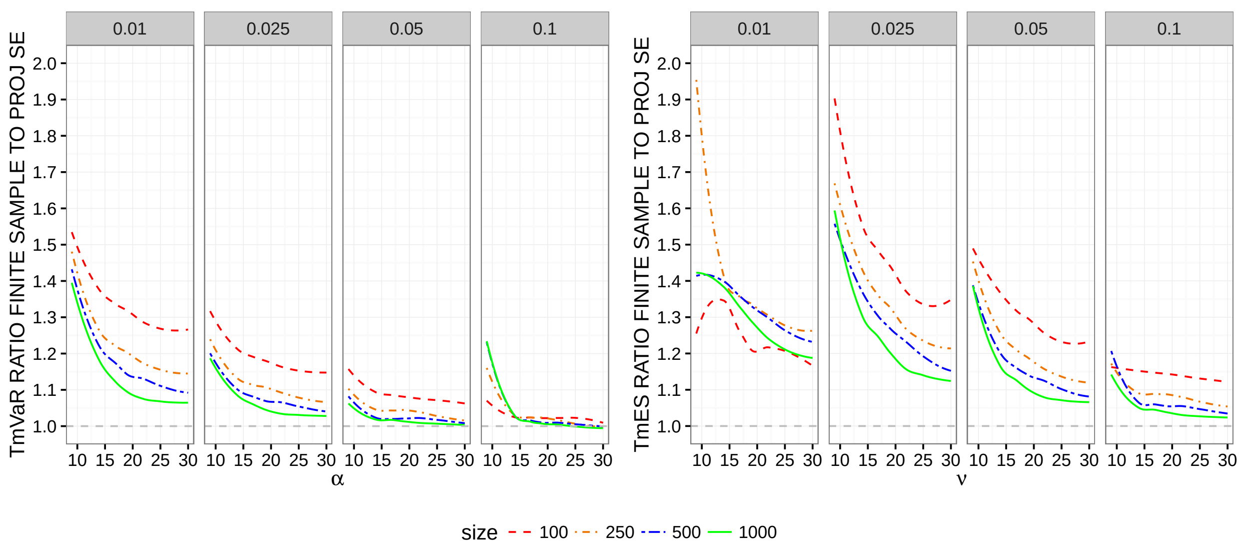

Figure 4.2: t-Dist mVaR and mES Ratios of Finite-Sample to Projected S.E.'s

Chapter 5

CONCLUSION

Our overall goal was to study the asymptotic and finite-sample variance of mVaR and mES estimators with a focus on returns having both normal distributions and t-distributions. Initially we derived general expressions for the asymptotic variance of modified estimators of VaR and ES. A key to the derivation of the large sample variance of the mVaR and mES estimators was use of the asymptotic covariance matrix of the first four sample central moment estimators along with the multivariate delta method. A small but useful step in deriving the large sample variance of the mES estimators was to translate the complex form of the estimator given by Boudt et al. (2008) into a form very similar to the well-known form of mVaR estimators, i.e., with skewness and kurtosis correction terms that are polynomials in certain quantiles.

With respect to projected S.E.'s for mVaR at t-distributions we see that projected S.E.'s under-estimating true finite-sample S.E.'s and that choice of 5% tail probability is better than 2.5%. For mES estimator at t-distributions projected S.E.'s under-estimates true finite-sample S.E.'s and 10% tail probability though not standard for risk measurement is preferred.

One limitation to our study is that we do not consider skewed fat-tailed distributions. An extension of this thesis with study for returns having skewed t-distributions studied by Azzalini and Capitanio (2003) is in progress. The question of performance of risk estimators in general in the presence of serial correlation in returns is open problem for which there currently exists relatively little literature for tail risk estimators. See for example Ledoit and Wolf (2008) for Sharpe ratio hypothesis testing with serial correlation, and Ardia and Boudt (2015) for a similar study of modified Sharpe ratio hypothesis testing with serial correlation.

Appendix A

A.1 Modified Expected Shortfall Preferred Form

$$E_{G_2}\left[z\middle|\, z \leq g_\alpha\right] = -\frac{1}{\alpha}\left\{\begin{array}{l} \phi\left(g_\alpha\right) + \frac{1}{6}\left[\mathrm{I}^3 - 3\mathrm{I}\right]\gamma_1 \\ +\frac{1}{72}\left\{\mathrm{I}^6 - 15\mathrm{I}^4 + 45\mathrm{I}^2 - 15\phi\left(g_\alpha\right)\right\}\gamma_1^2 \\ +\frac{1}{24}\left[\mathrm{I}^4 - 6\mathrm{I}^2 + 3\phi\left(g_\alpha\right)\right]\kappa \end{array}\right\}$$

$$\mathrm{I}^q = \begin{cases} \displaystyle\sum_{i=1}^{q/2}\left(\frac{\prod_{j=1}^{q/2} 2j}{\prod_{j=1}^{i} 2j}\right) g_\alpha^{2i}\phi\left(g_\alpha\right) + \left(\prod_{j=1}^{q/2} 2j\right)\phi\left(g_\alpha\right) & q\,\text{even} \\ \displaystyle\sum_{i=0}^{(q-1)/2}\left(\frac{\prod_{j=0}^{(q-1)/2}(2j+1)}{\prod_{j=0}^{i}(2j+1)}\right) g_\alpha^{2i+1}\phi\left(g_\alpha\right) - \left(\prod_{j=0}^{(q-1)/2}(2j+1)\right)\Phi\left(g_\alpha\right) & q\,\text{odd} \end{cases}$$

$$\mathrm{I}^1 = g_\alpha\phi\left(g_\alpha\right) - \Phi\left(g_\alpha\right)$$

$$\mathrm{I}^2 = \phi\left(g_\alpha\right)\left(g_\alpha^2 + 2\right)$$

$$\mathrm{I}^3 = \phi\left(g_\alpha\right)\left(3g_\alpha^1 + g_\alpha^3\right) - 3\Phi\left(g_\alpha\right)$$

$$\mathrm{I}^4 = \phi\left(g_\alpha\right)\left(4g_\alpha^2 + g_\alpha^4 + 8\right)$$

$$\mathrm{I}^6 = \phi\left(g_\alpha\right)\left(24g_\alpha^2 + 6g_\alpha^4 + g_\alpha^6 + 48\right)$$

$$E_{G_2}\left[z|\, z \le g_\alpha\right] = -\frac{1}{\alpha}\left\{\begin{array}{l} \phi\left(g_\alpha\right) \\ +\frac{1}{6}\left[\phi\left(g_\alpha\right)\left(3g_\alpha^1+g_\alpha^3\right)-3\Phi\left(g_\alpha\right)-3\left(g_\alpha^1\phi\left(g_\alpha\right)-\Phi\left(g_\alpha\right)\right)\right]\gamma_1 \\ +\frac{1}{72}\left\{\begin{array}{c} \phi\left(g_\alpha\right)\left(24g_\alpha^2+6g_\alpha^4+g_\alpha^6+48\right)-15\phi\left(g_\alpha\right)\left(4g_\alpha^2+g_\alpha^4+8\right) \\ +45\phi\left(g_\alpha\right)\left(g_\alpha^2+2\right)-15\phi\left(g_\alpha\right) \end{array}\right\}\gamma_1^2 \\ +\frac{1}{24}\left[\phi\left(g_\alpha\right)\left(4g_\alpha^2+g_\alpha^4+8\right)-6\phi\left(g_\alpha\right)\left(g_\alpha^2+2\right)+3\phi\left(g_\alpha\right)\right]\kappa \end{array}\right\}$$

$$= -\frac{\phi\left(g_\alpha\right)}{\alpha}\left\{1+\frac{\gamma_1}{6}g_\alpha^3+\frac{\gamma_1^2}{72}\left(g_\alpha^6-9g_\alpha^4+9g_\alpha^2+3\right)+\frac{\kappa}{24}\left(g_\alpha^4-2g_\alpha^2-1\right)\right\}$$

A.2 Joint Asymptotic Distribution of First Four Sample Moments

Consider a sequence, $\boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n}$, of k-dimensional random vectors that converge in probability to a mean vector $\boldsymbol{\mu}$. This sequence of random vectors is asymptotically normal if

$$\sqrt{n}\left(\boldsymbol{Y_n}-\boldsymbol{\mu}\right) \xrightarrow{d} \mathcal{N}_k\left(0, \boldsymbol{\Sigma}\right) \tag{A.1}$$

where $\boldsymbol{\Sigma}$ is the asymptotic covariance matrix. Let $g : \mathbb{R}^k \to \mathbb{R}^m$ be a once differentiable function with Jacobian $\boldsymbol{J_g}\left(\boldsymbol{\mu}\right)$ in the neighborhood of $\boldsymbol{\mu}$. We can approximate the variance of any smooth function $g\left(\boldsymbol{Y_n}\right)$ using the multivariate delta rule as follows

$$\sqrt{n}\left(g\left(\boldsymbol{Y_n}\right)-g\left(\boldsymbol{\theta}\right)\right) \xrightarrow{d} \mathcal{N}_m\left(0, \boldsymbol{J}_g\left(\boldsymbol{\theta}\right)\Sigma\boldsymbol{J}_g\left(\boldsymbol{\theta}\right)^\top\right) \tag{A.2}$$

Let $\hat{m}_j = \frac{1}{n}\sum_{i=1}^{n} Y_i^j$ be the j^{th} sample raw moment and $\hat{\mu_j} = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^j$ be the corresponding j^{th} sample central moment. For each $n = 1, 2, ..$let us consider a sequence of the following 4-dimensional random vectors

$$\boldsymbol{Y_n} = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n} Y_i \\ \frac{1}{n}\sum_{i=1}^{n} Y_i^2 \\ \frac{1}{n}\sum_{i=1}^{n} Y_i^3 \\ \frac{1}{n}\sum_{i=1}^{n} Y_i^4 \end{pmatrix} = \begin{pmatrix} \hat{m}_1 \\ \hat{m}_2 \\ \hat{m}_3 \\ \hat{m}_4 \end{pmatrix}$$

We define a transformation, $g : \mathbb{R}^4 \to \mathbb{R}^4$ which expresses the central moments $\hat{\mu_k}$ using raw moments $\{\hat{m}_1, ..., \hat{m}_k\}$

$$g\left(\boldsymbol{Y_n}\right) = \begin{pmatrix} \bar{Y} \\ \hat{\mu}_2 \\ \hat{\mu}_3 \\ \hat{\mu}_4 \end{pmatrix}$$

$$g\begin{pmatrix} \hat{m}_1 \\ \hat{m}_2 \\ \hat{m}_3 \\ \hat{m}_4 \end{pmatrix} = \begin{pmatrix} \hat{m}_1 \\ \hat{m}_2 - \hat{m}_1^2 \\ \hat{m}_3 - 3\hat{m}_1\hat{m}_2 + 2\hat{m}_1^3 \\ \hat{m}_4 - 4\hat{m}_1\hat{m}_3 + 6\hat{m}_1^2\hat{m}_2 - 3\hat{m}_1^4 \end{pmatrix}$$

Let $m_j = E\left[Y_1^j\right]$ be the j^{th} population raw moment and $\mu_j = E\left[(Y_1 - \mu)^j\right]$ be the corresponding j^{th} central moment. Then we can express raw moments in terms of central moments

$$\begin{pmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \end{pmatrix} = \begin{pmatrix} \mu \\ \sigma^2 + \mu^2 \\ \mu_3 + 3\mu\sigma^2 + \mu^3 \\ \mu_4 + 4\mu_3\mu + 6\sigma^2\mu^2 + \mu^4 \end{pmatrix}$$

where under transformation $\boldsymbol{g}$, we have

$$g\begin{pmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \end{pmatrix} = \begin{pmatrix} \mu \\ \sigma^2 \\ \mu_3 \\ \mu_4 \end{pmatrix}$$

Using A.1 we have

$$\sqrt{n}\left(\boldsymbol{Y_n} - \begin{pmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \end{pmatrix}\right) \xrightarrow{d} \mathcal{N}_k\left(0, \boldsymbol{\Sigma}\right)$$

where the covariance of raw moments is

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2 & & & \\ \operatorname{cov}\left(Y_1, Y_1^2\right) & var\left(Y_1^2\right) & & \\ \operatorname{cov}\left(Y_1, Y_1^3\right) & \operatorname{cov}\left(Y_1^3, Y_1^2\right) & var\left(Y_1^3\right) & \\ \operatorname{cov}\left(Y_1, Y_1^4\right) & cov\left(Y_1^4, Y_1^2\right) & \operatorname{cov}\left(Y, Y_1^3\right) & var\left(Y_1^4\right) \end{bmatrix}$$

$$\operatorname{cov}\left(Y_1^j, Y_1^k\right) = E\left[Y_1^{j+k}\right] - E\left[Y_1^j\right] E\left[Y_1^k\right]$$

It follows from A.2 that

$$\sqrt{n}\left(g\left(\boldsymbol{Y_n}\right) - g\begin{pmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \end{pmatrix}\right) = \sqrt{n}\begin{pmatrix} \bar{Y} - \mu \\ \hat{\sigma}^2 - \sigma^2 \\ \hat{\mu}_3 - \mu_3 \\ \hat{\mu}_4 - \mu_4 \end{pmatrix}$$

is asymptotically normally distributed with mean zero and covariance matrix given by

$$\boldsymbol{\Sigma_g} = \boldsymbol{J_g}\begin{pmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \end{pmatrix} \boldsymbol{\Sigma} \boldsymbol{J_g}\begin{pmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \end{pmatrix}^\top$$

where the Jacobian of the function evaluated at $\begin{pmatrix} m_1 & m_2 & m_3 & m_4 \end{pmatrix}^\top$ is given as follows

$$\boldsymbol{J_g}\begin{pmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \end{pmatrix} = \boldsymbol{J_g}\begin{pmatrix} \mu \\ \mu^2 + \sigma^2 \\ \mu_3 + 3\mu\sigma^2 + \mu^3 \\ \mu_4 + 4\mu_3\mu + 6\mu^2\sigma^2 + \mu^4 \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2\mu & 1 & 0 & 0 \\ 3\left(\mu^2 - \sigma^2\right) & -3\mu & 1 & 0 \\ -4\mu_3 - 4\mu^3 & 6\mu^2 & -4\mu & 1 \end{bmatrix}$$

After some algebra and matrix multiplication we can deduce the covariance of first four central moments

$$\boldsymbol{\Sigma_g} = \begin{bmatrix} \sigma^2 & & & \\ \mu_3 & \mu_4 - \sigma^4 & & \\ \mu_4 - 3\sigma^4 & \mu_5 - 4\sigma^2\mu_3 & \mu_6 - \mu_3^2 - 6\sigma^2\mu_4 + 9\sigma^6 & \\ \mu_5 - 4\mu_3\sigma^2 & \mu_6 - \sigma^2\mu_4 - 4\mu_3^2 & \mu_7 - 5\mu_3\mu_4 - 3\sigma^2\mu_5 + 12\sigma^4\mu_3 & \mu_8 - \mu_4^2 - 8\mu_3\mu_5 + 16\sigma^2\mu_3^2 \end{bmatrix}$$

BIBLIOGRAPHY

[1] D. Ardia and K. Boudt. “Testing equality of modified Sharpe ratios”. In: *Finance Research Letters* 13 (2015), pp. 97–104.

[2] A. Azzalini and A. Capitanio. “Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65.2 (May 2003), pp. 367–389. ISSN: 1369-7412. DOI: `10.1111/1467-9868.00391`. URL: `http://doi.wiley.com/10.1111/1467-9868.00391`.

[3] K. Boudt, B. Peterson, and C. Croux. “Estimation and Decomposition of Downside Risk for Portfolios with Non-normal Returns.” In: *Journal of risk* 11 (2008), pp. 79–103. ISSN: 1556-5068.

[4] O. Ledoit and M. Wolf. “Robust performance hypothesis testing with the Sharpe ratio”. In: *Journal of Empirical Finance* 15.5 (2008), pp. 850–859.

[5] D. Maillard. “A Users Guide to the Cornish Fisher Expansion”. In: *SSRN Electronic Journal* (Feb. 2012). ISSN: 1556-5068. URL: `http://papers.ssrn.com/abstract=1997178`.

[6] P. Zangari. “A VaR methodology for portfolios that include options”. In: *RiskMetrics Monitor* 1.First Quarter (1996), pp. 4–12.