

Value at Risk Estimation: A Review and Extension

By

Donald R Chambers, PhD.*
Walter E. Hanson KPMG Chair in Finance
Lafayette College

Qin Lu, PhD., CFA**
Associate Professor
Department of Mathematics
Lafayette College

Jan., 2011

Contact: Qin Lu
Pardee 211
Lafayette College
Easton PA 18042
(610)330-5569
luq@lafayette

Value at Risk Estimation: A Review and Extension

ABSTRACT

This paper: (1) provides a review and analysis of VaR (Value at Risk) estimation methodologies, and (2) extends VaR estimation with two new models. Both of the new models use the Cornish-Fisher-Approximation methodology in the context of the Delta-Gamma normal model to include underlying factors. The first model utilizes the normal mixture distribution for the underlying factors while the second model uses a jump diffusion model.

Value at Risk Estimation: A Review and Extension

1. INTRODUCTION

Value at risk (VaR for short) is widely used in financial risk measurement and management as an estimate of the worst loss that can occur for a given time period and at a given confidence level. VaR can be discussed both in the context of percentage losses (i.e., returns) or dollar losses. Statistically speaking, VaR is the p^{th} quantile of a distribution (where $1-p$ is the desired confidence level) as illustrated in Exhibit 1 for a case of $p=1\%$ and a VaR of \$100,000.

As a highly simplified example, consider a position consisting entirely of \$100,000 of GE common stock. One might wish to know the VaR for a one day period with a confidence level of 95% (a reasonable value might be -3% or \$3,000). This VaR would indicate that there is a 5% chance that on a given day the position will drop in value by 3% or more and a 95% chance that it would do better than losing 3%.

While the definition of VaR is quite unambiguous, the methodologies for estimating VaR are diverse. There are varied approaches to estimating VaR that can generate varied results.

The purpose of this paper is twofold: (1) to provide a review and analysis of VaR estimation methodologies, and (2) to extend VaR estimation with two new models. The two new models use the Cornish-Fisher-Approximation methodology in the context of the Delta-Gamma normal model to include underlying factors with two different models: the normal mixture distribution and the jump diffusion model.

Sections 2, 3 and 4 provide a review and analysis of major VaR estimation methodologies. Sections 5 and 6 introduce two new VaR estimation models, and Section 7 provides a summary and offers conclusions.

2. VAR ESTIMATION METHODOLOGIES OVERVIEW

We begin this overview of VaR estimation methodologies by identifying three major categories of VaR estimation as illustrated in Exhibit 2. The list in Exhibit 2 is not intended to be mutually exclusive or exhaustive: there are methodologies that combine elements from two or more categories and some methodologies do not fit in any of the three categories. Exhibit 2 simply attempts to organize the great majority of the methodologies into a more comprehensible structure.

We provide some detail regarding each of the three categories in each of the three subsections that follow. However, it is the second category, analytical/parametric approaches, that is further explored in later sections.

2.1 Historical/Empirical/Quantile Approaches

The simplest example of the historical approach (i.e., empirical quantile estimation) involves forming a frequency distribution of past outcomes and estimating VaR as the loss level that divides past outcomes into $p\%$ outcomes that are lower and $(1-p)\%$ outcomes that are higher (i.e., the p^{th} quantile).

Returning to the one day VaR for \$100,000 of GE stock (with 95% confidence), one would compute the daily percentage profits and losses on such a position over a previous time interval, rank the outcomes from worst to best, and estimate VaR as the loss that was exceeded in 5% of the outcomes.

This simple approach has two major drawbacks. First, it assumes that the return distribution of the position has been and will continue to be unchanged. Second, if it is based on limited data, the estimation of the quantile is not efficient. See Engle and Mangaelli's survey for other drawbacks ([8]).

There are more sophisticated methodologies within this category including quantile regression and its extensions which include the conditional autoregressive VaR (CAViaR) model. We will briefly review these two methods, both of which attempt to provide improved estimations of future volatility.

Tsay ([17]) provides a review of the empirical quantile estimation and basic quantile regression models introduced by Koenker and Bassett ([15]). Basic quantile regression analyzes returns using a user-defined set of explanatory (observable economic) variables. Koenker and Bassett [15] provide a general formula that is used in the analysis that follows. The p^{th} regression quantile q_t is defined as the solution of the minimization problem:

$$q_t = \min_{\beta \in R^k} \left\{ \sum_{t \in \{t: r_t \geq X_t \beta\}} p |r_t - X_t \beta| + \sum_{t \in \{t: r_t \leq X_t \beta\}} (1-p) |r_t - X_t \beta| \right\},$$

where r_t is the return at time t , X_t is a k dimensional vector of observable economic variables at time t , and β is a k dimensional vector of parameters.

The idea is to include current values of relevant economic variables in making forecasts of future volatility. For example, a fixed income portfolio may have returns that are correlated with interest rate levels. Further, the level of unexplained returns may vary based on the interest rate level. The idea of quantile regression approaches in this example would be to base forecasts of future return volatility on the slope coefficients

estimated from the historical regression, β , and on the current set of explanatory variables, X , that would include interest rate levels.

The CAViaR model (where CAViaR is an acronym Conditional Autoregressive VaR model with the “i” inserted for aesthetics) was introduced by Engle and Mangaelli [8]. The CAViaR model extends a static quantile approach based on the distribution of past returns to an evolutionary or dynamic view of the quantile over time. A general formula given in Engle and Mangaelli’s paper ([9]) is:

$$q_t = \beta_0 + \sum_{i=1}^m \beta_i q_{t-i} + l(\beta_{m+1}, \dots, \beta_{m+n}, \Omega_{t-1}),$$

where q_t is the p^{th} quantile at time t , $\beta_i \{0 \leq i \leq m+n\}$ contains the $m+n+1$ parameters, l are some specifically chosen functions, and Ω_{t-1} is the information available at time $t-1$ (which could be r_{t-1} , the return at time $t-1$ or other explanatory variables such as observable economic variables at time $t-1$). CAViaR models have many specializations as well as some generalizations. For example, a specialization is that $m=1, n=1$, and $l = \beta_2 |r_{t-1}|$ wherein the current VaR is based on past VaRs and the most recent return. An example of a generalization is the use of non-linear regression techniques.

In summary, historical or quantile approaches (illustrated in Exhibit 2 with the left most box) vary from rather simple approaches that assume static return distributions to more sophisticated approaches that attempt to model dynamic behavior. Pure historical or quantile estimation methods are *nonparametric* approaches since they do not model the return distribution using a specific return distribution with a specific set of parameters (e.g., mean and variance). The following subsection briefly reviews parametric approaches which are also illustrated in Exhibit 2 as one of the three major types of VaR estimation.

2.2 Analytic Estimation (Parametric and Semi-parametric models)

This sub-section introduces the analytic or parametric VaR estimation approaches which are probably the most common approaches in academia and industry. These approaches attempt to estimate VaR using the probability distribution of returns as expressed through the parameters of that distribution. This category is introduced in this sub-section and is further detailed as the focus of Section 3 and much of the remainder of this paper.

In the simplest case, the underlying return distribution is assumed to follow a normal distribution and therefore VaR can be estimated using the parameters of that distribution:

$$VaR = \hat{\mu} + z_c \hat{\sigma},$$

where $\hat{\mu}$ is the mean return, $\hat{\sigma}$ is the volatility (standard deviation) and z_c is the critical number for the p^{th} quantile of the standard normal distribution. The value of z_c is found in tables of the normal distribution using the given confidence interval. For example, if $p=5\%$ (indicating a 95% confidence level), then $z_c = -1.645$. The idea is that outcomes that lie more than 1.645 standard deviations below the mean are only expected to occur 5% of the time.

Returning to the GE stock example and continuing to use a 95% confidence level so that $z_c = -1.645$, if the GE stock returns are assumed to have a daily standard deviation of 1.5% (and ignoring the small daily mean return), the VaR of the \$100,000 position would be -2.4675% or -\$2,467.50. The primary problem with this approach is that returns are not normally distributed. In particular, returns tend to have larger probabilities of extreme outcomes than is found in normally distributed variables (i.e., return distributions are typically fat tailed or leptokurtic). Since VaR focuses on the tail of a distribution, most extensions to the above model address the issue of leptokurtosis.

For the purposes of our paper, the second category from Exhibit 2 (analytic or parametric) also includes semi-parametric models (models that use parameters to describe part of the return distribution rather than the entire distribution) and models that have both analytical or semi-analytical solutions. This category includes three major methods: variance-covariance models and their extensions (the delta-gamma normal models); GARCH related models; and extreme value theorem models. These extensions will be detailed in Section 3.

2.3 Simulation Approaches

Another popular major category of VaR estimation approaches (as illustrated in Exhibit 2) utilizes simulations such as the Monte-Carlo simulation method. The idea is to develop models of returns and then utilize computers to project a multitude of possible future paths. VaR is typically estimated as a quantile of the projected outcomes.

The projections can be based directly on past outcomes or on sets of projected outcomes or scenarios. To the extent that past observations are directly used as projections of the future, the approach tends towards being an historical or empirical approach as reviewed in sub-section 2.1.

The projections or scenarios can be based on current market data (e.g., implied volatilities from options) or based on professional judgment. The projections may be enumerated or drawn from distributions based on given parameters. To the extent that the simulation is based on parameters, the method tends towards generating similar results to the parametric approaches discussed in sub-section 2.2. Simulations often utilize combinations of parametric and non-parametric techniques and can be based on historical and/or expectational data.

3. ANALYTICAL/PARAMETRIC MODELS

Analytic models were introduced in Section 2 as the most popular of the three major VaR estimation approaches listed in Exhibit 2. In the simplest parametric approach, returns of the asset for which VaR is being estimated are assumed to form a normal distribution with known parameters and the solution is therefore straightforward:

$$VaR = \hat{\mu} + z_c \hat{\sigma} .$$

This section provides detail with regard to relatively simple extensions of the most basic analytic/parametric approach. Exhibit 3 divides these extensions into three major sub-categories: GARCH related models, multiple factor linear normal models, and extreme value theorem models. Each of the three sub-sections that follow discuss the three sub-categories in Exhibit 3. The second sub-section (regarding multiple factors with normal distributions) serves as an introduction to several extended approaches that are detailed in Section 4 and that serve as foundations to the new models introduced in Sections 5 & 6.

3.1 GARCH related VAR

GARCH related VAR is a parametric approach that is simply an extension of the previously discussed normal model with greater attention paid to the estimation of the volatility of the underlying asset. Recall that the normal model expresses VaR as:

$$VaR = \hat{\mu} + z_c \hat{\sigma} ,$$

To allow for a time horizon denominated in days (or other small units) and to measure VaR in dollar units (or any other currency), the above formula can also be written as

$$VaR = W(k\hat{\mu} + z_c \sqrt{k}\hat{\sigma}),$$

where W is the value of the portfolio for which VaR is being estimated, $\hat{\mu}$ is the forecast mean daily return of W , k is the number of days for which the VaR is being estimated, z_c is the critical number for the p^{th} quantile of the standard normal distribution and $\hat{\sigma}$ is the daily volatility of the return of W .

The volatility in the above formula can be estimated using a variety of techniques including historical data and implied volatilities from the option market. However, a GARCH type model generates the estimate of the volatility from a time variant model.

There are many studies that utilize a GARCH approach to estimate VaR. Most GARCH related VaR models assume that the return distribution r_t , conditioned on information of time $t-1$, is normally distributed. Assuming normality, the above VaR formula can be used. This sub-section follows approaches in the books by Dowd [4] and Tsay[17].

RiskMetrics uses an exponentially weighted moving average model (EWMA) to forecast the volatility $\hat{\sigma}_t$ (which is equivalent to an IGARCH model with zero mean). The approach involves two steps. The first step in the EWMA model approach is to forecast variance by modeling the volatility cluster.

$$\begin{aligned} r_t &= a_t, \\ a_t &= \sigma_t \varepsilon_t, \\ \sigma_t^2 &= \lambda \sigma_{t-1}^2 + (1 - \lambda) r_{t-1}^2, \end{aligned}$$

where λ is usually set to 0.94 for daily data and 0.97 for monthly data, and ε_t is a standard Gaussian white noise series (independent standard normal distribution). The forecast variance $\hat{\sigma}_t^2$ relies on the previous day variance σ_{t-1}^2 and the square of the previous day return, r_{t-1}^2 .

In the second step of the EWMA model approach, the VaR is calculated (with a time horizon adjustment). Since the mean is zero, it can be dropped. To match the volatility time horizon and VaR time horizon, RiskMetrics uses the *square root of time rule*:

$$VaR_{EWMA} = z_c \sqrt{k} \hat{\sigma}_t,$$

where z_c is the critical number for p^{th} quantile of standard normal distribution, k is the time horizon (typically in days) for VaR and $\hat{\sigma}_t$ is the EWMA model's forecasted volatility for the daily return. In theory, any GARCH related model can be used to forecast the volatility $\hat{\sigma}_t$. So there are a lot of possible models. For example, EGARCH can be used to model asymmetrical effects. The simplest case is the GARCH (1, 1) case which also consists of two steps.

In the first step, variance is forecast using a GARCH (1, 1) model:

$$\begin{aligned} r_t &= \mu + a_t \\ a_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \beta_1 a_{t-1}^2 \end{aligned}$$

In the second step of the GARCH (1, 1) model approach, the VaR is calculated (with a time horizon adjustment).

$$VaR_{GARCH} = k\hat{\mu} + z_c \sqrt{\sum_{l=1}^k \sigma_h^2(l)}$$

where $\sigma_h^2(l)$ is the GARCH l^{th} step forecast volatility starting at origin h , and other variables as previously defined. The calculation of l^{th} step forecast volatility $\sigma_h^2(l)$ is by the following:

$$\begin{aligned}\sigma_h^2(1) &= \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2 \\ \sigma_h^2(l) &= \alpha_0 + (\alpha_1 + \beta_1) \sigma_h^2(l-1), \quad l = 2, \dots, k.\end{aligned}$$

RiskMetrics and GARCH (1, 1) tend to generate similar short term variance forecasts, but do not have similar long term forecasts. The parameters for this category can be estimated by maximal likelihood or quasi-maximum likelihood methods. The advantage of GARCH models is that they consider the time variation in risk. However, GARCH models in general need a large number of observations to get reliable results and a misspecification of the GARCH variance model will cause an inaccuracy of a VaR calculation. Also, the standard residual, ε_t , may not be i.i.d. (independently and identically distributed) and normal, so variance can not be used to calculate the p^{th} quantile. Monte-Carlo simulation can be used to deal with non Gaussian standard errors. A variety of GARCH models based on non Gaussian standard errors such as the t-distribution have been developed.

3.2. Multiple linear normally distributed factors (i.e., the variance-covariance approach)

Sub-section 2.2 described the simplest parametric approach used in the case of a return that formed a normal distribution.

$$VaR = \hat{\mu} + z_c \hat{\sigma}$$

This “normal model” can be generalized by modeling the return of the underlying asset as a linear combination of normally distributed factors. The resulting model, the delta normal model, assumes that the return of the position for which VaR is being computed is a linear combination of returns on m underlying assets or factors. The model is simply a multivariate extension of the normal model (reviewed in sub-section 2.2) as detailed below.

Assume that the asset for which the VaR is being computed is a linear combination of m underlying assets or factors

$$R_{p,t} = w^T R_t = (w_1, \dots, w_m) \begin{pmatrix} R_{1t} \\ \dots \\ R_{mt} \end{pmatrix},$$

where w is the m dimensional vector of weights of underlying assets in the portfolio. Formally, assume the m dimensional vector of individual asset log returns R_t follows multivariate Brownian motion process with some $m \times m$ covariance matrix.

$$R_t = \alpha t + \sqrt{t} \Sigma^{1/2} Z_0$$

where Z_0 is standard normal in R^m , so R_t is a multivariate normal distribution with mean αt and covariance matrix $t\Sigma$. (Hence, Σ is a symmetric positive semi-definite matrix of dimension $m \times m$).

It is easy to see that this linear combination of normal distributions is also a normal distribution with mean given by $\mu_p = w^T \alpha t$, and variance given by $\sigma_p^2 = w^T (\Sigma t) w$. Hence, following the normal model, the formula for VaR is:

$$VaR_{\text{delta-normal}} = \mu_p + z_c \sigma_p = w^T \alpha t + z_c \sqrt{w^T \Sigma t w}$$

In dollars units, the formula is:

$$VaR_{\text{delta-normal}} = W(\mu_p + z_c \sigma_p) = W(w^T \alpha t + z_c \sqrt{w^T \Sigma t w}),$$

where W is the value of the portfolio.

Both the normal model and the delta-normal model are called variance-covariance approaches. This is because the VaR (p^{th} quantile) depends only on variance and covariance for the normal distribution (and does not require higher moments).

3.3 Extreme value theorem (EVT) models:

The third VaR approach illustrated in Exhibit 3 is extreme value theorem (EVT) models. These models are semi-parametric approaches because they use parameters to describe only the tails of the probability distribution. The focus of EVT models on the distributions of tails makes their use appropriate for VaR estimation since in estimating VaR only the distribution of the tail matters. This review of EVT models is based on three studies ([1], [16], and [17]).

A starting point for understanding the EVT method involves limit distributions in general and the central limit theorem in particular. The limit distribution is the distribution for a sequence of random variables, X_n , as n goes to infinity. The central limit theorem provides a good example of applying limit distributions in the case of the distribution of sample means. The central limit theorem derives the limit distribution of the sample mean statistic, $\frac{r_1 + \dots + r_n}{n}$, without knowing the distribution of r_t (it requires some

conditions such as the r_t having independent identical distributions). The crucial part for the central limit theorem is that the limit distribution of the sample mean statistics does not depend on the distribution of r_t . Thus, researchers can use the central limit theorem to understand the distribution of the sample mean without knowing the exact distribution of the variables underlying the means. Similarly, EVT models can be used to estimate VaR without knowledge of or specification of the return distribution of the position for which the VaR is being estimated.

There are two primary approaches of EVT: the block maxima method and the peak-over-threshold method. The former focuses on extremes (maximum or minimum), while the latter focuses on exceedances of a threshold and the frequency with which exceedances occur. The block maxima method uses the limit distribution approach to estimate VaR based on minimum returns over sub-periods, while the peak-over-threshold estimates VaR based on frequencies with which minimum values have been exceeded.

3.3.1 The block maxima method

The EVT allows researchers to understand the distribution of VaR while knowing only the distribution of the tails of the underlying returns (i.e., without knowing the full distribution of the underlying returns).

Consider a normalized minimum statistic:

$$\frac{r_{(1)} - \beta_n}{\alpha_n},$$

where $r_{(1)} = \min\{r_t : 1 \leq t \leq n\}$ and α_n, β_n are scale parameters and location parameters which depend on the tail distribution of r_t . (Some conditions are necessary such as r_t having an independent identical distribution, see [17] and its references).

The crucial result from the EVT is that the limit distribution of a normalized minimum statistic only depends on the tail distribution of r_t . While the central limit theorem establishes that the limit distribution is a normal distribution, the EVT establishes that the limiting distribution of the normalized minima, $\frac{r_{(1)} - \beta_n}{\alpha_n}$, belongs to one of three categories depending on a shape parameter, k , which further depends on the tail behavior of r_t .

- (1) if $k=0$, the limiting distribution is the Gumbel family,
- (2) if $k<0$, the limiting distribution is the Frechet family, and
- (3) if $k>0$, the limiting distribution is the Weibull family.

The left tail of r_t declines exponentially for the Gumbel family, by a power function for the Frechet family and finitely for the Weibull family.

In order to estimate VaR using the block maxima method of the EVT, we need to divide r_t into sub-periods of size n and the shape, location and scale parameters ($k_n, \alpha_n,$ and β_n) must be estimated¹. After estimating the parameters, the Gumbel, Frechet or Weibull distribution CDFs (cumulative density functions) may be used to find the quantile for the normalized sub-period minima. It is then necessary to connect the normalized sub-period minima to the observed return's VaR. We can find the VaR, namely the p^{th} quantile for r_t by the p^{th} quantile for normalized sub-period minima. The VaR formula² is:

$$VaR_{BM} = \begin{cases} \beta_n - \frac{\alpha_n}{k_n} \{1 - [-n \ln(1-p)]^{k_n}\} & k_n \neq 0 \\ \beta_n + \alpha_n \ln[-n \ln(1-p)] & k_n = 0 \end{cases}$$

3.3.1 The Peak-over-threshold method

The second major method of estimating VaR using the EVT, the Peak-over-threshold method, first defines a prespecified high threshold u . It considers a conditional excess distribution:

$$P_{r|u}(r_t \leq x+u | r_t > u) = \frac{F_{r_t}(x+u) - F_{r_t}(u)}{1 - F_{r_t}(u)}.$$

This conditional excess distribution has a limit distribution. If the threshold u is high, the limit distribution is close to a generalized Pareto distribution as follows:

$$G_{k,\alpha,\beta}(x) = \begin{cases} 1 - [1 + k(\frac{x-\beta}{\alpha})]^{-1/k} & k \neq 0 \\ 1 - \exp(-\frac{x-\beta}{\alpha}) & k = 0 \end{cases}$$

¹ Tsay ([29]) mentioned several methods to get parameters such as maximum likelihood and regression methods for all parameters, and the Pickands estimator and the Hill estimator for the shape parameter k .

² For a time series r_t of size T , we know that $T = ng$, where g is sub-period numbers and n is the size of sub-period. On one hand, we need n to be large to have the limit distribution, on the other hand, we need g to be large to get better estimations for parameters k_n, α_n and β_n and there is a trade off.

For $k \neq 0$, we have $\frac{F_{r_i}(x+u) - F_{r_i}(u)}{1 - F_{r_i}(u)} \approx 1 - [1 + k(\frac{x-\beta}{\alpha})]^{-1/k}$

So, $F_{r_i}(x+u) \approx (1 - F_{r_i}(u)) (1 - [1 + k(\frac{x-\beta}{\alpha})]^{-1/k}) + F_{r_i}(u)$

Note that $F_{r_i}(u) = P(r_i \leq u)$ is the probability such that the return is less than the threshold, which can be modeled by a Poisson process. We denote n_u as the number of exceedences over the threshold. We have

$$F_{r_i}(x+u) \approx \frac{n_u}{n} (1 - [1 + k(\frac{x-\beta}{\alpha})]^{-1/k}) + 1 - \frac{n_u}{n} = 1 - \frac{n_u}{n} [1 + k(\frac{x-\beta}{\alpha})]^{-1/k} + 1$$

This gives the approximate tail distribution for the return r_i (note that u is high so that $x+u$ is in the tail).

From the tail distribution, it is easy to get the VaR:

$$VaR_{GPD} = \hat{\beta} + \frac{\hat{\alpha}}{\hat{k}} \left[\left(\frac{n}{n_u} (1-p) \right)^{-\hat{k}} - 1 \right],$$

where n is sample size, n_u is the number of exceedences over the threshold, p corresponds to the p^{th} quantile. The parameters $\hat{\alpha}, \hat{\beta}, \hat{k}$ are scale, location and shape parameters respectively which can be estimated by a maximum likelihood or regression method.

If we do not use a generalized Pareto distribution as the limit distribution, instead, we use other generalized extreme value distribution, we have³

$$VaR_{GEV} = \hat{\beta} + \frac{\hat{\alpha}}{\hat{k}} \left(\left[-\ln \left(1 - \frac{(1-p)n}{n_u} \right) \right]^{-\hat{k}} - 1 \right).$$

³ VaR calculation using EVT becomes more adequate when we move father into the tail and it allows the asymmetries in the positive and negative tails. However, the block maxima method needs the assumption of independent returns and it overlooks the volatility clustering in the return. Also, we can not make sample sizes n and sub-periods g both large for fixed T since $ng=T$. For the peak-over-threshold method, different choices of the threshold lead to different estimates of the shape parameter. Also, we need to check the model fitting (see [29] for details).

SECTION 4. NON-LINEAR AND ADVANCED ANALYTICAL EXTENSIONS

Section 3 reviewed several analytical/parametric VaR estimation models as illustrated in Exhibit 3. However, perhaps the most important category for VaR estimation using a parametric or analytic approach involves attempts to model the non-normality of returns that is observed especially in, for example, derivatives and hedge funds. This section provides a moderately detailed review and analysis of four important approaches for modeling these complexities.

Exhibit 4 illustrates the four methods. Each of these methods is detailed in the four sub-sections that follow. The first sub-section reviews a method for handling non-normality. Sub-section 4.2 reviews a method for handling a non-linear relationship between a price and its underlying factor(s). The last two sub-sections discuss non-normal underlying factors. These four sub-sections provide foundations for two new models introduced in Sections 5 and 6.

4.1 Cornish-Fisher approximation: Computing VaR for Non-Normal Distributions

When returns are normally distributed, VaR estimation is simplified as demonstrated in sub-sections 2.2 and 3.2 since the critical value, z_c can be found in tables constructed for the standard normal distribution. When returns are somewhat near being normally distributed, but not extremely near being normally distributed, the Cornish-Fisher approximation can be used. The Cornish-Fisher approximation is used by determining a critical value, z_{CF} , to be used in place of z_c in the formula for VaR. The approximation utilizes higher moments of the distribution to adjust for the distribution's non-normality in finding quantiles.

The Cornish-Fisher approximation states that the p^{th} quantile of a distribution can be approximated by using mean, variance, skewness, excess kurtosis and higher cumulants. Farve ([10]) advocates using the first four cumulants (mean, variance, skewness, and excess kurtosis) to estimate the VaR which is called modified VaR. The following equation utilizes four cumulants since most studies indicate that four is enough.

$$VaR_{CF} = \mu + z_{CF}\sigma$$

where

$$z_{CF} = z_c + \frac{1}{6}(z_c^2 - 1)S + \frac{1}{24}(z_c^3 - 3z_c)K - \frac{1}{36}(2z_c^3 - 5z_c)S^2,$$

and $\mu, \sigma, S,$ and K are the mean, standard deviation, skewness and excess kurtosis respectively, and z_c is the critical number for p^{th} quantile of the standard normal distribution.

For any unknown return distribution, if the sample skewness, S , and sample excess kurtosis, K , can be estimated, then they may be used in the above formula to estimate VaR.

The Cornish-Fisher approach has several advantages and disadvantages. The Cornish-Fisher approach is generally easy to implement. Jaschke [14] compares the Cornish-Fisher approach for a delta-gamma normal model with the numerical Fourier inversion approach (and when and when not to use Cornish-Fisher in favor of a saddle points method or a partial Monte-Carlo method). Jaschke notes that the numerical Fourier inversion approach uses the characteristic function which includes all moments (cumulants) while the Cornish-Fisher approach only uses the first four cumulants. So, the numerical Fourier inversion approach may generally be more accurate than the Cornish-Fisher approach. However, since the Cornish-Fisher approach can be computed faster and more simply, Jaschke concludes that in many practical situations, the Cornish-Fisher approximation's actual accuracy is "more than sufficient" and Holton [12] notes that even though the Cornish Fisher approach has limitations,⁴ because of its tractability and sufficiency in practical use, it is advocated.

4.2 Delta-Gamma Approach: Handling a Non-linear Relationship Between Factors and Prices

Many portfolios contain non-linear contracts such as options whose returns are non-linear functions of the returns of the underlying assets. Since a non-linear function of a normally distributed variable is not generally normally distributed, the normal model can not be used directly to estimate VaR (the return of the portfolio generally needs to be a

⁴ To summarize the limitations of Cornish-Fisher approximation as in Jaschke's paper ([14]):

1. The cumulative distribution function approximate by Cornish-Fisher approximation is not necessarily monotonic. Occasionally, a 95% VaR estimate may be higher in absolute value than a 99% VaR estimate. Recently, Chernozhukov et al ([3]) increase rearrangement to monotonize the Cornish-Fisher expansion,
2. It is not very reliable as the quantile approach to 0% or 100%. EVT offers better result for far tails,
3. It is not reliable when the distribution is substantially different than normality,
4. If the fourth moment does not exist, the numerical Fourier inversion approach must be used, and
5. It does not necessarily increase the convergence by involving high cumulants.

linear combination of the return of underlying assets that are normally distributed to use the normal or delta normal model).

The most popular method for addressing the problem of VaR estimation for options and other non-linear positions is the delta-gamma approach. In the delta gamma approach, a quadratic approximation relating portfolio value to risk factors (e.g., underlying assets) is used. For simplicity in illustrating this approach, we assume that the underlying factors returns are jointly normally distributed. In sections 5 and 6 the delta-gamma method is used with other underlying distributions.

In the case of underlying factors with normal distributions, the delta-gamma VaR estimation model is called the delta-gamma normal model. To illustrate, we assume the vector of individual asset log returns R_t follows multivariate Brownian motion process:

$$R_t = \alpha t + \sqrt{t}\Sigma^{1/2}Z_0, \quad (4.2.1)$$

where Z_0 is a vector that is standard normally distributed in R^m . So, R_t is has a multivariate normal distribution with mean αt and covariance matrix $t\Sigma$.

Now we consider a portfolio W_t which is a (potentially) non-linear function of R_t , which can be well approximated as

$$W_t = \Delta^T R_t + \frac{1}{2} R_t^T \Gamma R_t, \quad (4.2.2)$$

where Δ and R_t are $m \times 1$ vectors, and Γ is a symmetric $m \times m$ matrix. Note that W_t is in dollars units. Δ is not the portfolio's delta with respect to an individual asset, but is closely related to that delta. Γ is related to the portfolio's gamma with respect to individual asset. (Details can be found in Hull [13].)

In this model, the underlying assets (risk factors) follow a multi-normal distribution and the portfolio has a quadratic relation with its underlying assets. The difficulty of calculating VaR is that, unlike in the delta normal case, W_t does not form a normal distribution, but in fact, forms the non-central χ^2 distribution.

There are several approaches to estimating the VaR of W_t . Britton [2] offers the Solomon-Stephens approximation approach. Britton's approach uses a distribution matching the first three moments of the density function of the sum of the non-central chi-squares. Jaschke [14] details and reviews other approaches. This sub-section overviews the characteristic function approach and details the Cornish-Fisher approximation approach, both of which are reviewed in Jaschke [14] and Holton [12].

The characteristic function approach for the delta-gamma normal model is begun by calculating the characteristic function for the non central χ^2 distribution. Then through

the inversion theorem, the cumulative distribution function for W_t is derived. From the cumulative distribution function for W_t , the p^{th} quantile of W , namely the VaR, can be found. The problem with this approach is that the inversion theorem involves definite integrals which can only be calculated by numerical approximation. This approach is also called numerical Fourier inversion (see Jaschke [14] and Holton [12] for details).

The second approach to estimating VaR for the delta-gamma normal model uses the Cornish-Fisher approximation method detailed in sub-section 4.1. The Cornish-Fisher approximation method for the estimation of VaR requires computation of the mean, standard deviation, skewness and excess kurtosis of the portfolio. Since W_t is a non-linear function of the normally distributed factors (R_t), computation of these cumulants is challenging.

The first step in calculating the cumulants begins with the quadratic approximation for W and the assumed return distribution for the factors (R_t) as in (4.2.1).

$$\begin{aligned} W_t &= \Delta^T R_t + \frac{1}{2} R_t^T \Gamma R_t \\ &= (\Delta^T \alpha t + \frac{1}{2} (\alpha t)^T \Gamma \alpha t) + ((\alpha t)^T \Gamma + \Delta^T)(R_t - \alpha t) + \frac{1}{2} (R_t - \alpha t)^T \Gamma (R_t - \alpha t) \end{aligned}$$

This shifts the means: from R_t mean at αt to $R_t - \alpha t$ mean at 0. Now we define

$A = \Delta^T \alpha t + \frac{1}{2} \alpha t^T \Gamma \alpha t$, $B = \Gamma \alpha t + \Delta$, $X = R - \alpha t$ and substituting into the above equation yields:

$$W = A + B^T X + \frac{1}{2} X^T \Gamma X$$

Appendix A demonstrates, among other things, the derivation of the variance, skewness and excess kurtosis for the above equation. Using the results in the appendix A and with Σt replacing Σ , the cumulants can be written as

$$\text{The variance } \sigma_{W_t}^2 = \kappa_2 = \frac{1}{2} \text{trace}((\Gamma \Sigma t)^2) + (\Gamma \alpha t + \Delta)^T \Sigma t (\Gamma \alpha t + \Delta) \quad (4.2.3)$$

$$\text{The skewness } S(W_t) = \frac{\kappa_3}{\kappa_2^{3/2}} = \frac{\text{trace}((\Gamma \Sigma t)^3) + 3(\Gamma \alpha t + \Delta)^T \Sigma (\Gamma \Sigma t)(\Gamma \alpha t + \Delta)}{\left\{ \frac{1}{2} \text{trace}((\Gamma \Sigma t)^2) + (\Gamma \alpha t + \Delta)^T \Sigma t (\Gamma \alpha t + \Delta) \right\}^{3/2}} \quad (4.2.4)$$

The excess kurtosis

$$K(W_t) = \frac{\kappa_4}{\kappa_2^2} = \frac{3\text{trace}((\Gamma \Sigma t)^4) + 12(\Gamma \alpha t + \Delta)^T \Sigma t (\Gamma \Sigma t)^2 (\Gamma \alpha t + \Delta)}{\left\{ \frac{1}{2} \text{trace}((\Gamma \Sigma t)^2) + (\Gamma \alpha t + \Delta)^T \Sigma t (\Gamma \alpha t + \Delta) \right\}^2} \quad (4.2.5)$$

The VaR is estimated by inserting the above variance, skewness and excess kurtosis in the Cornish-Fisher approximation function (detailed in sub-section 4.1). The result is the estimation of VaR for a non-linear function of normally distributed factors using the delta-gamma normal and Cornish-Fisher approaches.

The delta-gamma approach detailed above used the Cornish-Fisher approach and used an underlying normal distribution. There are other major approaches to implementing a delta-gamma approach to estimating VaR. Duffie and Pan's survey [5] uses Monte-Carlo simulation with a delta-gamma context (and with normally distributed underlying factors, and jump diffusion underlying factors). They compared the delta approximation, delta-gamma approximation and actual derivative pricing to calculate VaR. The delta approximation VaR was far away from that using the actual derivative pricing VaR. It over-estimates the loss on a long position and under-estimates the loss on a short position. The delta-gamma approximation under-estimates the loss for both long positions and short positions, but by much less. They also found that using a jump-diffusion underlier produces a large VaR than using a normally distributed underlier.

Glasserman and Heidelberger [11] discuss the delta-gamma model where the underliers are multivariate t distributions and some other extensions. They have two models. In the first model they derive the characteristic function for a delta-gamma t model and use a numerical Fourier inversion to find VaR. In the second model, they use the delta-gamma approximation to guide the selection of an effective importance sampling distribution which yields a low variance Monte Carlo method. Another extension related to the delta-gamma models is EI-Jahel's paper [7] where the underlying factors are assumed to have stochastic volatility. Finally, Duffie and Pan [6] analytically extended the delta-gamma model to having the underlying factors follow a jump diffusion process. They use the numerical Fourier inversion approach (i.e., the characteristic function approach) rather than the Cornish-Fisher expansion.

In Sections 5 and 6, we use the delta-gamma approach (along with the Cornish-Fisher approximation) to estimate VaR for two underlying non-normal distributions: the normal mixture model and the jump diffusion model. These two non-normal models are detailed in the following two sub-sections, which can also be found in Duffie and Pan's [5].

4.3 Normal mixture model

While the joint normally distributed model detailed in sub-section 3.2 and utilized in sub-section 4.2 is relatively simple, it has been found to provide poor modeling of the tails of the actual distributions of financial data of underlying assets/factors (where VaR is focused). The normal mixture model is the first of the two models that we use that have been found to offer improved modeling of financial data. The second approach (a jump diffusion process) is detailed in the next sub-section).

We assume the portfolio log returns R_t follows a normal mixture process wherein two normal distributions are “mixed” together (summed). Zangari [12] first used the normal mixture model in the context of VaR calculation. The algorithm works with any finite number of normal distributions mixed together.

$$R_t = \alpha t + Y \sqrt{t} \sigma_1 Z_1 + (1 - Y) \sqrt{t} \sigma_2 Z_2$$

where Z_1 and Z_2 are two independent standard normal distributions, namely $N(0,1)$. Y is a Bernoulli random variable which takes 1 with probability p and 0 with probability $1-p$. Note that Y is independent of Z_1 and Z_2 .

Note that this mixture of normal distributions is not a normal distribution. However, it is not far away from normal. The conditional distributions $R_{t|Y=0}$ and $R_{t|Y=1}$ are both normal distributions. For this simple normal mixture model, VaR can be estimated using a trial and error approach. However, in order to construct a model in Section 5 and to handle more than two mixtures, we will illustrate the Cornish-Fisher approach. In order to use Cornish-Fisher approximation, we need to calculate the mean (μ_p), variance (σ_p^2), skewness (S) and excess kurtosis (K).

$$\begin{aligned} \mu_p &= E(R_t | Y = 1)P(Y = 1) + E(R_t | Y = 0)P(Y = 0) \\ &= \alpha t p + \alpha t (1 - p) \\ &= \alpha t \end{aligned}$$

$$\begin{aligned} \sigma_p^2 &= E((R_t - ER_t)^2 | Y = 1)P(Y = 1) + E((R_t - ER_t)^2 | Y = 0)P(Y = 0) \\ &= t\sigma_1^2 p + t\sigma_2^2 (1 - p) \end{aligned}$$

Note that conditioning on $Y=1$, R_t is a normal distribution with mean αt and variance $t\sigma_1^2$ and conditioning on $Y=0$, R_t is a normal distribution with mean αt and variance $t\sigma_2^2$. Similarly, we can find skewness, S:

$$S = \frac{E((R_t - ER_t)^3)}{\sigma_p^3} = 0$$

And excess kurtosis, K:

$$K = \frac{E((R_t - ER_t)^4)}{\sigma_p^4} - 3$$

We use the same technique to calculate:

$$\begin{aligned}
& E((R_t - ER_t)^4) = \\
& E((R_t - ER_t)^4 | Y = 1)P(Y = 1) + E((R_t - ER_t)^4 | Y = 0)P(Y = 0) \\
& = 3t^2\sigma_1^4 p + 3t^2\sigma_2^4(1-p)
\end{aligned}$$

So,

$$\begin{aligned}
K &= \frac{E((R_t - ER_t)^4)}{\sigma_p^4} - 3 \\
&= \frac{3t^2\sigma_1^4 p + 3t^2\sigma_2^4(1-p)}{(t\sigma_1^2 p + t\sigma_2^2(1-p))^2} - 3 \\
&= \frac{3p(1-p)(\sigma_1^2 - \sigma_2^2)^2}{(\sigma_1^2 p + \sigma_2^2(1-p))^2}
\end{aligned}$$

We can plug the above results into the Cornish-Fisher formula to get VaR:

$$VaR_{CF} = W(\alpha t + z_{CF}\sigma_p)$$

where

$$z_{CF} = z_c + \frac{1}{6}(z_c^2 - 1)S + \frac{1}{24}(z_c^3 - 3z_c)K - \frac{1}{36}(2z_c^3 - 5z_c)S^2,$$

where z_c is the critical number for p^{th} quantile of the standard normal distribution and W is the portfolio value.

4.4 Jump diffusion model

The second underlying distribution that is being examined due to its ability to model fat tailed distributions is the jump diffusion model. The jump diffusion model is used in Section 6 with a delta-gamma model and the Cornish-Fisher approximation approach to derive the second of our two new VaR estimation approaches.

We assume the portfolio log returns R_t follows a jump diffusion model:

$$R_t = \alpha t + \sqrt{t}\sigma Z_0 + \sum_{i=1}^{N(t)} (\nu Z_i + \mu)$$

where Z_0, \dots, Z_i, \dots are independent standard normal distributions, namely $N(0,1)$. $N(t)$ is a Poisson process with density λ . This process describes a jump process. Without the

jump, namely, $N(t)=0$, the process goes to Brownian motion. The summation term on the right hand side adds to this Brownian motion the potential for the return to jump. The jump frequency is governed by a Poisson process $N(t)$ and the jump size is a normal distribution with mean μ and standard deviation ν . Note that $N(t)$ is independent of Z_i .

The return from the jump diffusion model is not normally distributed. However, the return is not far away from normal. The conditional distribution $R_{t|N(t)=j}$ is a normal distribution. In fact it is the sum of $j+1$ independent normal distributions condition on $N(t)=j$, where the first independent normal distribution is $N(\alpha t, \sqrt{t}\sigma)$ which comes from the Brownian motion part; the next j distributions are $N(\mu, \nu)$ which come from the j independent jumps with a normal distribution mean μ and standard deviation ν . In fact, the conditional distribution $R_{t|N(t)=j}$ is normally distributed with mean $\alpha t + \mu j$ and variance $t\sigma^2 + \nu^2 j$, namely $R_{t|N(t)=j} \sim N(\alpha t + \mu j, \sqrt{t\sigma^2 + j\nu^2})$. But VaR cannot be estimated directly through the cumulative distribution function because unconditional distribution of R_t is not normal. The Cornish-Fisher approach is illustrated next.

In order to use Cornish-Fisher approximation, we need to calculate the mean (μ_p), variance (σ_p^2), skewness (S) and excess kurtosis (K).

$$\begin{aligned}\mu_p &= \sum_{j=1}^{\infty} E(R_t | N(t) = j) P(N(t) = j) \\ &= \sum_{j=1}^{\infty} (\alpha t + j\mu) P(N(t) = j) \\ &= \alpha t + \mu E(N(t)) \\ &= \alpha t + \mu \lambda t\end{aligned}$$

$$\begin{aligned}\sigma_p^2 &= \sum_{j=1}^{\infty} E((R_t - ER_t)^2 | N(t) = j) P(N(t) = j) \\ &= \sum_{j=1}^{\infty} \text{Variance}(R_t | N(t) = j) P(N(t) = j) \\ &= \sum_{j=1}^{\infty} (t\sigma^2 + j\nu^2) P(N(t) = j) \\ &= t\sigma^2 + \nu^2 \lambda t\end{aligned}$$

$$S = \frac{E((R_t - ER_t)^3)}{\sigma_p^3} = 0$$

$$K = \frac{E((R_t - ER_t)^4)}{\sigma_P^4} - 3$$

We use the same technique to calculate $E((R_t - ER_t)^4)$:

$$\begin{aligned} &= \sum_{j=1}^{\infty} E((R_t - ER_t)^4 | N(t) = j) P(N(t) = j) \\ &= \sum_{j=1}^{\infty} \text{centralized - fourth - moments}(R_t | N(t) = j) P(N(t) = j) \\ &= \sum_{j=1}^{\infty} 3 \text{Variance}(R_t | N(t) = j)^2 P(N(t) = j) \\ &= \sum_{j=1}^{\infty} 3(t\sigma^2 + v^2 j)^2 P(N(t) = j) \\ &= \sum_{j=1}^{\infty} (3t^2\sigma^4 + 3v^4 j^2 + 6t\sigma^2 v^2 j) P(N(t) = j) \\ &= 3t^2\sigma^4 + 3v^4(\lambda t + \lambda^2 t^2) + 6t\sigma^2 v^2(\lambda t) \end{aligned}$$

So,

$$\begin{aligned} K &= \frac{E((R_t - ER_t)^4)}{\sigma_P^4} - 3 \\ &= \frac{3t^2\sigma^4 + 3v^4(\lambda t + \lambda^2 t^2) + 6t\sigma^2 v^2(\lambda t)}{(t\sigma^2 + v^2 \lambda t)^2} - 3 \\ &= \frac{3v^4 \lambda t}{(t\sigma^2 + v^2 \lambda t)^2} \\ &= \frac{3v^4 \lambda}{(\sigma^2 + v^2 \lambda)^2} \frac{1}{t} \end{aligned}$$

As t goes to large, the excess kurtosis K goes to 0. This is exactly the central limit effect. Namely, as t goes to large, R_t goes close to being normally distributed.

In Duffie and Pan's survey [5], similar results as section 4.3 and 4.4 can be found, but with $t=1$. We provide the details and allow t to be any length time period.

We can insert the Cornish-Fisher formula to get VaR:

$$VaR_{CF} = W(\alpha t + \mu \lambda t + z_{CF} \sigma_P)$$

where

$$z_{CF} = z_c + \frac{1}{6}(z_c^2 - 1)S + \frac{1}{24}(z_c^3 - 3z_c)K - \frac{1}{36}(2z_c^3 - 5z_c)S^2.$$

and where z_c is the critical number for p^{th} quantile of the standard normal distribution and W is the portfolio value.

5. DELTA-GAMMA NORMAL MIXTURE MODEL

The normal mixture model (introduced in Section 4.3 and to be used in this Section) has been found to provide more accurate modeling of actual underlying factors. The purpose of Section 5 is to extend VaR estimation to the mixed normal model using the delta-gamma and Cornish-Fisher approaches as illustrated in Exhibit 5.

To apply the Cornish-Fisher approximation we need to find the mean, standard deviation, skewness and excess kurtosis in the context of the delta-gamma model and for the mixed normal model. We assume the vector of individual asset log returns R_t follows a mixed joint normal distribution:

$$R_t = \alpha t + Y\sqrt{t}\Sigma_1^{1/2}Z_1 + (1-Y)\sqrt{t}\Sigma_2^{1/2}Z_2 \quad (5.1)$$

where Z_1 and Z_2 are independent standard normal in R^m and Y is a Bernoulli (p) distribution which is independent of Z_1 and Z_2 . Σ_1 and Σ_2 are symmetric positive semi-definite matrices of dimension $m \times m$.

We know that conditioning on $Y=1$, the distribution of R_t is the multivariate normal distribution with mean αt and covariance matrix $t\Sigma_1$ while conditioning on $Y=0$, the distribution of R_t is the multivariate normal distribution with mean αt and covariance matrix $t\Sigma_2$.

Using the delta-gamma modeling approach, we consider a portfolio W_t ,

$$W_t = \Delta^T R_t + \frac{1}{2}R_t^T \Gamma R_t, \quad (5.2)$$

where Δ is $m \times 1$ vector, R_t is $m \times 1$ vector defined in (5.1), and Γ is a $m \times m$ matrix.

We need to find $\sigma_{W_t}^2$, $S(W_t)$ and $K(W_t)$.

$$\begin{aligned}
\sigma_{W_t}^2 &= E((W_t - EW_t)^2) \\
&= E((W_t - EW_t)^2 | Y = 1)p + E((W_t - EW_t)^2 | Y = 0)(1 - p) \\
&= \left\{ \frac{1}{2} \text{trace}((\Gamma \Sigma_1 t)^2) + (\Gamma \alpha t + \Delta)^T \Sigma_1 t (\Gamma \alpha t + \Delta) \right\} p + \\
&\quad \left\{ \frac{1}{2} \text{trace}((\Gamma \Sigma_2 t)^2) + (\Gamma \alpha t + \Delta)^T \Sigma_2 t (\Gamma \alpha t + \Delta) \right\} (1 - p)
\end{aligned} \tag{5.3}$$

$$\begin{aligned}
S(W_t) &= E(W_t - EW_t)^3 / \sigma_{W_t}^3 \\
&= \{E((W_t - EW_t)^3 | Y = 1)p + E((W_t - EW_t)^3 | Y = 0)(1 - p)\} / \sigma_{W_t}^3 \\
&= [\{\text{trace}((\Gamma \Sigma_1 t)^3) + 3(\Gamma \alpha t + \Delta)^T \Sigma(\Gamma \Sigma_1)(\Gamma \alpha t + \Delta)\} p \\
&\quad + \{\text{trace}((\Gamma \Sigma_2 t)^3) + 3(\Gamma \alpha t + \Delta)^T \Sigma(\Gamma \Sigma_2)(\Gamma \alpha t + \Delta)\} (1 - p)] / \sigma_{W_t}^3
\end{aligned} \tag{5.4}$$

$$K(W_t) = E(W_t - EW_t)^4 / \sigma_{W_t}^4 - 3$$

We first find $E(W_t - EW_t)^4$

$$\begin{aligned}
E((W_t - EW_t)^4) &= \\
&= E((W_t - EW_t)^4 | Y = 1)p \\
&\quad + E((W_t - EW_t)^4 | Y = 0)(1 - p) \\
&= E[(W_t - EW_t)^4 - 3(E(W_t - EW_t)^2)^2 | Y = 1]p + E\{3(E(W_t - EW_t)^2)^2 | Y = 1\}p \\
&\quad + E[(W_t - EW_t)^4 - 3(E(W_t - EW_t)^2)^2 | Y = 0](1 - p) + E\{3(E(W_t - EW_t)^2)^2 | Y = 0\}(1 - p) \\
&= \kappa_4(\alpha t, \Sigma_1 t)p + 3\sigma_{W_t}^4(\alpha t, \Sigma_1 t)p + \kappa_4(\alpha t, \Sigma_2 t)(1 - p) + 3\sigma_{W_t}^4(\alpha t, \Sigma_2 t)(1 - p) \\
&= \{3\text{trace}((\Gamma \Sigma_1 t)^4) + 12(\Gamma \alpha t + \Delta)^T \Sigma_1 t (\Gamma \Sigma_1 t)^2 (\Gamma \alpha t + \Delta)\} p \\
&\quad + \{3\text{trace}((\Gamma \Sigma_2 t)^4) + 12(\Gamma \alpha t + \Delta)^T \Sigma_2 t (\Gamma \Sigma_2 t)^2 (\Gamma \alpha t + \Delta)\} (1 - p) \\
&\quad + 3\left\{ \frac{1}{2} \text{trace}((\Gamma \Sigma_1 t)^2) + (\Gamma \alpha t + \Delta)^T \Sigma_1 t (\Gamma \alpha t + \Delta) \right\}^2 p \\
&\quad + 3\left\{ \frac{1}{2} \text{trace}((\Gamma \Sigma_2 t)^2) + (\Gamma \alpha t + \Delta)^T \Sigma_2 t (\Gamma \alpha t + \Delta) \right\} (1 - p)
\end{aligned}$$

So,

$$\begin{aligned}
K(W_t) = & \{3\text{trace}((\Gamma\Sigma_1 t)^4) + 12(\Gamma\alpha t + \Delta)^T \Sigma_1 t (\Gamma\Sigma_1 t)^2 (\Gamma\alpha t + \Delta)\} p \\
& + \{3\text{trace}((\Gamma\Sigma_2 t)^4) + 12(\Gamma\alpha t + \Delta)^T \Sigma_2 t (\Gamma\Sigma_2 t)^2 (\Gamma\alpha t + \Delta)\} (1-p) \\
& + 3\left\{\frac{1}{2}\text{trace}((\Gamma\Sigma_1 t)^2) + (\Gamma\alpha t + \Delta)^T \Sigma_1 t (\Gamma\alpha t + \Delta)\right\}^2 p \\
& + 3\left\{\frac{1}{2}\text{trace}((\Gamma\Sigma_2 t)^2) + (\Gamma\alpha t + \Delta)^T \Sigma_2 t (\Gamma\alpha t + \Delta)\right\}^2 (1-p) / \sigma_{w_i}^4 - 3
\end{aligned} \tag{5.5}$$

If we let $\Gamma = 0, \Sigma_1 = \sigma_1^2, \Sigma_2 = \sigma_2^2, \Delta = 1$, it will reduce to model in section 4.3.

$$\text{We have } \sigma_{w_i}^2 = \Sigma_1 t p + \Sigma_2 t (1-p) = \sigma_1^2 p t + \sigma_2^2 (1-p) t \tag{5.6}$$

$$S(W_t) = 0 \tag{5.7}$$

$$\begin{aligned}
K(W_t) &= [3(\Sigma_1 t)^2 p + 3(\Sigma_2 t)^2 (1-p)] / \sigma_{w_i}^2 - 3 \\
&= [3(\sigma_1^2 t)^2 p + 3(\sigma_2^2 t)^2 (1-p)] / \sigma_{w_i}^2 - 3 \\
&= \frac{3(\sigma_1^2 t)^2 p + 3(\sigma_2^2 t)^2 (1-p)}{(\sigma_1^2 p t + \sigma_2^2 (1-p) t)^2} - 3
\end{aligned} \tag{5.8}$$

The final VaR estimation equation for a mixed normal distribution using the delta-gamma approach simply requires that the above cumulants be inserted into the Cornish-Fisher approximation equation.

6. DELTA-GAMMA JUMP DIFFUSION MODEL:

Like Section 5, this section utilized the delta-gamma approach (for non-linear portfolios) and the Cornish-Fisher expansion (for non-normal distributions) to estimate VaR for a non-normal distribution that has been found to provide superior modeling of actual financial data of underliers. While Section 5 used the mixed normal distribution for the underlying return factors, this section uses the jump diffusion model (which was introduced in Section 4.4) as illustrated in Exhibit 6.

We assume the vector of individual asset log returns R_t follows a jump diffusion model:

$$R_t = \alpha t + \sqrt{t} \Sigma^{1/2} Z_0 + \sum_{i=1}^{N(t)} (\mu + V^{1/2} Z_i) \tag{6.1}$$

where Z_i and Z_j are independent standard normal in R^m for $i \neq j$ and $N(t)$ is a jump counting Poisson process which is independent of Z_i for $i = \{0, 1, \dots\}$, Σ and V are symmetric positive semi-definite matrices of dimension $m \times m$.

We know that conditioning on $N(t) = j$, the distribution of R_t is a multivariate normal distribution with mean $\alpha t + \mu j$ and covariance matrix $t\Sigma + jV$.

Using the delta-gamma approach to the modeling of a non-linear relationship between a portfolio, W_t and its underlying return factors,

$$W_t = \Delta^T R_t + \frac{1}{2} R_t^T \Gamma R_t, \quad (6.2)$$

where Δ , R_t are $m \times 1$ vectors, and Γ is a symmetric $m \times m$ matrix.

To handle the non-normality using the Cornish-Fisher approach, it is necessary to find $\sigma_{W_t}^2$, $S(W_t)$, and $K(W_t)$. We first define the Poisson process probability $p_j = P(N(t) = j)$

$$\begin{aligned} \sigma_{W_t}^2 &= E((W_t - EW_t)^2) \\ &= \sum_{j=0}^{\infty} E((W_t - EW_t)^2 | N(t) = j) p_j \\ &= \sum_{j=0}^{\infty} \left\{ \frac{1}{2} \text{trace}((\Gamma \Sigma t + j \Gamma V)^2) + (\Gamma \alpha t + \Gamma j \mu + \Delta)^T (\Sigma t + jV) (\Gamma \alpha t + \Gamma j \mu + \Delta) \right\} p_j \end{aligned} \quad (6.3)$$

$$\begin{aligned} S(W_t) &= E(W_t - EW_t)^3 / \sigma_{W_t}^3 \\ &= \left[\sum_{j=0}^{\infty} E((W_t - EW_t)^3 | N(t) = j) p_j \right] / \sigma_{W_t}^3 \\ &= \left[\sum_{j=0}^{\infty} \text{trace}((\Gamma \Sigma t + j \Gamma V)^3) + 3(\Gamma \alpha t + \Gamma j \mu + \Delta)^T (\Sigma t + jV) (\Gamma \Sigma t + \Gamma jV) (\Gamma \alpha t + \Gamma j \mu + \Delta) \right] p_j / \sigma_{W_t}^3 \end{aligned} \quad (6.4)$$

$$K(W_t) = E(W_t - EW_t)^4 / \sigma_{W_t}^4 - 3$$

We first find $E(W_t - EW_t)^4$

$$\begin{aligned}
& E((W_t - EW_t)^4) = \\
& = \sum_{j=0}^{\infty} E((W_t - EW_t)^4 | N(t) = j) p_j \\
& = \sum_{j=0}^{\infty} \{E[(W_t - EW_t)^4 - 3(E(W_t - EW_t)^2)^2 | N(t) = j] p_j + E\{3(E(W_t - EW_t)^2)^2 | N(t) = j\} p_j\} \\
& = \sum_{j=0}^{\infty} \{\kappa_4(\alpha t + \mu j, \Sigma t + jV) p_j + 3\sigma_{W_t}^4 (\alpha t + \mu j, \Sigma t + jV) p_j\} \\
& = \sum_{j=0}^{\infty} \{3\text{trace}((\Gamma\Sigma t + j\Gamma V)^4) + 12(\Gamma\alpha t + \Gamma j\mu + \Delta)^T (\Sigma t + jV)(\Gamma\Sigma t + \Gamma jV)^2 (\Gamma\alpha t + \Gamma j\mu + \Delta)\} p_j + \\
& 3 \sum_{j=0}^{\infty} \left\{ \frac{1}{2} \text{trace}((\Gamma\Sigma t + \Gamma jV)^2) + (\Gamma\alpha t + \Gamma j\mu + \Delta)^T (\Sigma t + jV)(\Gamma\alpha t + \Gamma j\mu + \Delta) \right\}^2 p_j
\end{aligned}$$

So,

$$\begin{aligned}
K(W_t) &= \left[\sum_{j=0}^{\infty} \{3\text{trace}((\Gamma\Sigma t + j\Gamma V)^4) + 12(\Gamma\alpha t + \Gamma j\mu + \Delta)^T (\Sigma t + jV)(\Gamma\Sigma t + \Gamma jV)^2 (\Gamma\alpha t + \Gamma j\mu + \Delta)\} p_j + \right. \\
& \left. 3 \sum_{j=0}^{\infty} \left\{ \frac{1}{2} \text{trace}((\Gamma\Sigma t + \Gamma jV)^2) + (\Gamma\alpha t + \Gamma j\mu + \Delta)^T (\Sigma t + jV)(\Gamma\alpha t + \Gamma j\mu + \Delta) \right\}^2 p_j \right] / \sigma_{W_t}^4 - 3
\end{aligned}$$

(6.5)

If we let $\Gamma = 0$, $\Sigma = \sigma^2$, $V = \nu^2$ and $\Delta = 1$, it will reduce to model in section 4.4. We have

$$\sigma_{W_t}^2 = \sum_{j=0}^{\infty} (\sigma^2 t + j\nu^2) p_j = \sigma^2 t + \nu^2 E(N(t)) = \sigma^2 t + \nu^2 \lambda t \quad (6.6)$$

$$S(W_t) = 0 \quad (6.7)$$

$$\begin{aligned}
K(W_t) &= \left[3 \sum_{j=0}^{\infty} (\sigma^2 t + \nu^2 j)^2 p_j \right] / \sigma_{W_t}^2 - 3 \\
&= [3\sigma^4 t^2 + 6\sigma^2 \nu^2 t(\lambda t) + 3\nu^4 (\lambda t + \lambda^2 t^2)] / \sigma_{W_t}^2 - 3 \\
&= \frac{3\sigma^4 t^2 + 6\sigma^2 \nu^2 \lambda t^2 + 3\nu^4 \lambda t + 3\nu^4 \lambda^2 t^2}{(\sigma^2 t + \nu^2 \lambda t)^2} - 3 \quad (6.8)
\end{aligned}$$

$$= \frac{3\nu^4 \lambda t}{(\sigma^2 t + \nu^2 \lambda t)^2}$$

$$= \frac{3\nu^4 \lambda}{(\sigma^2 + \nu^2 \lambda)^2 t}$$

This matches models in section 4.4. If we let $N(t) = 0$, it is easy to check that the formulas of $\sigma_{W_t}^2$, $S(W_t)$, and $K(W_t)$ will match with models in section 4.2, the normal case which does not have a jump. The final VaR estimation equation for a jump diffusion process using the delta-gamma approach simply requires that the above cumulants be inserted into the Cornish-Fisher approximation equation.

7. SUMMARY AND CONCLUSIONS

Exhibit 7 illustrates the relationships between the VaR estimation methodologies reviewed and derived in this paper.

VaR estimation is made complex by two common challenges: (1) many financial returns such as hedge fund returns and other dynamic portfolio strategies are distributed with skew and leptokurtosis, and (2) many portfolios such as those containing derivatives or hedge funds contain non-linear contracts that generate non-normal returns even if the underlying securities are normally distributed. Section 4 reviewed the Cornish-Fisher approach for addressing non-normality and the delta-gamma approach for addressing non-linearity.

Sections 5 and 6 used the delta-gamma approach and the Cornish-Fisher approach to derive formulas for estimating VaR in the case of two distributions that have been found to provide potentially superior modeling of underlying returns factors: (1) the mixed normal model, and (2) the jump diffusion process. Both models are extensions of the existing literature and may be useful in providing reasonably simple, fast and accurate estimation or VaR.

APPENDIX A: SKEWNESS AND KURTOSIS FOR DELTA-GAMMA-NORMAL PORTFOLIOS.

This appendix provides a detailed derivation of skewness and excess kurtosis in the case of applying the delta-gamma approach to a portfolio with normally distributed underlying return factors. The higher cumulants enable the application of the Fisher-Cornish approximation equation. The approach detailed in this appendix also serves as a foundation to the models derived in Sections 5 and 6 using the mixed normal and jump diffusion models for the underlying return factors.

The result in this appendix is in Jaschke's paper without details ([14]).

$$W = A + B^T X + \frac{1}{2} X^T \Gamma X \quad (A1)$$

where A is a constant, B, X are $m \times 1$ vectors, and Γ is a symmetric $m \times m$ matrix. We assume that X follows multivariate Normal distribution with mean 0 and covariance matrix Σ where Σ is symmetric positive semi-definite matrix of dimension $m \times m$. We can view X as m assets and W is the portfolio related to these m assets. We would like to find σ_W^2 , S(W) and K(W), which are the variance, the skewness and the excess kurtosis for portfolio W. We first investigate the easiest case:

$$\text{Case 1: } V = \delta Y + \frac{1}{2} \lambda Y^2 \quad (A2)$$

where δ, λ are constants and Y follows a one variable standard normal distribution N(0,1). We know that

$$\begin{aligned} V &= \delta Y + \frac{1}{2} \lambda Y^2 \\ &= \frac{1}{2} \lambda \left(Y^2 + \frac{2\delta}{\lambda} Y \right) \\ &= \frac{1}{2} \lambda \left(Y + \frac{\delta}{\lambda} \right)^2 - \frac{1}{2} \lambda \frac{\delta^2}{\lambda^2} \\ &= \frac{1}{2} \lambda \left(Y + \frac{\delta}{\lambda} \right)^2 - \frac{1}{2} \frac{\delta^2}{\lambda} \end{aligned}$$

Since Y follows N(0,1), we know that $\left(Y + \frac{\delta}{\lambda} \right)^2$ is a non central chi-square distribution

with $\mu = \frac{\delta}{\lambda}, \sigma = 1$ and degree of freedom k=1.

The non-central chi-square distribution has two parameters θ and k . In this case, $\theta = \frac{\delta^2}{\lambda^2}$ and $k=1$.

It is well known that non-central chi-square distribution has a moment generating function:

$$\frac{\exp\left(\frac{\theta t}{1-2t}\right)}{(1-2t)^{k/2}}$$

So, $(Y + \frac{\delta}{\lambda})^2$ has a moment generating function: $M(t) = \frac{\exp\left(\frac{\frac{\delta^2}{\lambda^2}t}{1-2t}\right)}{(1-2t)^{1/2}}$

Similarly, $(Y + \frac{\delta}{\lambda})^2$ has a characteristic function: $\varphi(t) = \frac{\exp\left(\frac{\frac{\delta^2}{\lambda^2}it}{1-2it}\right)}{(1-2it)^{1/2}}$

It is easy to see $V = \frac{1}{2} \lambda (Y + \frac{\delta}{\lambda})^2 - \frac{1}{2} \frac{\delta^2}{\lambda}$ with a moment generating function:

$$\begin{aligned} M_V(t) &= \exp\left(-\frac{1}{2} \frac{\delta^2}{\lambda} t\right) \frac{\exp\left(\frac{\frac{\delta^2}{\lambda^2} \left(\frac{1}{2} \lambda t\right)}{1-2\left(\frac{1}{2} \lambda t\right)}\right)}{\left(1-2\frac{1}{2} \lambda t\right)^{1/2}} = \frac{1}{\sqrt{1-\lambda t}} \exp\left(\frac{\frac{1}{2} \frac{\delta^2 t}{\lambda}}{1-\lambda t} - \frac{1}{2} \frac{\delta^2}{\lambda} t\right) \\ &= \frac{1}{\sqrt{1-\lambda t}} \exp\left(\frac{\frac{1}{2} \delta^2 t^2}{1-\lambda t}\right) \end{aligned} \quad (A3)$$

$$\text{Similarly, } \varphi_V(t) = \frac{1}{\sqrt{1-i\lambda t}} \exp\left(\frac{\frac{1}{2} \frac{\delta^2 it}{\lambda}}{1-i\lambda t} - \frac{1}{2} \frac{\delta^2}{\lambda} it\right) = \frac{1}{\sqrt{1-i\lambda t}} \exp\left(\frac{-\frac{1}{2} \delta^2 t^2}{1-i\lambda t}\right) \quad (A4)$$

Now we consider the next case:

Case 2:

Let $W = A + \sum_{i=1}^m (\delta_i Y_i + \frac{1}{2} \lambda_i Y_i^2)$, where A, δ_i, λ_i are constants and Y_i are independent standardized normal distributed random variables. (A5)

From the properties of moment generating functions, it is easy to get

$$M_W(t) = \exp(At) \frac{1}{\prod_{i=1}^m \sqrt{1-\lambda_i t}} \exp\left(\sum_{i=1}^m \left(\frac{\frac{1}{2} \delta_i^2 t}{1-\lambda_i t} - \frac{1}{2} \frac{\delta_i^2}{\lambda_i} t\right)\right) = \exp(At) \frac{1}{\prod_{i=1}^m \sqrt{1-\lambda_i t}} \exp\left(\sum_{i=1}^m \frac{\frac{1}{2} \delta_i^2 t^2}{1-\lambda_i t}\right)$$

$$= \frac{\exp\left(At + \sum_{i=1}^m \frac{\frac{1}{2} \delta_i^2 t^2}{1-\lambda_i t}\right)}{\prod_{i=1}^m \sqrt{1-\lambda_i t}} \quad (A6)$$

Similarly, we can get the characteristic function:

$$\varphi_W(t) = \frac{\exp\left(At + \sum_{i=1}^m \frac{-\frac{1}{2} \delta_i^2 t^2}{1-\lambda_i it}\right)}{\prod_{i=1}^m \sqrt{1-\lambda_i it}} \quad (A7)$$

The cumulant generating function (defined as the natural log of the moment generating function) is:

$$\kappa_W(t) = \ln M_W(t) = At + \sum_{i=1}^m \left\{ \frac{\frac{1}{2} \delta_i^2 t^2}{1-\lambda_i t} - \frac{1}{2} \ln(1-\lambda_i t) \right\} \quad (A8)$$

We know that $\kappa_W(t) = \sum_{r=1}^{\infty} \frac{\kappa_r t^r}{r!}$, where κ_r are cumulants. (A9)

Cumulants are closely related to variance, skewness and excess kurtosis.

It is well known that κ_2 is the variance, κ_3 is third central moments, so $\kappa_3 / \kappa_2^{3/2}$ is the skewness and κ_4 / κ_2^2 is the excess kurtosis.

In order to get κ_r from equation (A8), we need to have a power series representation of the following:

$$\frac{1}{1-\lambda_i t} = 1 + \sum_{r=1}^{\infty} \lambda_i^r t^r, \quad (\text{A10})$$

$$-\frac{1}{2} \ln(1-\lambda_i t) = -\frac{1}{2} \sum_{r=1}^{\infty} -\lambda_i^r t^r / r = \sum_{r=1}^{\infty} \frac{1}{2} \frac{\lambda_i^r}{r} t^r = \frac{1}{2} \lambda_i t + \sum_{r=2}^{\infty} \frac{1}{2} \frac{\lambda_i^r}{r} t^r \quad (\text{A11})$$

$$\frac{\frac{1}{2} \delta_i^2 t^2}{1-\lambda_i t} = \frac{1}{2} \delta_i^2 t^2 + \frac{1}{2} \delta_i^2 t^2 \sum_{r=1}^{\infty} \lambda_i^r t^r = \frac{1}{2} \delta_i^2 t^2 + \sum_{r=1}^{\infty} \frac{1}{2} \delta_i^2 \lambda_i^r t^{r+2} = \sum_{r=2}^{\infty} \frac{1}{2} \delta_i^2 \lambda_i^{r-2} t^r \quad (\text{A12})$$

Therefore, we can insert equations (A11) and (A12) into equation (A8),

$$\begin{aligned} \kappa_W(t) &= At + \sum_{i=1}^m \left\{ \frac{\frac{1}{2} \delta_i^2 t^2}{1-\lambda_i t} - \frac{1}{2} \ln(1-\lambda_i t) \right\} \\ &= At + \left(\sum_{i=1}^m \frac{1}{2} \lambda_i \right) t + \sum_{r=2}^{\infty} \left\{ \sum_{i=1}^m \left(\frac{1}{2} \frac{\lambda_i^r}{r} + \frac{1}{2} \delta_i^2 \lambda_i^{r-2} \right) \right\} t^r \end{aligned}$$

So, by equation (A9), we have $\kappa_1 = A + \frac{1}{2} \sum_{i=1}^m \lambda_i$ (A13)

$$\kappa_r = r! \sum_{i=1}^m \left(\frac{1}{2} \frac{\lambda_i^r}{r} + \frac{1}{2} \delta_i^2 \lambda_i^{r-2} \right) = \frac{1}{2} (r-1)! \sum_{i=1}^m \lambda_i^r + \frac{1}{2} r! \sum_{i=1}^m \delta_i^2 \lambda_i^{r-2} \quad \text{for } r \geq 2 \quad (\text{A14})$$

Now we consider our last case, namely the equation (A1):

Case 3:

Recall equation (A1),

$$W = A + B^T X + \frac{1}{2} X^T \Gamma X$$

where X follows a normal distribution with mean 0 and covariance matrix Σ .

Now we need to express equation (A1) as a sum of independent random variables, namely equation (A5) in our case 2:

$$W = A + \sum_{i=1}^m (\delta_i Y_i + \frac{1}{2} \lambda_i Y_i^2) = A + (\delta_1, \dots, \delta_m) Y + \frac{1}{2} Y^T \begin{pmatrix} \lambda_1 & & \\ & \dots & \\ & & \lambda_m \end{pmatrix} Y.$$

We can use standard linear algebra steps to transform A1 to A5. The steps are as follows:

- (1) Do a Cholesky decomposition about Σ , the covariance matrix of X : Let $\Sigma = CC^T$.
- (2) For matrix $C^T \Gamma C$, solve the eigenvalue problem, namely find orthogonal matrix Q such that $Q^T C^T \Gamma C Q = \Lambda$, where Λ is the diagonal eigenvalue matrix.

The relations between A1 and A5 are:

$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \dots & \\ & & \lambda_m \end{pmatrix} \text{ and } Q^T C^T B = \begin{pmatrix} \delta_1 \\ \dots \\ \delta_m \end{pmatrix}$$

Now we would like to write κ_1 and κ_r in equations (A13) and (A14) into the matrix forms:

By matrix algebra, it is easy to see that $\begin{pmatrix} \lambda_1^r & & \\ & \dots & \\ & & \lambda_m^r \end{pmatrix} = \Lambda^r = Q^T C^T (\Gamma \Sigma)^{r-1} \Gamma C Q$

$$\sum_{i=1}^m \lambda_i^r = \text{trace}(\Lambda^r) = \text{trace}(Q^T C^T (\Gamma \Sigma)^{r-1} \Gamma C Q) = \text{trace}((\Gamma \Sigma)^{r-1} \Gamma C Q Q^T C^T)$$

$$= \text{trace}((\Gamma \Sigma)^r) \text{ (Because } QQ^T = I \text{ and } CC^T = \Sigma)$$

$$\begin{aligned} \sum_{i=1}^m \delta_i^2 \lambda_i^{r-2} &= (\delta_1 \quad \dots \quad \delta_m) \Lambda^{r-2} \begin{pmatrix} \delta_1 \\ \dots \\ \delta_m \end{pmatrix} = (Q^T C^T B)^T Q^T C^T (\Gamma \Sigma)^{r-3} \Gamma C Q (Q^T C^T B) \\ &= B^T C Q Q^T C^T (\Gamma \Sigma)^{r-3} \Gamma C Q (Q^T C^T B) \\ &= B^T \Sigma (\Gamma \Sigma)^{r-3} \Gamma \Sigma B \\ &= B^T \Sigma (\Gamma \Sigma)^{r-2} B \end{aligned}$$

$$\text{So, } \kappa_1 = A + \frac{1}{2} \text{trace}(\Gamma \Sigma) \tag{A15}$$

$$\kappa_r = \frac{1}{2} (r-1)! \text{trace}((\Gamma \Sigma)^r) + \frac{1}{2} r! B^T \Sigma (\Gamma \Sigma)^{r-2} B \text{ for } r \geq 2 \tag{A16}$$

$$\text{It is easy to see } \kappa_2 = \frac{1}{2} \text{trace}((\Gamma \Sigma)^2) + B^T \Sigma B$$

$$\kappa_3 = \text{trace}((\Gamma \Sigma)^3) + 3B^T \Sigma (\Gamma \Sigma) B$$

$$\kappa_4 = 3\text{trace}((\Gamma \Sigma)^4) + 12B^T \Sigma (\Gamma \Sigma)^2 B$$

So, for portfolio W, we have

$$\text{The variance } \sigma_w^2 = \kappa_2 = \frac{1}{2} \text{trace}((\Gamma\Sigma)^2) + B^T \Sigma B \quad (\text{A17})$$

$$\text{The skewness } S(W) = \frac{\kappa_3}{\kappa_2^{3/2}} = \frac{\text{trace}((\Gamma\Sigma)^3) + 3B^T \Sigma(\Gamma\Sigma)B}{\left(\frac{1}{2} \text{trace}((\Gamma\Sigma)^2) + B^T \Sigma B\right)^{3/2}} \quad (\text{A18})$$

$$\text{The excess kurtosis } K(W) = \frac{\kappa_4}{\kappa_2^2} = \frac{3\text{trace}((\Gamma\Sigma)^4) + 12B^T \Sigma(\Gamma\Sigma)^2 B}{\left(\frac{1}{2} \text{trace}((\Gamma\Sigma)^2) + B^T \Sigma B\right)^2} \quad (\text{A19})$$

REFERENCES:

1. Turan G. Bali, An extreme value approach to estimating volatility and Value at risk, *Journal of Business*, vol. 76, no.1, 2003.
2. M. Britton-Jones, S. Schaefer, Non-linear Value-at-Risk, p.161-187, *European Finance Review* vo.2, 1999.
3. Victor Chernozhukov, Ivan Fernandez-Val, Alfred Galichon, Rearranging Edgeworth-Cornish-Fisher Expansions, *submitting paper*, 2007.
4. Kevin Dowd, *Beyond Value at Risk: The New Science of Risk Management*, *Wiley Frontiers in Finance Series*, 1998.
5. Darrell Duffie, Pan Jun, An Overview of Value at Risk, p. 7-49, *Journal of Derivatives*, Spring 1997, reprinted in *Options Markets*, edited by G. Constantinides and A. G. Malliaris, London: Edward Elgar, 2001.
6. Darrell Duffie, Pan Jun, Analytical Value-At-Risk with Jumps and Credit Risk, p155—180, *Finance and Stochastics*, Volume 5, 2001.
7. Lina El-Jahel, William Perraudin, Peter Sellin, Value at Risk for Derivatives, *Journal of Derivatives*, Spring 1999.
8. Robert F. Engle, Simone Manganelli, Value at risk models in finance, European central bank working paper no.75, August 2001.
9. Robert F. Engle, Simone Manganelli, CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles; *Journal of Business & Economic Statistics*, Vol. 22, 2004. (It has early equivalent version.)
10. Laurent Favre; Jose-Antonio Galeano, Mean-Modified Value-at-Risk Optimization with Hedge Funds, *The Journal of Alternative Investments*, Fall 2002.
11. P. Glasserman, P. Heidelberger, P. Shahabuddin, Portfolio Value-at-Risk with Heavy-Tailed Risk Factors, 239-270, *Mathematical Finance*, vol. 12, 2002.
12. Glyn A. Holton, *Value-at-Risk: Theory and Practice* by Academic Press, Elsevier, 2003.
13. John C. Hull, *Options, Futures, and Other Derivatives*, Pearson Prentice Hall, 6th edition, 2006.

14. Stefan R Jaschke, The Cornish-Fisher-Expansion in the Context of Delta - Gamma - Normal Approximations, p.33-52, *Journal of Risk* v4, n4, Summer 2002.
15. Roger W. Koenker, Gilbert W. Bassett, Jr., Regression Quantiles, p33-50, *Econometrica*, Econometric Society , 46(1), 1978.
16. Natalya Lyzanets, Maksym Senchyna , Comparing different Value-at-Risk models for hedge funds, *Master thesis in Banking and Finance*, 2005.
17. Ruey S. Tsay , Analysis of Financial Time Series, *Wiley Series in Probability and Statistics*, 2005.
18. Zangari, P., A VaR Methodology for Portfolios that Include Options, p. 4-12, *RiskMetrics Monitor*, First Quarter 1996.

Exhibit 1: Illustration of VaR

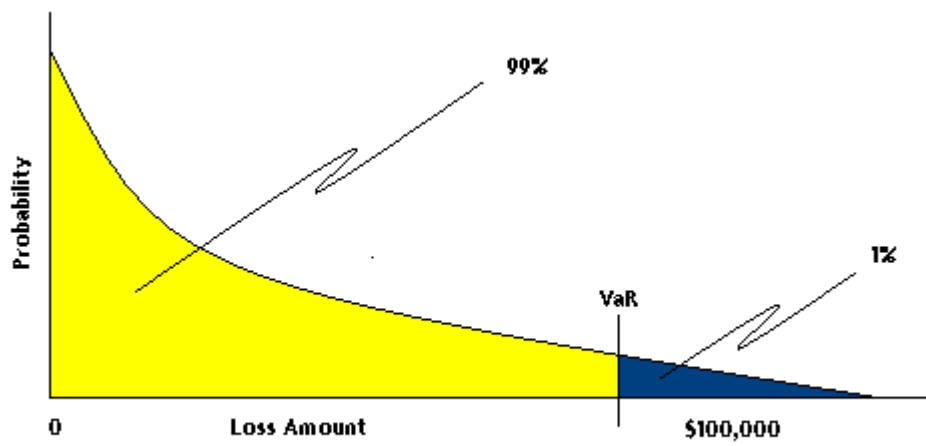


Exhibit 2: Major Categories of VaR Estimation Approaches

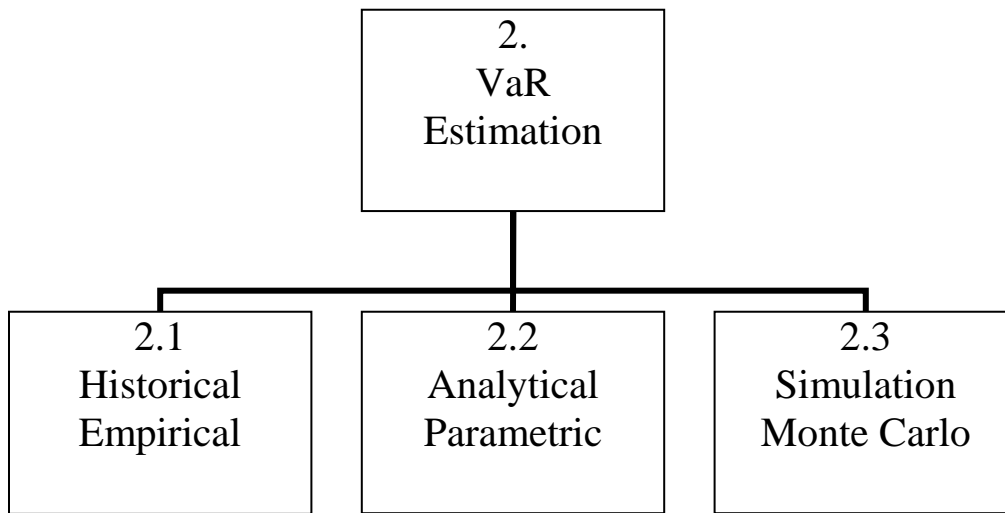


Exhibit 3: Categories of Analytical/Parametric VaR Estimation Approaches

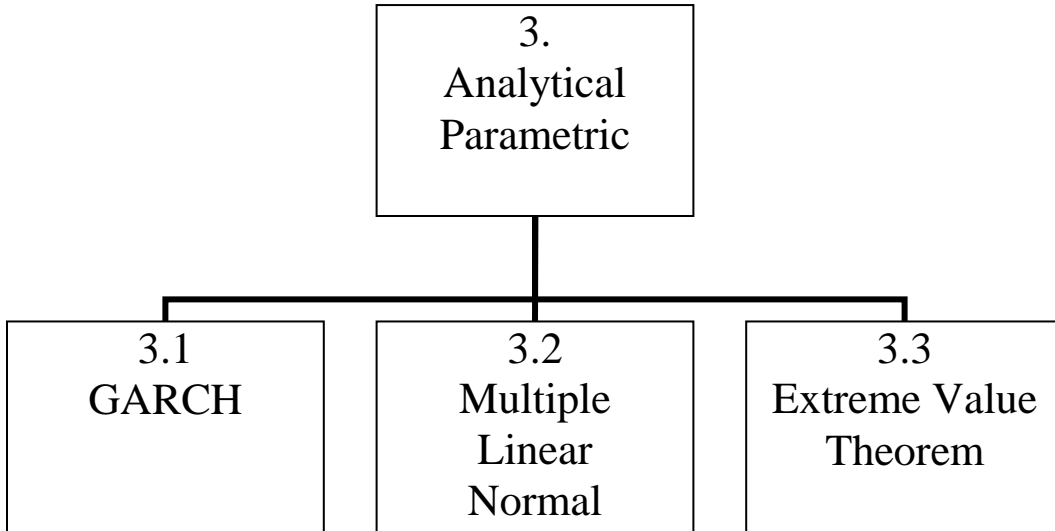


Exhibit 4: Advanced Analytical/Parametric Extensions of VaR

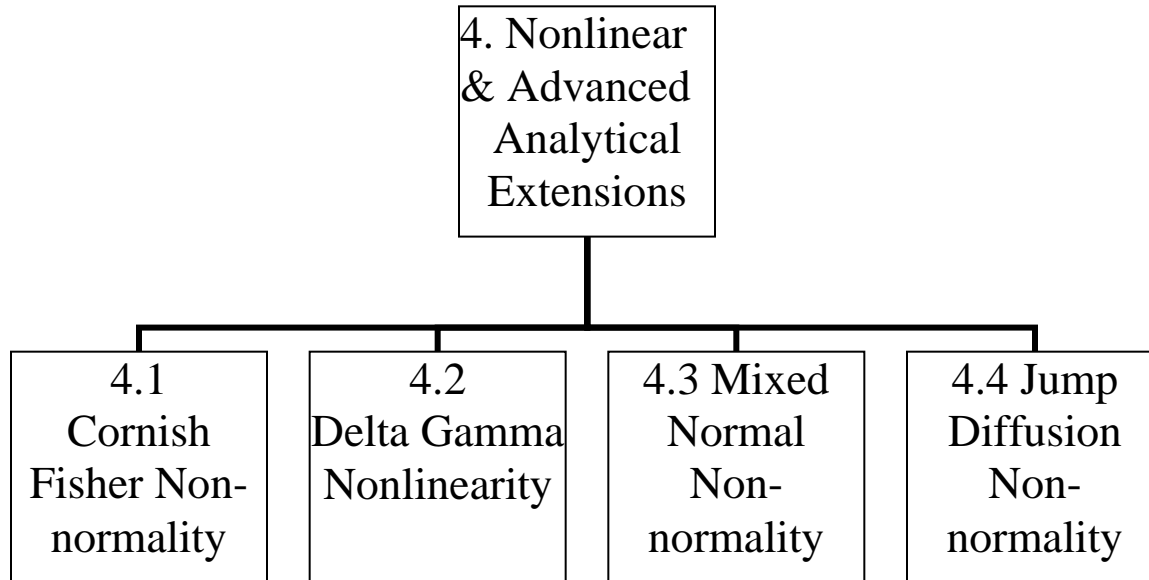


Exhibit 5: Delta-Gamma Normal Mixture Model

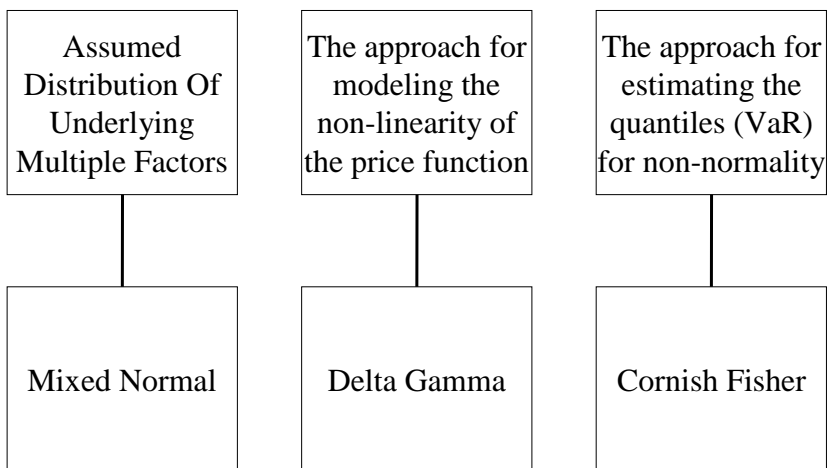


Exhibit 6: Delta-Gamma Normal Jump Diffusion Model

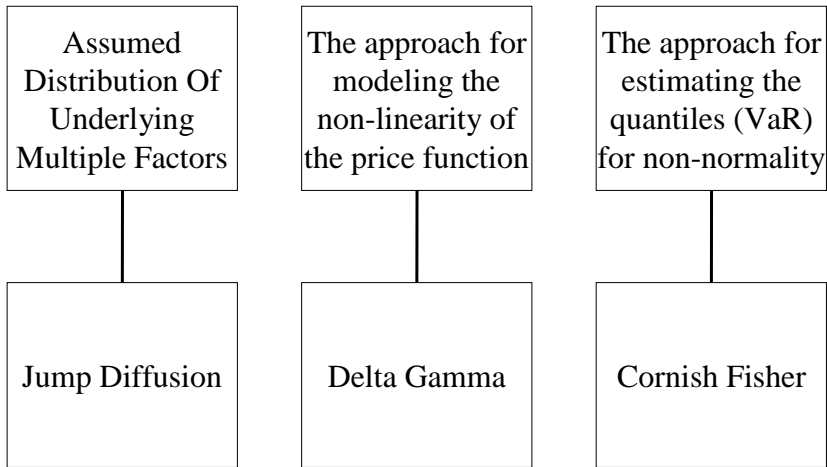


Exhibit 7: Overview of Major VaR Estimation Methodologies

