Value at Risk Based on the Volatility, Skewness and Kurtosis*

David X. Li
Riskmetrics Group
44 Wall Street
New York, NY 10005
Phone: (212)-981-7453
Fax: (212)-981-7402
Email:david.li@riskmetrics.com

March 4, 1999

Abstract

This paper presents a new approach to calculating Value-at-Risk (VaR) in which skewness and kurtosis as well as the standard deviation or volatility are explicitly used. Based on the theory of estimating functions in statistics we construct an approximate confidence interval from the first two moment conditions. The final result shows explicitly how the confidence interval is affected by the standard deviation, skewness and kurtosis. We test our method using ten years of daily observations on twelve different foreign exchange spot rates and find the new approach captures the extreme tail much better than the standard VaR calculation method used in Riskmetrics^{TM}.

*The author thanks his colleagues, Chris Finger, Joonwoo Kim, Allan Malz, Jim Ye at RiskMetrics Group, and Peter J. Zangari at Goldman, Sachs & Co. for helpful discussions and suggestions.

1 Introduction

Value at Risk (VaR) has become a popular risk management technique in the last few years. One driving force behind the popularity of this technique is the release to the public of JP Morgan RiskMetrics^{TM} Technical Document (1997) and the subsequent BIS adoption of VaR risk report for all trading portfolios of financial institutions. The technical document provides a benchmark for VaR calculation based on the statistical confidence interval constructed under the assumption of a normal distribution.

However many empirical studies of time series data show that the rate of return or percentage change of many financial variables are not normally distributed. These series tend to be skewed and leptokurtic. As shown by Zangari (1996), the VaR calculated under the normal assumption underestimates the actual risk since the distribution of many observed financial return series have tails that are "fatter" than those implied by conditional normal distribution. How to incorporate these observations into the VaR calculation is an important issue.

There exist mainly two methods to construct VaR if we do not assume normality, parametric and non-parametric approaches. In the parametric approach alternative distributions are explicitly assumed instead of normal distribution. For example, Hull and White (1998) suggest to use alternative distributions, such as a mixture of two normal distributions, to model the return of financial assets, and then use a percentile-to-pecentile mapping between this alternative distribution and the normal distribution to obtain the VaR. In the nonparametric approach no particular distribution assumption is made, and VaR is calculated using the standard theory of order statistics (see Kupiec (1995) or Monte Carlo simulation. In standard Monte Carlo simulations, it is well-known that the precision of the estimated VaR increases with the square root of the number of simulation runs. Large sample is needed to have a stable VaR result for 99% confidence interval, which makes the Monte Carlo approach a quite expensive practice.

This paper presents an alternative approach to the construction of confidence intervals based on a semiparametric setting. In general, we need either a fully-fledged distribution assumption or a Monte Carlo simulation to build confidence intervals. Since this new semiparametric approach uses only moment conditions up to the fourth order, it allows us to incorporate empirical findings on moments directly. In the mean time it is not as restrictive as the parametric model, or as expensive as the Monte Carlo simulation approach.

Since we do not assume any sample distribution the confidence interval we obtain is an approximate confidence interval based on large sample asymptotic theory. The proposed approach is consistent with the statistical method of confidence interval construction using pivotal quantities. A pivotal quantity or ancillary statistic is defined as a function of the data and parameter having a fixed distribution the same for all parameter values. For example in the case of obtaining a confidence interval for the mean parameter μ of a normal distribution with a known variance σ^2, we can take the pivotal quantity

$$\frac{X-\mu}{\sigma}$$

to construct a confidence interval since this statistic follows the standard normal distribution for all parameter values of μ. The pivotal quantity we use here involves a higher order item of the observation X, then the skewness and the kurtosis come into the final confidence interval expression explicitly.

The underlying theory we use here is called the theory of estimating functions, which has been becoming a popular statistical theory in the last decade. The theory of estimating functions generalizes and unifies many existing statistical theories and has extensive application in generalized linear statistical models, sampling theory and biostatistics. For an exposure to the theory of estimating functions we refer to Godambe (1991).

2 The Current RiskMetrics^{TM} Approach

Before we present our new approach, it is beneficial to review the RiskMetrics^{TM} approach. RiskMetrics^{TM} assumes that returns follow a conditional normal distribution. Suppose that the return series is $X_t, t = 1, 2, \cdots n$, and the volatility series of the return is $\sigma_t, t = 1, 2, \cdots n$. The variable X_t is not normally distributed, but the ratio of the return over the volatility, X_t/σ_t, follows a standard normal distribution. RiskMetrics^{TM} uses the exponentially smoothed historical data to estimate the volatility series. This approach has the following two advantage:

1. The unconditional series X_t has a fatter tail than the conditional one, X_t/σ_t
2. The explicit modeling of the volatility series captures the time-varying, persistent volatility observed in real financial markets

In summary we essentially assume that the original return series is not normally distributed, but that a transformed series is. In the RiskMetrics^{TM} framework the transformation is

$$f(x_t) = \frac{x_t}{\sigma_t}.$$

This approach has the same motivation as the pivotal quantity approach to the construction of confidence intervals. Essentially we need to construct a normally distributed pivotal quantity, which is a function of both the sample data and the parameter to be estimated. Then we can solve for the confidence interval of the parameter. In the RiskMetrics^{TM} approach the pivotal quantity is simply the original return series divided by the time varying volatility. Using the theory of estimating functions we can find an alternative pivotal quantity involving higher order terms of the return as follows.

3 The Estimating Function Approach

Suppose we have a random variable X, whose mean, variance, skewness and kurtosis are defined as follows

$$\begin{aligned}
\mu &= E(X), \\
\sigma^2 &= Var(X), \\
\gamma_1 &= \frac{E(X-\mu)^3}{\sigma^3}, \\
\gamma_2 &= \frac{E(X-\mu)^4}{\sigma^4} - 3.
\end{aligned} \tag{1}$$

We consider one sample from the distribution of X using the theory of estimating functions in statistics. For the basic concepts of the theory, see Godambe (1991). For a concise summary we refer to Li and Turtle (1997).

From the first two moment conditions in (1) we have two basic estimating functions as follows

$$\begin{aligned}
h_1 &= X - \mu, \\
h_2 &= (X-\mu)^2 - \sigma^2.
\end{aligned}$$

But h_1 and h_2 are not orthogonal to each other. We adopt the othogonalization procedure in Doob (1953) to produce an orthogonal estimating function to h_1

$$h_3 = (X - \mu)^2 - \sigma^2 - \gamma_1 \sigma (X - \mu).$$

Then we need to find an optimal linear combination of estimating functions h_1 and h_3 as follows

$$l_\mu = \alpha h_1 + \beta h_3.$$

Godambe and Thompson (1989) shows that the optimal coefficients α and β based on the theory of estimating functions are given as follows

$$\begin{aligned}
\alpha^* &= \frac{E(\frac{\partial h_1}{\partial \mu})}{E(h_1^2)} = -\frac{1}{\sigma^2}, \\
\beta^* &= \frac{E(\frac{\partial h_3}{\partial \mu})}{E(h_3^2)} = \frac{\gamma_1 \sigma}{\sigma^4(\gamma_2 + 2 - \gamma_1^2)}.
\end{aligned}$$

In general, $\frac{l_\mu^*}{\sqrt{Var(l_\mu^*)}}$ can be approximated by a standard normal distribution. So a $(1 - \alpha)$ percent confidence interval for $\frac{l_\mu^*}{\sqrt{Var(l_\mu^*)}}$ would be

$$\left| \frac{l_\mu^*}{\sqrt{Var(l_\mu^*)}} \right| < C_\alpha, \tag{2}$$

where C_α is the critical value corresponding to the confidence level α. For example, if $\alpha = 0.05$, $C_\alpha = 1.96$. From the inequality (2) we can solve for a confidence interval for X if all moments are known, i.e.,

$$X_L < X < X_U.$$

Some tedious mathematical derivations result in the following result

$$X_U = \mu + \frac{\frac{\gamma_2+2}{\gamma_1} + \sqrt{\left(\frac{\gamma_2+2}{\gamma_1}\right)^2 + 4\left[\frac{C_\alpha\sqrt{(\gamma_2+2)(\gamma_2+2-\gamma_1^2)}}{|\gamma_1|} + 1\right]}}{2}\sigma, \quad (3)$$

$$X_L = \mu + \frac{\frac{\gamma_2+2}{\gamma_1} - \sqrt{\left(\frac{\gamma_2+2}{\gamma_1}\right)^2 + 4\left[\frac{C_\alpha\sqrt{(\gamma_2+2)(\gamma_2+2-\gamma_1^2)}}{|\gamma_1|} + 1\right]}}{2}\sigma, \qquad \gamma_1 \neq 0$$

In the case of normal distribution

$$\gamma_1 = \gamma_2 = 0,$$

the optimal estimating function is

$$l_\mu^* = -\frac{X-\mu}{\sigma^2}$$

and

$$\left|\frac{l_\mu^*}{\sqrt{Var(l_\mu^*)}}\right| = \left|\frac{X-\mu}{\sigma}\right|.$$

In this case our approximation approach leads to the same confidence interval constructed under the assumption of normal distribution.

4 The Properties of the Model

Next we study the properties of this model. We also investigate how the length of confidence intervals changes with the moment inputs. The length of the confidence interval can be defined as the difference between the upper limit and lower limit, i.e.

$$L = X_U - X_L.$$

Under the estimating function approach we have

$$L = \sqrt{\left(\frac{\gamma_2 + 2}{\gamma_1}\right)^2 + 4\left[\frac{C_\alpha\sqrt{(\gamma_2+2)(\gamma_2+2-\gamma_1^2)}}{|\gamma_1|} + 1\right]}\sigma, \qquad \gamma_1 \neq 0. \quad (4)$$

In the special case of normal distribution we have

$$L = 2C_\alpha\sigma.$$

We use a numerical example with the following parameters

$$\begin{aligned} \sigma &= 0.0060, \\ \gamma_1 &= -0.2244, \\ \gamma_2 &= 3.1556 \end{aligned}$$

which are the average volatility, skewness and kurtosis of the twelve major currencies we study later on. We can make the following observations

- As in the case of a normal distribution, the length of the confidence interval is positively related to the standard deviation or volatility. This is consistent with our intuition since the standard deviation measures the dispersion of the distribution. The more disperse the distribution, the longer the confidence interval we need for a given confidence level.

- Unlike in the normal case, the confidence interval is not symmetrical around the mean value. It is tilted toward the direction of skewness. If the skewness is positive, the confidence interval covers more value on the right-hand side of the mean value than the left-hand side. If the skewness is negative, the confidence interval covers more on the left hand of the mean value than the left-hand side. Skewness acts as an indicator as to which side the confidence interval should be stretched so that a given percentage of the underlying distribution is covered.

- The skewness parameter γ_1 characterizes the degree of asymmetry of the distribution around its mean. Positive (negative) skews indicates asymmetric tail extending toward right-hand (left-hand) side. Intuitively, if the skewness is very large or small we know which side the asymmetrical fat tail extends toward, so the length of confidence interval should be narrow. But when the skewness is very small, but not zero, we know the distribution is not symmetrical, but are not sure which side the fatter tail is. In this case we need a longer confidence interval to cover a certain percentage of the density function. Based on the formula (4) we find that the length of the confidence interval is inversely related to the absolute value of the skewness. Figure 1 shows this property. The length of confidence intervals is a bell shape type function of the value of its skewness with the minimum attained at zero skewness. There is a discontinuous point at zero skewness since the confidence interval in the case of zero skewness is different than that in the case of non-zero skewness.

- The standardized kurtosis measures the relative peakedness or flatness of a given distribution compared to a normal distribution. High kurtosis or leptokurtosis indicates there are more occurrences far away from the mean than predicted by a standard normal distribution. From equation (4) we see that L is positively related with the excess kurtosis. Figure 2 also shows the positive relationship.

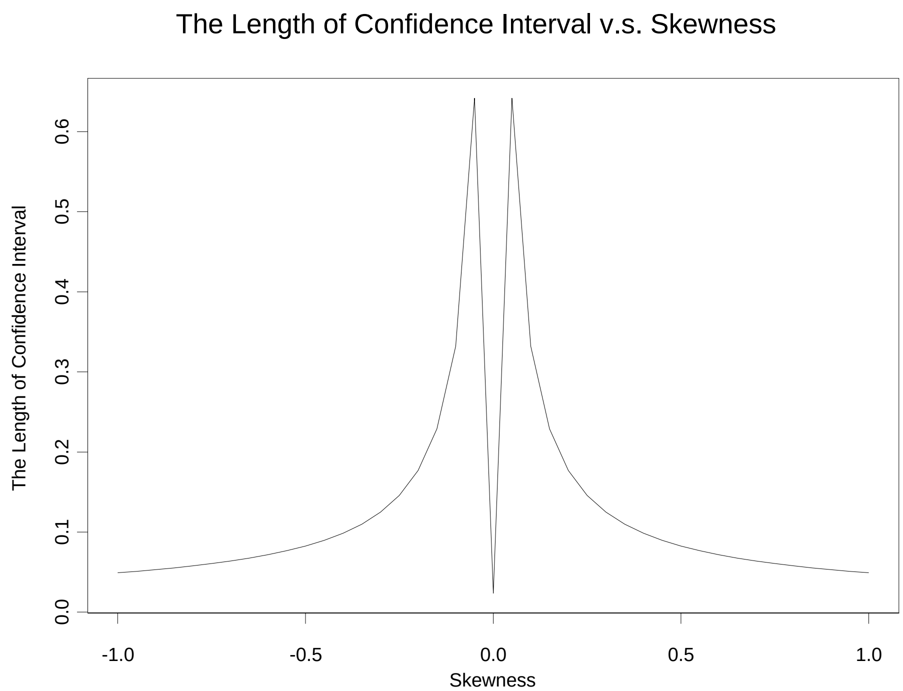
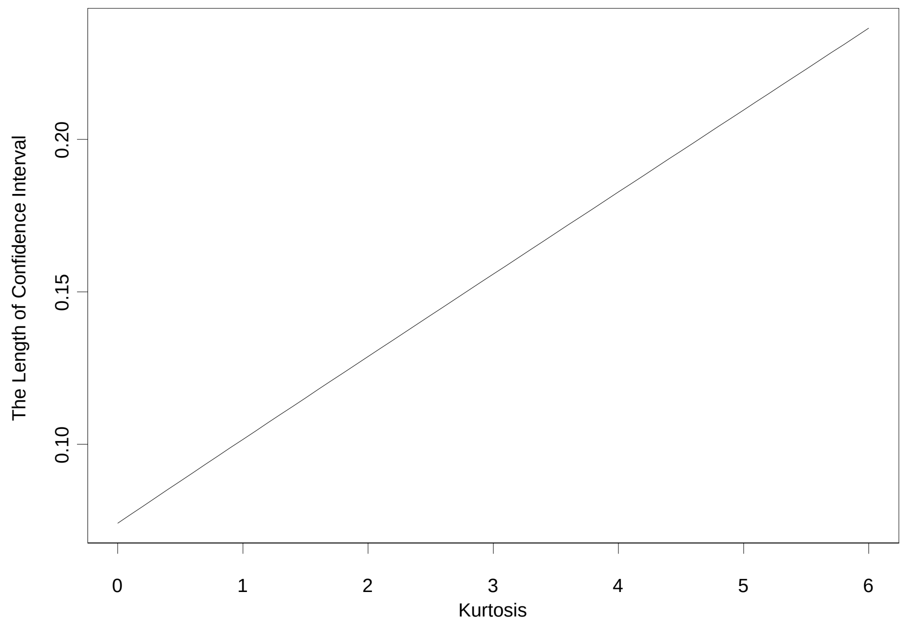

The Length of Confidence Interval v.s. Skewness
The Length of Confidence Interval
0.0
0.1
0.2
0.3
0.4
0.5
0.6
-1.0
-0.5
0.0
0.5
1.0
Skewness

Figure 2. The Length of Confidence Interval v.s. Kurtosis



5 An Empirical Study

Next we backtest our model to see how well it works in practice. To backtest our model we use daily exchange rates for twelve major currencies between February 17, 1989 and February 8, 1999. The total number of trading days covered by the data is 2584. We first calculate the daily logarithm change $X_t = \ln(S_t/S_{t-1})$, where S_t is the spot exchange rate at time t. Then, as in Hull and White (1998), we use two approaches to estimate the volatility of the series:

- We estimate one volatility σ using the entire return series for each currency. We simply call this approach the constant-variance model.

Table 1: Skewness and Kurtosis for the Return x_t

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NLG	SEK
Constant Volatility Model												
γ_1	-0.17	0.03	0.16	0.04	-0.08	-0.41	-0.08	-0.21	-0.74	0.66	0.09	-0.87
γ_2	6.04	2.07	1.75	2.19	2.44	5.34	1.88	3.02	7.05	5.58	2.35	11.15
RiskMetrics EWMA Model												
γ_1	-0.54	0.18	0.34	0.17	0.13	-0.03	0.24	0.05	0.07	0.58	0.21	-0.28
γ_2	3.14	2.11	2.33	1.88	2.27	2.67	2.03	2.40	2.10	3.24	2.01	3.29

- We use an exponentially weighted moving average (EWMA) with a smoothing parameter $\lambda = 0.94$ and 74 past observations to estimate the time-varying volatility σ_t. This is the standard RiskMetrics^{TM} approach. We call this approach the RiskMetrics^{TM} EWMA method.

Then we obtain a new time series data by dividing the return series by the standard deviation. In the constant-variance approach $x_t = \frac{X_t}{\sigma}$, and in the RiskMetrics^{TM} EWMA approach $x_t = \frac{X_t}{\sigma_t}$. We calculate the sample moments based on the series X_t and x_t. Table 1 shows the skewness and kurtosis for the transformed series x_t, of each currency.

From Table 1 we have the following observations for the return series when the constant variance model is used.

- Most currencies have a non-zero skewness. For a normal distribution the skewness is zero.

- All currencies exhibits significant excess kurtosis. The kurtosis for the twelve currencies varies from 1.75 to 11.15. The excess kurtosis is zero for a normal distribution.

From the table we see the following properties for the transformed data using the RiskMetrics^{TM} EWMA method.

- Each transformed return series still has a non-zero skewness and kurtosis. If the RiskMetrics^{TM} assumption is accepted, the series x_t should be iid standard normally distributed

- Both the skewness and kurtosis are generally reduced after transformation using the RiskMetrics^{TM} EWMA method., but the skewness of

Table 2: Backtesting Result when Constant Volatility Model is Used

AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NLG	SEK
$\alpha = 5\%$											
5.57	5.41	5.65	6.00	5.77	5.69	6.11	5.96	5.38	6.04	5.84	5.38
2.74	2.83	2.94	2.83	3.21	3.79	2.94	3.02	3.67	3.29	2.90	3.60
$\alpha = 3\%$											
4.02	4.33	4.14	4.22	3.91	3.64	4.37	4.49	3.87	4.06	4.10	3.91
2.09	2.17	2.17	2.01	2.32	2.86	2.44	2.44	3.13	2.63	2.13	2.79
$\alpha = 1\%$											
2.23	2.83	2.24	2.09	2.24	2.48	2.24	2.79	2.09	2.48	2.09	2.36
0.97	1.63	1.32	1.16	1.39	1.55	1.16	1.51	1.74	1.51	1.28	1.59

some currencies and the kurtosis of all currencies are still significantly different from zero.

These observations show that the conditional normality assumption by RiskMetrics^{TM} is not consistent with our empirical findings. The majority of the transformed return series are still skewed and leptokurtic, in contrast to the normal distribution which we assume in the usual RiskMetrics^{TM} VaR calculation. It also shows that it is important to incorporate the skewness and kurtosis into the VaR calculation.

To backtest our approach we first construct a two-tailed confidence interval based on the volatility alone under the normal assumption. We also construct a confidence interval based on equation (3). Finally we calculate the actual percentage of observations which fall outside the confidence interval and compare it with the significance level. We choose three confidence levels $\alpha = 0.05, 0.03, 0.01$. To study the effect of the transformation or degarching, the same procedure has applied to the transformed series x_t using two transformation methods. The result is summarized in Table 2 and Table 3.

For each significance level the first row gives the percentage of the number of observations outside the confidence interval based on the normal assumption, and the second one based on the estimating function approach.

From Table 2 and 3 we see that the normal assumption approach always underestimates the number of observations outside the confidence interval. For 1 percent significance level, there are twice as many observations out-

Table 3: Backtesting Result When EWMA Model is Used

AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NLG	SEK
$\alpha = 5\%$											
5.26	5.06	5.46	5.70	5.70	5.22	5.86	5.82	5.74	5.38	5.86	5.42
3.43	2.99	3.15	3.19	2.95	2.75	3.15	2.95	2.83	3.83	3.27	3.71
$\alpha = 3\%$											
4.07	3.51	3.87	4.07	4.15	3.83	4.03	4.34	4.39	3.75	3.91	3.79
2.79	2.27	2.59	2.43	2.11	1.87	2.35	2.23	2.11	2.67	2.35	2.55
$\alpha = 1\%$											
2.27	1.87	2.15	2.03	2.11	2.07	1.83	2.31	2.11	2.19	2.03	1.87
1.59	1.24	1.43	1.24	1.28	1.08	1.47	1.28	1.20	1.63	1.20	1.20

side the confidence interval band than are predicted by a confidence interval based on a normal assumption. In contrast the estimating function approach overestimates for both the significance levels of 0.05 and 0.03, and slightly underestimates for the significance level of 0.01. This suggests that our new approach does capture the extreme cases better than the normal assumption method.

Comparing Table 2 and Table 3 we also observe the effect of de-ARCHing of the RiskMetrics^{TM} EWMA method. It shows that the degarching procedure does improve the accuracy of the RiskMetrics^{TM} VaR calculation based on our backtesting result. However, it has much smaller impact on our estimating function method. It shows that our non-linear transformation involving skewness and kurtosis as well as volatility produces a series which is closer to normal than the changing volatility approach used in RiskMetrics EWMA method.

6 Conclusion

We have shown how to incorporate the skewness and kurtosis explicitly into the construction of a confidence interval based on the theory of estimating functions in statistics. The final result of the confidence interval is an explicit function of the skewness and kurtosis as well as the standard deviation or volatility. The length of the confidence interval is positively related to the kurtosis and negatively related to the absolute value of skewness, consistent with our intuition on the relationship between the confidence interval and the skewness and kurtosis. The new approach enables us to directly take account of empirical findings on most financial time series data. It is a semiparametric approach where only moments up to fourth order need to be empirically estimated. No full distribution assumption is required.

We back test the model using 10-year of daily exchange rate data for twelve major currencies, and find that it is able to capture the extreme situation much better than the normal assumption approach. This model can be adapted to calculating VaR relatively easily under the general framework of the RiskMetrics^{TM}. A Taylor series expansion or a Monte Carlo simulation approach can also be adopted since moments can be readily obtained in these methods. Then we can use the obtained standard deviation, skewness and kurtosis to construct an approximate confidence interval.

References

[1] Doob, J. L. (1953), *Stochastic Processes*, New York: John Wiley and Sons.

[2] Godambe, V. P. (1991), *Estimating Functions*, Oxford: Oxford University Press.

[3] Godambe. V. P. and M. Thompson, (1989), An Extension of Quasi-Likelihood Estimation (with discussion), *Journal of Statistical Planning and Inference*, 22, pp. 137-72.

[4] Hull, J. and A. White (1998), Value at Risk When Daily Changes in Market Variables are not Normally Distributed, *The Journal of Derivatives*, Spring 1998, pp. 9-19.

[5] Kupiec, P. (1995), Techniques for Verifying the Accuracy of Risk Measurement Models, *The Journal of Derivatives*, Winter 1995, pp. 73-84.

[6] Li, D. X. and H. Turtle (1997), Semiparametric ARCH Models: An Estimating Function Approach to ARCH Model Estimation, Working Paper 97-6, Global Analytics, Risk Management Division, CIBC/Wood Gundy.

[7] RiskMetrics - Technical Document, J. P. Morgan, April 1997.

[8] Zangari, P. (1996), "An Improved Methodology for Measuring VaR." RiskMetrics Monitor, Reuters/J. P. Morgan.