EFFECT OF HIGHER ORDER MOMENTS ON HEDGING LOSS VALUE-AT-RISK AND CONDITIONAL VALUE-AT-RISK

Yuji Yamada
Graduate School of Business Sciences
University of Tsukuba
Tokyo 112-0012, Japan
email: yuji@gssm.otsuka.tsukuba.ac.jp

James A. Primbs
Management Science and Engineering
Stanford University
Stanford, CA 94305-4026
email: japrimbs@stanford.edu

ABSTRACT
In this paper, we present a Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) estimation technique for dynamic hedging and investigate the effect of higher order moments in the underlying on the hedging loss distributions and kurtosis using a general parameterization of multinomial lattices, and solve the mean square optimal hedging problem. At first, we approximate the underlying stock process through its first four moments including skewness and kurtosis using a general parameterization of multinomial lattices given in [16]. Suppose that we first model the underlying asset using its (higher order) moments, and combine multinomial tree recombining higher order moments. Then a multinomial tree with possible prices at time \( t_n \) is the mean of i.i.d. random variables whose moments are finite and \( \tau \) is a basic time period. Note that equation (1) can be rewritten in log-coordinates as

\[
\ln S_t - \ln S_{t-1} = X_t.
\]

2 Modeling the underlying asset process using higher order moments

Here we introduce a derivative security market whose underlying is given by a stock in the time interval \( t \in [0, t_N] \). Let \( S_t \) be the price of the stock at \( t = t_n := n\tau, \ n = 0, 1, \ldots, N \) satisfying

\[
S_n = S_{n-1}e^{X_n}, \ n = 1, \ldots, N, \tag{1}
\]

where \( X_n \) is a sequence of i.i.d. random variables whose moments are finite and \( \tau \) is a basic time period. Note that equation (1) can be rewritten in log-coordinates as

\[
\ln S_n - \ln S_{n-1} = X_n.
\]

2.1 Multinomial lattices
We first summarize a parameterization of multinomial lattices given in [16]. Suppose that \( u_n \) and \( d_n \) satisfy \( u_n > d_n > 0 \). Then a multinomial tree with \( L \) branches at each node is given by

\[
S_{n+1} = u_n^{L-1}d_n^{L-1}S_n, \ n = 1, \ldots, L, \tag{2}
\]

where \( p_l, l = 1, \ldots, L \) are the corresponding probabilities which satisfy \( p_1 + \cdots + p_L = 1 \). To make the multinomial tree recombine, we further assume that \( u_n/d_n = c \) for all \( n = 0, \ldots, N - 1 \) for some constant \( c (> 1) \). One can verify that the process in (2) consists of a lattice (or a recombining multinomial tree), where the stock may achieve \( n(L-1) + 1 \) possible prices at time \( t = t_n, n = 0, \ldots, N \). For example, in the case of \( u_n = u \) and \( d_n = d \) for all \( n = 1, \ldots, N - 1, \) the price of the stock at the \( k \)-th node from the top of the lattice is given by

\[
S_n^{(k)} = u^{(L-1)+k}d^{L-1}S_0, \ k = 1, \ldots, n(L-1) + 1. \tag{3}
\]

We will provide a parameterization of multinomial lattice random walks which takes higher order moments into account. Suppose that \( \nu_n\tau \) is the mean of \( X_n \), i.e.,

\[
\nu_n\tau = \mathbb{E}(X_n), \tag{4}
\]
and let
\[
\begin{align*}
    u_n &:= \exp \left( \frac{\nu_n \tau}{L-1} + \alpha \sqrt{\tau} \right), \\
    d_n &:= \exp \left( \frac{\nu_n \tau}{L-1} - \alpha \sqrt{\tau} \right),
\end{align*}
\]
where \( \alpha > 0 \) is some constant. One can readily see that \( u_n/d_n \) is constant for all \( n = 0, \ldots, N - 1 \) if \( \alpha \) is fixed. With these choices for \( u_n \) and \( d_n \), \( X_n \) may be computed as
\[
X_n = \ln S_{n+1} - \ln S_n = \nu_n \tau + (L - 2l + 1) \alpha \sqrt{\tau}.
\]
In this case, the probabilities are unsymmetric if skewness is not zero (\( \kappa_n \neq 0 \), i.e.,
\[
- \text{positive (negative) skewness causes } p_1 \text{ and } p_4 \text{ to increase (decrease)},
\]
and
\[
- \text{the corresponding probabilities } p_2 \text{ and } p_3 \text{ to decrease (increase) by an equal amount.}
\]
Moreover, it is possible to show that there exists \( \alpha \) such that all the probabilities are positive if
\[
\kappa_n > 8 \kappa_n^2 - 3.
\]

2.2 The Kolmogorov backward equation with higher order moments

Here we use a continuous time notation to introduce Markov property and the propagator moment function. The reader may assume that \( X_n = X(t_n) \) and \( S_n = S(t_n) \) for \( n = 0, 1, \ldots, N \).

Let \( X(t), \ t \in [0, T] \) be a Markov process, whose density (in the \( \xi_j \) variable of) of \( X(t) \) conditioned on \( X(t_i) = \xi_i \) is given by
\[
p(\xi_j; \ t_j; \ \xi_i, \ t_i), \quad 0 \leq t_i < t_j \leq T.
\]
Define the \( k \)-th propagator moment function of the Markov process \( X(t) \) as
\[
\Phi^{(k)}(\xi_i; \ t_i), \quad \xi_i = X(t_i), \ t_i \in [0, T]
\]
Informally, the \( k \)-th propagator moment of \( X(t) \) is given by the \( k \)-th order conditional moment of the instantaneous return \( X(t + dt) - X(t) \) divided by the infinitesimal time \( dt \) [6]. Then we have the backward Kramers-Moyal equation:
\[
-\frac{\partial}{\partial t_i} p(\xi_j; \ t_j; \ \xi_i, \ t_i) = \sum_{k=1}^{\infty} \frac{1}{k!} \phi^{(k)}(\xi_i; \ t_i) \frac{\partial^k}{\partial \xi_i^k} p(\xi_j; \ t_j; \ \xi_i, \ t_i).
\]
For simplicity, we consider the case where each propagator moment is constant, i.e., \( \Phi^{(k)}(\xi, \ t) = \phi_k \). We will show that \( \phi_k \) is given by the \( k \)-th order cumulant of \( X(t) \) using the following general relation between moments and cumulants:
\[
m_k = \sum_{j=1}^{k} \frac{k!}{j! (k-j)!} \phi_{j} m_{k-j}.
\]
where \( m_k \) and \( e_k \) are, respectively, the \( k \)-th order moment and the \( k \)-th order cumulant of a random variable.

Assume that \( \Delta(0) = 1, T = 1 \) without loss of generality and that each increment of \( X(t) \) is independent. Also, let \( \hat{e}_k \) be the annualized cumulant of \( X(t) \). Because cumulants have the additive property for independent random variables, the cumulant of \( X(t_{n+1}) - X(t_n) \) is given by \( \tau \hat{e}_k \), where \( t_n = \tau \times n, n = 0, \ldots, N \) and \( \tau = 1/N \). From (13), the \( k \)-th order moment of \( X(t_{n+1}) - X(t_n) \) may be computed as

\[
\tau \hat{e}_k + \text{“higher order terms in } \tau \text{”} \quad (14)
\]

Since the propagator moment is given by the \( k \)-th order conditional moment of the instantaneous return \( X(t + dt) - X(t) \) divided by the infinitesimal time \( dt \), it is obtained by dividing (14) by \( \tau \) and letting \( \tau \to 0 \). Therefore, we have \( \phi_k = \hat{e}_k \), and the backward Kramers-Moyal equation (12) is finally written as

\[
-\frac{\partial}{\partial t_i} p(\xi_j, t_j; \xi_i, t_i) = \sum_{k=1}^{\infty} \hat{e}_k k! \frac{\partial^k}{\partial \xi_i^k} p(\xi_j, t_j; \xi_i, t_i). \quad (15)
\]

**Remark 2.1** Equation (15) addresses the case where the underlying asset follows a geometric Brownian motion, or its log return process \( X(t) = \log \frac{S(t)}{S(0)} \) is given as

\[
dX(t) = \left( r - \frac{\sigma^2}{2} \right) dt + \sigma dB(t) \quad (16)
\]

where \( B(t) \) is a Brownian motion under the risk neutral probability measure. Condition (16) implies that \( X(T) \) is normal with mean \( \left( r - \frac{\sigma^2}{2} \right) T \) and variance \( \sigma^2 T \), and thus, the annualized cumulants are given by

\[
\hat{e}_1 = r - \frac{\sigma^2}{2}, \quad \hat{e}_2 = \sigma, \quad \hat{e}_k = 0, \quad k = 3, 4, \ldots
\]

Then, we have the Black-Scholes equation for a European contingent claim in log-coordinate as follows:

\[
ru(\xi, t) = \frac{\partial}{\partial \xi} u(\xi, t) + \left( r - \frac{\sigma^2}{2} \right) \frac{\partial}{\partial \xi} u(\xi, t) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial \xi^2} u(\xi, t), \quad u(\xi, T) = h(\xi).
\]

**2.3 Jump-diffusion models**

Finally, we will introduce a jump-diffusion model (see e.g., [11]) in discrete time as

\[
S_n = S_{n-1} e^{X_n}, \quad n = 1, \ldots, N, \quad (17)
\]

where \( X_n \) consists of a Gaussian random variable \( G_n \) and a jump portion \( J_n Q_n \), i.e.,

\[
X_n = G_n + J_n Q_n, \quad n = 0, \ldots, N. \quad (18)
\]

Here \( G_n, J_n \) and \( Q_n \) are independent, and

\[
G_n \sim \mathcal{N}(\nu_G \tau, \sigma_G^2 \tau) \quad (19)
\]

\( Q_n \) is a random variable which becomes either 1 with probability \( \lambda \tau \) or 0 with probability \( 1 - \lambda \tau \) for some \( \lambda > 0 \). \( J_n \) determines the jump size which we assume follows another Gaussian distribution

\[
J_n \sim \mathcal{N}(\nu_J, \sigma_J^2) \quad (20)
\]

Mean \( \nu_T \), variance \( \sigma_T^2 \), skewness \( s \), and kurtosis \( \kappa \) of \( X_n \) are computed as

\[
\nu_T = \nu_G + \lambda \nu J, \quad \sigma_T^2 = \sigma_G^2 + \lambda \tau \sigma_J^2 s = \frac{\lambda}{\nu_T^2} E\left[J_T^2\right], \quad \kappa = \frac{\lambda}{\nu_T^2} E\left[J_T^4\right]. \quad (21)
\]

**3 Estimation of VaR and CVaR for hedging loss distributions**

In this section, we estimate VaR and CVaR for hedging loss distributions resulting from mean square optimal hedging, and investigate the effect of higher order moments (i.e., skewness and kurtosis) in the underlying.

**3.1 Mean square optimal price**

Assume that traders are allowed to purchase and sell two basic securities, a risky asset (or stock) as given by (1) and a risk free asset (or bond) with fixed interest rate \( r \) at discrete times \( t_n = nt, \quad n = 0, 1, \ldots, N \). We will consider a self-financing portfolio which consists of the assets in the underlying market. The portfolio value \( \Omega_n (n = 0 \ldots N) \) satisfies the following difference equation:

\[
\Omega_{n+1} = \Delta_n S_{n+1} + R \left( \Omega_n - \Delta_n S_n \right)
= R \Omega_n + \Delta_n (S_{n+1} - R S_n), \quad (22)
\]

where \( \Delta_n \) is the number of shares of the stock held from \( n \) to \( n+1 \), and \( R := 1 + r \).

The objective of mean square optimal hedging (MSOH) is to optimally replicate or hedge the payoff of a European derivative security \( Y_N \) of maturity \( N \) through a self-financing trading strategy with an adequate initial portfolio value \( \Omega_0 \). This involves solving the following optimization problem:

\[
\text{MSOH: } \Omega_0 \Delta_0 \ldots \Delta_{N-1} \min \mathbb{E} \left[ (Y_N - \Omega_N)^2 \right] \quad (23)
\]

subject to the dynamics of the underlying stock and the portfolio, i.e., (1) and (22).

This problem has been studied extensively [3, 4, 5, 7, 14, 15] and can be solved using dynamic programming. Although we only formulate the MSOH problem for a European call option, note that the same approach can be extended to other types of options, including exotics (such as
3.2 Hedging loss VaR and CVaR

The hedging loss distribution is the distribution of

\[ W_N = -(\Omega_N - V_N), \]

i.e., the loss caused by the hedging error, and its VaR is a maximum allowable loss (or a percentile point) at a specified confidence level. In this section, we estimate the hedging loss distribution and the VaR through their moments as follows:

1. First, compute the first four moments of the hedging loss distribution by our recently developed algorithm of [13] using a backward recursion on the underlying stock lattice, under a mean square optimal dynamic hedging strategy (or more generally the so-called “affine hedging strategy”).

2. Plug the moment information into the generalized lambda distribution to extract the hedging loss distribution, and estimate its VaR using the Cornish-Fisher expansion (see, e.g., [9] and [10]).

Given the first four moments of the hedging loss distribution, \( m_1, m_2, m_3 \) and \( m_4 \), mean \( \mu_w \), variance \( \sigma_w^2 \), skewness \( \kappa_w \), and kurtosis \( \kappa_w \) are, respectively, calculated as

\[
\begin{align*}
\mu_w &= m_1 \\
\sigma_w^2 &= m_2 - m_1^2 \\
s_w^2 &= \frac{m_3 - 3m_2m_1 + 2m_1^3}{\sigma_w^2} \\
\kappa_w &= \frac{m_4 - 4m_2m_1 + 12m_2m_1^2 - 6m_1^4}{\sigma_w^2} 
\end{align*}
\]  

Then, the Cornish-Fisher expansion estimates the \( \alpha \)-th percentile of the hedging loss distribution as

\[
\mu_w + \sigma_w Y_\alpha 
\]

where

\[
Y_\alpha = Z_\alpha + \frac{1}{6} (Z_\alpha^2 - 1) s_w + \frac{1}{24} (Z_\alpha^3 - 3Z_\alpha) \kappa_w - \frac{1}{36} (2Z_\alpha^3 - 5Z_\alpha) s_w 
\]

and \( Z_\alpha \) is the \( \alpha \)-th percentile of the standard normal distribution \( N(0, 1) \).

Let \( f : \mathbb{R} \to \mathbb{R} \) be a real valued function such that

\[
f(x) = \mu_w + \sigma_w \left( x + \frac{1}{6} (x^2 - 1) s_w + \frac{1}{24} (x^3 - 3x) \kappa_w - \frac{1}{36} (2x^3 - 5x) s_w \right). 
\]

Then, using the Cornish-Fisher expansion, \( 100 \times \alpha \%-\text{VaR}, \text{VaR}_\alpha \), of the hedging loss distribution is estimated as follows:

\[
\text{VaR}_\alpha \approx f(Z_\alpha). 
\]

Let \( \text{CVaR}_\alpha \) be \( 100 \times \alpha \%-\text{CVaR} \) of the hedging loss distribution. Then it holds that

\[
\text{CVaR}_\alpha = \frac{1}{1 - \alpha} \int_0^1 \text{VaR}_\alpha dq. 
\]

Similar to the VaR case, \( \text{CVaR} \) of the hedging loss distribution can also be estimated using the Cornish-Fisher expansion as follows:

\[
\begin{align*}
\text{CVaR}_\alpha &\approx \frac{1}{1 - \alpha} \int_0^1 f(Z_\alpha) dq \\
&= \mu_w + \sigma_w \left( M^{(1)}_\alpha + \frac{1}{6} (M^{(2)}_\alpha - 1) s_w + \frac{1}{24} \times \right. \\
&\left. \left( (M^{(3)}_\alpha - 3M^{(1)}_\alpha) \kappa_w - \frac{1}{36} (2M^{(3)}_\alpha - 5M^{(1)}_\alpha) s_w \right) \right) \\
M^{(i)}_\alpha &= \frac{1}{1 - \alpha} \int_{Z_\alpha}^\infty x^n \phi(x) dx. 
\end{align*}
\]

where \( Z_\alpha \) and \( \phi(\cdot) \) are the \( \alpha \)-th percentile and the probability density function of the standard normal distribution, respectively.

Note that we can use the generalized lambda distribution to estimate VaR and CVaR instead of the Cornish-Fisher expansion in the above steps. Although we merely provide the results of our numerical experiments for the VaR estimation in this section, more numerical results for the CVaR estimation are in preparation.

3.3 Numerical experiments

Consider a discrete time market with \( t_N = 12 \) weeks, where each time step is given by 1 week. Assume that the
underlying stock price process is modeled as in (1), and that the initial stock value \( S_0 \), the annualized risk free rate \( r \), and the annualized mean and standard deviation of \( X_n \) are given as \( S_0 = 100 \), \( \sigma = 0.15 \) and \( \nu_j = 0.16 \).

We first solve the problem of hedging three European call options on the stock, all with expiration at \( t_N \), and with strike prices \( K = 92, 102, 115 \) which correspond to the option being in-the-money (ITM), at-the-money (ATM), and out-of-the-money (OTM), respectively. In addition, we assume that skewness and excess kurtosis of \( X_n \) are zero, implying that the underlying distribution of \( X_n \) is nearly Gaussian. We computed up to the 4th moment of the hedging loss using the algorithm in [13], and obtained the hedging loss distributions (solid line) in Fig. 1 for the three cases, \( K = 92 \) (ITM), \( K = 102 \) (ATM), and \( K = 115 \) (OTM), where Gaussian distributions with the same mean and variance are also plotted. From this figure, we see that the hedging loss distributions tend to be leptokurtic (or heavy tailed) compared to the Gaussian distributions, particularly in the ITM and OTM cases.

Next, we estimate VaR of the hedging loss distribution by applying the Cornish-Fisher expansion. In this example, the strike price is varied in the range \([95, 110]\), while the initial stock price was held constant at \( K = 10 \). The dashed lines in Fig. 2 denote the relation between 95% VaR and different values of strike prices when the excess kurtosis \( \kappa \) of the underlying stock in log-coordinates is given by \( \kappa = 0, 1, 2, 3 \). In every case, 95% VaR took its maximum around \( K = 105 \), i.e., 95% VaR increases with \( K \) when \( K \leq 105 \) and decreases when \( K \geq 105 \), and the larger kurtosis provided the larger hedging loss VaR.

To compare VaR of the hedging loss distribution with the initial option value, we also plot the mean square optimal prices of options with respect to different strike prices for each kurtosis value as in Fig. 3. Since the value of a European call option decreases with larger strike price in general, the gap between the initial option value and the estimated value of 95% VaR for the hedging loss distribution becomes smaller. In particular, in the case of \( \kappa = 3 \), the estimated value of 95% VaR is bigger than the initial option price if \( K \geq 105 \), implying that there is at least 5% possibility of losing more than the initial option value. From these numerical experiments, we conclude that large kurtosis of the underlying provides a significant effect on hedging loss.

Finally, we compare delta hedging (see e.g., [1]) with mean square optimal hedging in terms of VaR estimated from Cornish-Fisher expansion. Fig. 4 shows the relation between VaR from a Cornish-Fisher expansion and kurtosis for two hedging schemes, delta-hedging and mean square optimal hedging. In this example, we varied the value of kurtosis for different values of strikes. The solid lines denote the mean square optimal hedge, whereas the dashed

![Fig. 1. Hedging loss distributions](image1)

![Fig. 2. VaR by Cornish-Fisher vs. Strike price](image2)

![Fig. 3. VaR (and option value) vs. Strike price](image3)
ones denote delta-hedging. In every case, the estimated VaR of both delta-hedging and mean square optimal hedging monotonically increases with kurtosis of the underlying. Although the difference between the two is not so significant, we see that mean square optimal hedging provides a better hedge in terms of VaR for the hedging loss distribution in this example.

![Fig. 4. VaR vs. Kurtosis for $K = 95$, 100, 105, 110](image)

### 4 Conclusion

In this paper, we demonstrated a Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) estimation technique for dynamic hedging and investigate the effect of higher order moments in the underlying on the hedging loss distributions. At first, we approximated the underlying stock process through its first four moments including skewness and kurtosis using a general parameterization of multinomial lattices, and solved the mean square optimal hedging problem. Then our recently developed technique was applied to extract the hedging loss distributions in option hedge positions. Finally, we showed how the hedging error distribution changes with respect to non-zero kurtosis and skewness in the underlying through numerical experiments, and examined the relation between VaR and CVaR of the hedging loss distributions and kurtosis of the underlying.

### References


