VITAL STATISTICS

Eric Jacquier and Robert Jarrow present a new framework for evaluating error in option pricing models

Like all models, option pricing models are, by definition, simplifications of reality. As such, they are misspecified and contain error.

The better the model, the smaller the error. In the best, the time series of the errors behave as independent and identically distributed random variables which are not correlated with any relevant economic variables. Unfortunately, few option pricing models are in this category. For instance, the Black-Scholes model contains systematic biases when compared with market prices (see Rubinstein, 1985, and Whaley, 1982).

Textbook treatments of option pricing theory ignore model error when discussing risk management (see Cox and Rubinstein, 1985, Jarrow and Rudd, 1983, and Hull, 1989). Yet model error is of immense importance in the application of option pricing and related risk management techniques. It can be argued, for example, that it is the only reason why both gamma and vega hedging exist.

This article describes a new methodology for evaluating model error in the context of option pricing models. It uses standard Bayesian statistical procedures augmented by new Monte Carlo simulation techniques.

Defining the model error

The crucial step in analysing model error is quantifying its existence. We do this in the context of option pricing models (and, more generally, derivatives security valuation models), through two related equations.

The first relates the market price for an option at time $t$, $C_t$, to the “true” arbitrage-free value of the option, $C^*$:

$$ C_t = c_t \times e^{r_t} $$

(1)

The quantity $e_t$ is called the market error because it represents an arbitrage opportunity. Standard option pricing models set the market error at zero, $e_t = 0$. In contrast, equation (1) explicitly recognizes its existence. Both $c_t$ and $e_t$ are unobservable quantities.

If we let $P(\cdot)$ denote the empirical (objective) probabilities associated with this system and $E_t(\cdot)$ its corresponding time $t$ expectation operator, then a market which is on average properly priced would satisfy:

$$ C_t = E_t(c_t \times e^{r_t}) $$

The second equation relates the “true” arbitrage-free price at time $t$, $c_t$, to the option pricing model’s value, $m_t(x_t, \theta)$:

$$ c_t = m_t(x_t, \theta) $$

(2)

The quantity $\eta_t$ represents model error as it is the difference (in logarithms) between the “true” arbitrage-free price and the model’s value.

The model’s value is postulated as a function of observable variables $(x_t)$ and unobservable parameters $(\theta)$. Both $x_t$ and $\theta$ can be vectors. For example, $m_t(x_t, \theta)$ could be the Black-Scholes model for equity options. The observable variables $x_t$ would consist of the stock price; the spot rate of interest; the dividends; the exercise price; and the option’s expiry date. The unobservable parameter $\theta$ would be the stock’s volatility. Both $\theta$ and $\eta_t$ are unobservable quantities.

An unbiased model would satisfy:

$$ c_t = E_t(m_t(x_t, \theta) \times e^{r_t}) $$

Combining equations (1) and (2) we see that:

$$ C_t = m_t(x_t, \theta) \times e^{r_t + \eta_t} $$

(3)

The error between the market price ($C_t$) and the model’s value ($m_t(x_t, \theta)$) can be decomposed into two components, the market error ($e_t$) and the model error ($\eta_t$). But in both practice and theory, it is difficult to separate the two.

Statistical model for the model error

To analyse statistically the error terms in equation (3), we need to postulate a statistical model for both the market and model errors. Let us assume that there is no market error ($e_t = 0$) and that the model error ($\eta_t$) is a sequence of independent and identically distributed normal random variables with mean 0 and variance $\sigma^2$.

Both assumptions are easily relaxed and their modifications are discussed in the final section of this article.

Purpose of the error analysis

Given the quantification in the previous sections, the purpose of the error analysis is to generate posterior distributions for the unobservables (model error $\eta_t$, parameters $\theta$) and functions of these unobservables, to answer the following three fundamental questions:

1. Mispricing error Given a market quote $C_{t+1}$ and an estimate of the model’s price, how likely is it that a difference represents an arbitrage opportunity?

2. Hedging error Given the current market quote $C_t$ and the option’s delta, what is a distribution for the next period’s dollar hedging error? For example, what is the likelihood of losing more than A dollars?

3. Model comparison Given a collection of different models ($m^1_t(x_t, \theta_1), m^2_t(x_t, \theta_2), \ldots$), eg. Black-Scholes versus stochastic volatility models, which one matches past market prices (or hedging errors) the best?

These questions are fundamental to risk management but standard option pricing techniques do not address them.

Results

The answers can be found using Bayes’ Theorem, if a prior distribution is given for the unobservable parameters $\theta$, denoted $P(\theta)$. In practice, $P(\theta)$ is set to reflect either diffuse priors or some analytically convenient probability distribution. Letting $y_t$ represent the time $t$ history of the observable variables $x_t$, note that $E_t(c_t) = E(\cdot | y_t)$. Bayes’ Theorem yields the key insight:

Equation (4) gives the posterior distribution for the unobservable parameters $\theta$, given the (observable) history $y_t$. It also provides the fundamental equation needed to answer the above questions. Before illustrating this, some remarks need to be made concerning the computation of equation (4).

First, to compute equation (4), an explicit representation of the distribution for $P(y_t | \theta)$ is needed. The distribution for $P(y_t | \theta)$ is an input to the statistical procedure. This distribution is usually provided by the theory underlying the model construction. For example, if $m_t(x_t, \theta)$

1. Note that $\text{log } C_t = \text{log } c_t + e_t$. We use an additive error structure in logarithms to guarantee non-negativity of option prices. If $e_t$ is homoscedastic, then it is the percentage error which is homoscedastic, not the dollar error.

2. This hedging error could easily be extended to cover gamma and vega hedging as well, see “Results”
represents the Black-Scholes formula, then \( P(Y_t | \theta) \) will be given by a lognormal distribution over stock prices \( (Y_t = S_t) \) with drift and volatility parameters \( (\mu, \sigma) = \theta \).

Second, we cannot usually get analytic expressions for the normalisation constants of the posterior densities as in equation (4). On the basis of the kernel only, we can still use simulation-based estimators, that is, we make draws from the posterior density of the parameter. These draws yield draws of the density of any function of interest, eg, hedge ratio, hedging error.

The advantage is that we obtain draws of the exact distribution of these quantities, and do not have to resort to standard approximations. This can be crucial if, for example, we are working in an updating set-up with a small cross-section of option prices. A technical problem arises because it is impossible to make direct draws of the distribution in (4). The Metropolis algorithm helps us get round this problem.

### 1. Mispricing error

To analyse the mispricing error of a yet unobserved call price \( C_t \) we need to compute the predictive density of \( C_t \) given \( Y_t \). By equation (3):

\[
C_t = m_t(x_t, \theta) \times e^{\eta + \xi},
\]

(5)

To compute \( P(C_t | Y_t) \), we use equations (4) and (5) to integrate out parameter uncertainty over \( \theta \), that is:

\[
P(C_t | Y_t) = \int P(m_t(x_t, \theta) e^{\eta + \xi} | \theta, Y_t) P(\theta | Y_t) \, d\theta
\]

(6)

Equation (6) is computable because the distribution for the error \( (\eta + \xi) \) is available from the statistical model postulated earlier. The output from the computation of equation (6) is a plot of the posterior distribution for the call values \( C_t \) like that given in figure 1.

The best estimate of the call’s value is, of course, \( E(C_t | Y_t) \). In figure 1, points A and B represent the call prices for which 5% of the distribution lies below and above those values respectively.

The idea underlying the usefulness of figure 1 is that if the observed call price \( C_{t+1} \) lies outside A and B, the likelihood that the difference between \( C_{t+1} \) and \( E(C_t | Y_t) \) represents model error is small (less than 5%). The difference is more likely attributable to market error, that is, an arbitrage opportunity. This information would signal a trade.

The difference between this probabilistic approach and the standard option pricing technique is easily explained. The standard technique provides a point estimate only for the market price, \( C_t \). Any difference, no matter how small, is attributable to an arbitrage opportunity. In contrast, this Bayesian approach provides a range of possible option prices (A to B) consistent with the past obser-

---

### RISK CONFERENCE AND COURSES

**DOCUMENTATION AND TAPES**

<table>
<thead>
<tr>
<th>RISK CONFERENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>■ ADVANCED MATHEMATICS for Derivatives</td>
</tr>
<tr>
<td>■ TAX ARBITRAGE using derivatives</td>
</tr>
<tr>
<td>■ Forecasting, hedging and trading VOLATILITY</td>
</tr>
<tr>
<td>■ Derivatives for the INSURANCE INDUSTRY</td>
</tr>
<tr>
<td>■ EMERGING MARKET derivatives</td>
</tr>
<tr>
<td>■ Advanced RISK MANAGEMENT TECHNIQUES</td>
</tr>
<tr>
<td>■ YIELD CURVES modelling, hedging and trading</td>
</tr>
<tr>
<td>■ Measuring and managing MARKET RISK</td>
</tr>
<tr>
<td>■ STRESS TESTING PORTFOLIOS for effective risk management</td>
</tr>
<tr>
<td>■ EQUITY DERIVATIVES, pricing, hedging and trading</td>
</tr>
<tr>
<td>■ Forecasting and hedging CORRELATION</td>
</tr>
<tr>
<td>■ Developing, pricing and hedging EXOTIC OPTIONS</td>
</tr>
<tr>
<td>■ RISK MANAGEMENT SYSTEMS for derivatives</td>
</tr>
<tr>
<td>■ OBJECT ORIENTED TECHNOLOGY for derivatives</td>
</tr>
<tr>
<td>■ Pricing, hedging and trading CORRELATION</td>
</tr>
</tbody>
</table>

---

**RISK COURSES**

Price for documentation only

- **US$370** **£250**
- **US$225** **£150**
- **US$485** **£325**
- **£353.95 (UK only)**

Price for audio tape recording only

- **US$225** **£150**
- **US$176.25 (UK only)**
- **£114.95 (UK only)**

Price for audio tape recording and documentation

- **US$675** **£450**
- **£585.75 (UK only)**

*Purchases from EC countries must supply their VAT number, otherwise VAT will be charged at the standard rate VAT/IVA/IBTIVA/Mehrwertsteuer/Meervwntsteuer. C8 511 122 9 07.

For companies in EU member states only. Please write your VAT/IVA/IBTIVA/Mehrwertsteuer/Mehrwertsteuer number.
2. Posterior distribution for $H_t$

$$P(H_t(x_t, \theta) | y_t)$$

**dollars**

of the relevant economic variables and previously observed model errors.

2. Hedging error

A similar computation can be made to analyse hedging error. For illustrative purposes, consider an option on a stock with stock price $S_t$. The delta is the derivative of the option model’s value with respect to the underlying stock price, hence it is also a function of the observables $x_t$ and the parameters $\theta$. Let $\Delta_t(x_t, \theta)$ represent the option’s delta at time $t$. The hedging error, for yet unobservable stock prices ($S_t$) and call values ($C_t$), denoted $H_t(x_t, \theta)$ is defined as:

$$H_t(x_t, \theta) = \left[ C_t - C_{t-1} - \Delta_t(x_t, \theta) (S_t - S_{t-1}) \right] \quad (7)$$

The predictive density for $H_t(x_t, \theta)$ given $y_t$, denoted $P(H_t(x_t, \theta) | y_t)$ can be computed via equation (8):

$$P(H_t(x_t, \theta) | y_t) = \int P(H_t(x_t, \theta) | \theta, y_t) P(\theta, y_t) \, d\theta \quad (8)$$

Equation (8) is computable because:

$$P(H_t(x_t, \theta) | \theta, y_t)$$

is known, given expressions (5), (7) and $P(\theta, y_t)$.

The best estimate of the hedging error is $E(H_t(x_t, \theta) | y_t)$.

The output from this computation would be a plot of the posterior distribution for the hedging error $P(H_t(x_t, \theta) | y_t)$, as shown in figure 2. Let the point $A$ represent the dollar magnitudes for which $5\%$ of the distribution lies below this value. Figure 2 provides the range and likelihood of the possible hedging losses due to a quantification of the model error. Standard option pricing models cannot provide estimates for this quantity.

In practice, if the possibility of a significant loss appears too great, the delta hedge can be augmented to include gamma or vega hedging and the posterior distribution appropriately recalculated. If the augmented hedge still leads to unacceptable probabilities of a large loss, then the trade can be closed or avoided.

3. Model comparison

Given various models, $\{m_1(x_t, \theta), m_2(x_t, \theta), \ldots\}$, this Bayesian approach is well suited to give a statistical comparison. Standard posterior odd ratios can be computed based on equations (3) and (4). These ratios can be used in standard ways to choose among competing models (see DeGroot, 1970). For example, a comparison of Black-Scholes and various stochastic volatility models can be easily implemented in this fashion.

A time series plot of these posterior odd ratios can also give information on trends regarding which model the “market” appears to be currently using to price traded options.

**Extensions**

This article describes a new method of analysing model errors in option pricing models. The advantage of this approach is that it can easily be modified or extended to accommodate different markets and modelling structures. The statistical model used was chosen for simplicity. Two extensions are worth mentioning.

The first is where the market error ($\epsilon_t$) is assumed to be non-zero. We have studied a model where the market error is assumed to occur only rarely and with small probability. Outlier detection procedures can then be used. They yield a probability of occurrence of this additional error for each observation.

The second extension is with respect to the randomness of the model error ($\eta_t$). This approach can easily incorporate known relationships (biases) between model error $\eta_t$ and market observables $x_t$ or unobservable parameters $\theta$, to obtain better-fitting models using existing computer software for the original model. We studied a general class of such dependencies which can provide both better pricing and hedging techniques than standard models. This extension is an “inexpensive” alternative to developing and programming a completely new model.

Eric Jacquier is assistant professor of finance and Robert Jarrow is the Donald P and Susan E Lynch professor of investment management at the Johnson Graduate School of Management, Cornell University.

Robert Jarrow is also director of research at Kamakura Corporation.