國 立 交 通 大 學

資訊管理研究所

博 士 論 文

使 用 時 變 機 率 分 配 之 歐 式 選 擇 權 定 價 模 型 研 究

Using Computational Methodology to Price European Options with Time Variant Distributions

研 究 生： 盛介中

指導教授： 陳安斌 博士

中 華 民 國 96 年 6 月

使 用 時 變 機 率 分 配 之 歐 式 選 擇 權 定 價 模 型 研 究

Using Computational Methodology to Price European Options with Time Variant Distributions

研 究 生： 盛介中　　　　　　　Student: Chieh-Chung Sheng

指導教授： 陳安斌 博士　　　　Advisor: An-Pin Chen

國立交通大學

資訊管理研究所

博士論文

A Dissertation

Submitted to Institute of Information Management

College of Management

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy in Information Management

June 2007

Hsinchu, Taiwan, the Republic of China

中 華 民 國 九 十 六 年 六 月

Acknowledgement

It takes me five years to accomplish my doctoral degree. And, it may be my last diploma of my entire life. I have to firstly thank my advisor Professor An-Pin Chen, he always supports me and gives me a lot of suggestions. Next, Professor Shrane-Koung Chou first taught me how to apply information technology to solve real world problems during my college time. He is always amiable to students and having abundant knowledge among cross field applications. It is also worth mentioned that Professor Han-Lin Li is a well disciplined scholar having strong logic concepts and theoretical backgrounds. I learned a lot from him about how to be a good scholar and how to perform deliberated research.

Apart from my teachers, I have to thank my parents and parents in law who always support me without any reservation. Also, I have to thank my wife Kates Chiu who gives me a lot of supports on research and takes care of my life. It is also worth mentioned that Professor Mu-Yan Chen is one of my best friends who gives me strong supports on discussing research ideas and we exchanges learning experiences with each other. Finally, I thank all my doctoral dissertation committee members. They gave me a lot of suggestions on improving my final dissertation.

Content

Graphs

Tables

Using Computational Methodology to Price European Options with Time Variant Distributions

Abstract. Most option pricing methods use mathematical distributions to approximate underlying asset behavior. However, pure mathematical distribution approaches have difficulty approximating the actual distribution. This study first introduces an innovative computational method for pricing European options based on time variant distributions of the underlying asset. The distribution can be either mathematically generated or simply apply real payoff distributions. Moreover, this computational approach can also be applied to applications related to expected value computation that require actual distributions rather than mathematical distributions. This study makes the following contributions: a) solving the risk neutral issue related to price options with none-mathematical distributions; b) proposing a simple method for adjusting standard deviation based on the need to apply short term volatility to real world applications; c) demonstrating option pricing algorithms that are easy to apply to cross field applications; d) helping traders to generate their own distributions to price new derivative instruments without the difficulty of deriving new closed form formulas.

Keywords: Option pricing, actual distribution, expected value, implied volatility

1. Introduction

An option is a tradable contract that confers the right, but not the obligation, to buy (call) or sell (put) an underlying asset at an agreed-upon price during a certain period or on a specific date. The value of such a contract is termed the option price or option value. Thus, an option price is the expected return of the underlying asset's final settlement price larger (call) or lesser (put) than the desired value (the agreed-upon price). Because option value is the expected return of a usually unpredictable underlying asset, option pricing methodologies have been widely adopted by cross fields applications that need to obtain the target's expected value under uncertainties. For example, the real options analysis (ROA) approach has been widely adopted for assessing information technology investments since the early 1990s [1][2]. Thus, improvements in option pricing methodology can significantly benefit expected value related applications.

Option pricing methods have been widely researched since the development of the Black-Scholes model (BS model) in 1973 [3]. Numerous studies have attempted to relax the restrictive assumptions of the BS model by using various methodologies to approximate the real payoff distribution on assets in a risk-neutral manner and thus obtain the fair option price. Although it seems natural to obtain the option price based on real asset payoff distribution, this idea has rarely been implemented because the actual distribution never behaves risk-neutrally. This characteristic limits the adoption of option pricing methodology in certain non-mathematical distribution applications because real world behavior frequently disobeys

mathematical distributions. Furthermore, the time value decreasing speed of an option accelerates considerably (non-linearly) as the maturity date approaches, yielding large pricing error, but high-frequency (time interval less than 1 minute) pricing methodologies have received little attention. This non-linear variation characteristic also limits high frequency applications. For example, applications with time to maturity less than one day are not suitable for traditional option pricing methodologies because its expected value varies significantly as the settlement time approaches. If an option pricing model can remove the above limitations, it will be more applicable not only in finance but also in cross field applications.

Accordingly, this study proposes a computational model for pricing European options (whose exercise is only permitted on expiry) using time variant distributions of the underlying asset, and verifies the high-frequency pricing performance based on empirical investigation. Experimental results indicate not only that the proposed method can use actual distribution for pricing options which outperforms the BS model, but also that modern computational methods can be adopted to implement option pricing applications rather than using mathematical distributions via closed form formulas. According to the test results, the proposed model contributes significantly to overcoming the limitations of traditional options pricing models when adopted by numerous cross field applications. For example, researchers must determine whether their target index exhibits geometric Brownian motion with lognormal returns when integrating the BS model (or most option pricing models) to calculate the desired expected values, as Benaroch did in his research on IT investment risks [4]. However, there is no need to justify the target's distribution when using the proposed computational model. Moreover, traders can generate their own distributions to emulate the real world behaviour of any underlying asset without the help of mathematicians to derive closed form formulas while it is also possible that no closed form formula existed in some situations.

The rest of this paper is organized as follows. Section 2 briefly discusses the traditional option pricing methodologies. Section 3 then discusses observations of asset real payoff distribution and the feasibility of applying the actual distribution map to price European options. The pricing methodology and algorithms are also presented in this section. Next, section 4 conducts an empirical study to verify effectiveness of applying real payoff distribution to price European options with the proposed methodology. Finally, conclusions and future research directions are presented in section 5.

2. Backgrounds on option pricing

Before constructing option pricing models, it is important to realize basic concepts about option and its fair price, option pricing models, and risk-neutral distributions.

2.1 Options and fair prices

An option is a tradable contract that confers the right, but not the obligation, to buy (call) or sell (put) an underlying asset at an agreed-upon price during a certain period or on a specific date. There are two major types of options existed: plain vanilla option, exotic options. Plain vanilla option is the first generation options, for example, European options and American options. American option allows early exercise which enables an investor executes his privilege before final settlement date. On the contrary, European option does not allow early exercise. Finally, the payoff of a plain vanilla option is determined by the final settlement price of the underlying asset at maturity.

Besides plain vanilla options, there are three types of exotic options [5]: 1) path-dependent options, 2) multi-factor options and 3) time-dependent options (illustrated in Table 1). The option price of a path-dependent option is determined by the "path" of the underlying asset

before maturity. The multi-factor option's price is determined by two or more underlying assets. The time-dependent option determines the price by time and the underlying asset's current value.

Table 1 Types of exotic options

Path-dependent options	Multi-factor options	Time-dependent & other options
Average rate option (Asian option)	Rainbow option	Chooser option
	Quanto option	Forward start option
Barrier option	Basket option	Binary option
Lookback option		Compound option
Ladder option		Pay-later option
Shout option		Bermudan option

This study focuses on developing a pricing model for European options because it can be modified to price American options. Moreover, cross field applications such as Real Option Analysis (ROA) often apply option pricing models related to European options as well.

For a typical European option, the contract is issued with a certain underlying asset to be settled after T days with an agreed upon price K to buy (call) or sell (put) a predetermined amount of the underlying asset. Suppose the current value of the underlying asset is S_0, the relationship among put value, call value, time to maturity T, and volatility is demonstrated in Figure 1. For European options, the contract holder must keep the contract until its final settlement day and then evaluate the underlying asset's price to determine whether the contract should be executed or not. It is obvious that the option price must be determined in a fair manner to deliver an unbiased insight of the contract's value. Thus, it is important to realize the "fair value" of options. That is the reason why almost all option pricing models

focus on obtaining a theoretical fair price of a certain option to provide traders with a rational basis that can be used to judge whether actual option prices in the marketplace are reasonable, and to help investors determine where to place bids and offers.
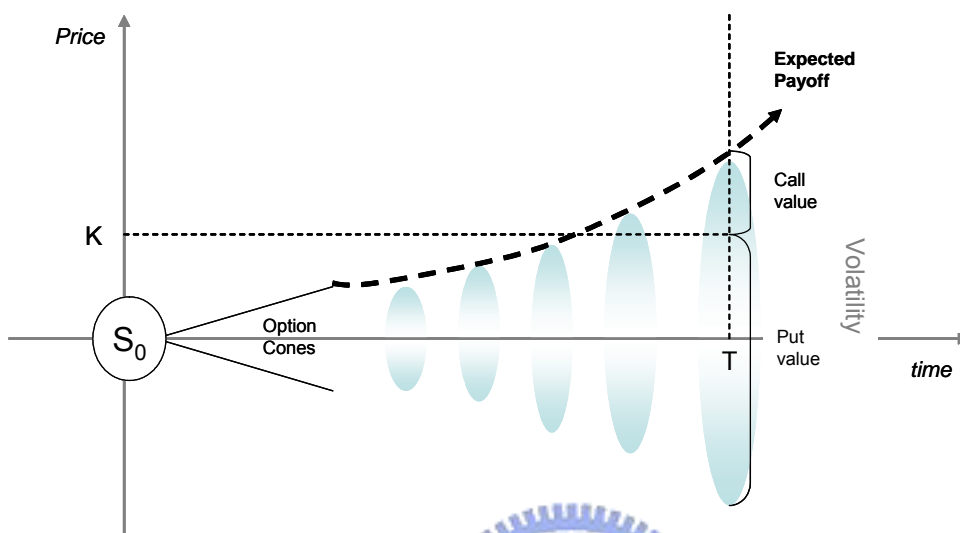
Figure 1. Relationship between option values and time to maturity

The fair value of an option is basically determined by market participants. Also, a theoretical option price must be fair in order not to benefit any participant or introduce arbitrage opportunities. That is, the fair price allows no exploitation by traders or investors for a profit when the option value is efficiently priced by the market [6]. As a result, a fair option price must: a) generate no arbitrage opportunities, b) make no profit for both buyers and writers.

a) Generates no arbitrage opportunities: If there is any arbitrage opportunity exists, market participants will execute arbitrage strategies in order to obtain riskless profits. As a result, biased prices will be influenced by investors and no more arbitrage opportunities exist. Thus a fair value must promise no arbitrage opportunities that are widely known as put-call parity. Put-call parity is a strong arbitrage relation first identified by Stoll [7], that between the prices of European put and call options having same underlying with the same strike and expiry, the combinations of options can create positions that are the same as holding the underlying itself. Thus, the call price C, put price P, strike price K and underlying stock price S must satisfy the

following equation: $C - P = S - PV(Dividends) - K e^{(-r\,T)}$, where T is the option holding period, r is a continuously compounded interest rate and $PV(Dividends)$ is the present value of the dividends received by the stock owner over the holding period. Although put-call parity identifies the relationship between options and underlying assets, in real world scenario, it is hard to duplicate a stock index with limited stock positions. However, the stock index can be easily duplicated with its relative futures. If there simultaneously exists both options and futures contract of the same underlying asset with the same strike and expiry, put-call parity becomes $C - P = FP - PV(Dividends) - K e^{(-r\,T)}$ where FP is the futures price of the relative underlying asset.

b) Makes no profit for both buyers and writers: If the theoretical option price is always under priced then there will be no option writers selling options with that theoretical value because no one wants to lose money; if the theoretical price is always overpriced then there will be no option buyers purchase that option with theoretical value. A fair option price must also promise an unbiased value that it is equally fair to buy or write an option contract.

To sum up, a fair option price guarantees equal profits to all market participants and make no arbitrage opportunities. As a result, every participant in the market will earn only riskless interest rate if the price is fair.

2.2 Risk-neutral distributions

Cox and Ross (1976) established the option price as the expected payoff value discounted at the risk-free interest rate over the risk-neutral distribution of the underlying asset [8]. Nowadays, almost all option pricing models apply risk neutral mathematical distributions to obtain the fair option price. Thus, it is important to realize what a risk neutral distribution is.

Risk-neutral means that investors are not risk-averse but instead they do not demand a discounting of the price to take account of risk. With this regard, most option pricing

methodologies estimate the future distribution of the underlying asset in terms of risks, but the pricing models take the average expected price to be the same as holding a riskless bond [9]. The term risk-neutral probability distribution is used to refer to probability distributions which when used as weights in an expected-value calculation will reproduce the market value of financial instruments [10]. Generally speaking, risk-neutral probabilities differ from real-world probabilities because the market price of the underlying asset does not assign value in the same way that a risk-neutral individual would. A risk-neutral distribution must satisfy the following formula [11]: $0 = E(p_n - p_0)$ where $E(.)$ denotes the expected value, p_n is the future value of the underlying asset, and p_0 is the current value of the underlying asset. More practically, if the mean value μ of a payoff distribution equals zero then this distribution is risk-neutral.
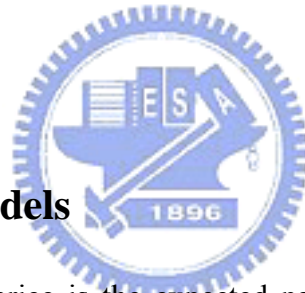
2.3 Option pricing models

As discussed before, option price is the expected payoff value discounted at the risk-free interest rate over the risk-neutral distribution of the underlying asset. However, applying none risk-neutral distributions like the real payoff distribution rather than mathematical risk-neutral distributions is difficult because such distributions rarely behaves in a risk-neutral manner. Applying a distribution with non risk-neutral characteristic will violate put-call parity rules [7] because of the arbitrage possibilities associated with the derived put and call prices. This characteristic limits the use of applying practical distribution to price options. Also, a researcher cannot justify if it is worthy to develop closed form formulas to price options with new types of distributions. A simple example is that if a distribution is risk-neutral then the mean value μ must equal zero. However, the μ in a actual distribution rarely equals zero. Another example is that if a researcher wants to integrate 2 or 3 mathematical distributions to

price an option, he may use normal distribution random number generator to produce the simulated distribution space but it does not work because the mean value of randomly generated samples will not be zero. Another example of demonstrating difficulty in applying actual distribution is that it needs different distribution maps for different time to maturity (because this study demonstrates that the actual payoff distribution is time variant, which will be illustrated in the next section). For example, at least n different distribution maps are needed to valuate the option price if it is n days before maturity. Thus, if the sampling data is huge then the pricing speed will be too slow for practical use. Moreover, short-term asset volatility is rarely consistent with that implied by the actual distribution map, leading to significant pricing errors. Consequently, even traders want to apply real asset return distributions, it cannot be practically used to obtain the option price, encouraging researchers to apply mathematically risk-neutral distributions instead. The most classical of these approaches is the BS model, which assumes that the payoff of the underlying asset follows the geometric Brownian motion and has a lognormal distribution with constant volatility and risk-free interest rate before maturity [3]. Since the development of the BS model, more realistic option pricing methodologies have been developed, including: (a) the stochastic interest-rate/volatility option model [12][13][14]; (b) jump-diffusion related models [15][16]; (c) Markovian models [17][18]; and (d) stochastic-volatility jump-diffusion models [19][20]. However, all these models focus on identifying the "right" distributions and pricing options using close form formulas. Consequently, the mathematical distribution never perfectly fits any underlying asset's actual payoff distribution.

Recently, researchers addressed two related empirical phenomena to demonstrate that the real-world distribution do not follow the basic assumption of Black-Scholes model. As a result, researchers forward continuous improvement of pricing methodologies. The first

phenomena is the asymmetric leptokurtic features indicating the return distribution is skewed to the left side and accompanied with a higher peak and two fat tails than those in normal distribution. The second is 'volatility smile' which discloses that the implied volatility curve is a convex curve of the strike price while the implied volatility should be constant in Black-Scholes model. In order to deal with the asymmetric leptokurtic features mentioned above, numerous models have been proposed. For example, chaos theory fractal Brownian motion and stable processes [21][22], generalized hyperbolic models[23][24], and time-changed Brownian motions[16][25][22][27]. These models might demonstrate some analytical formulae for standard European call and put options, but they may be unrealistic to obtain analytical solutions for option prices.

Meanwhile, different models are also propose to tackle the "volatility smile" in option pricing include: (1) stochastic volatility and ARCH models[28][29][30], (2) constant elasticity model [8][31], (3) normal jump models[12], (4) affine stochastic-volatility and affine jump-diffusion models[32][33][34], (5) implied binomial trees methods[35][36]. These models might not be easy to address analytical solutions for option pricing. Moreover, some of these modes may not solve the asymmetric leptokurtic feature.
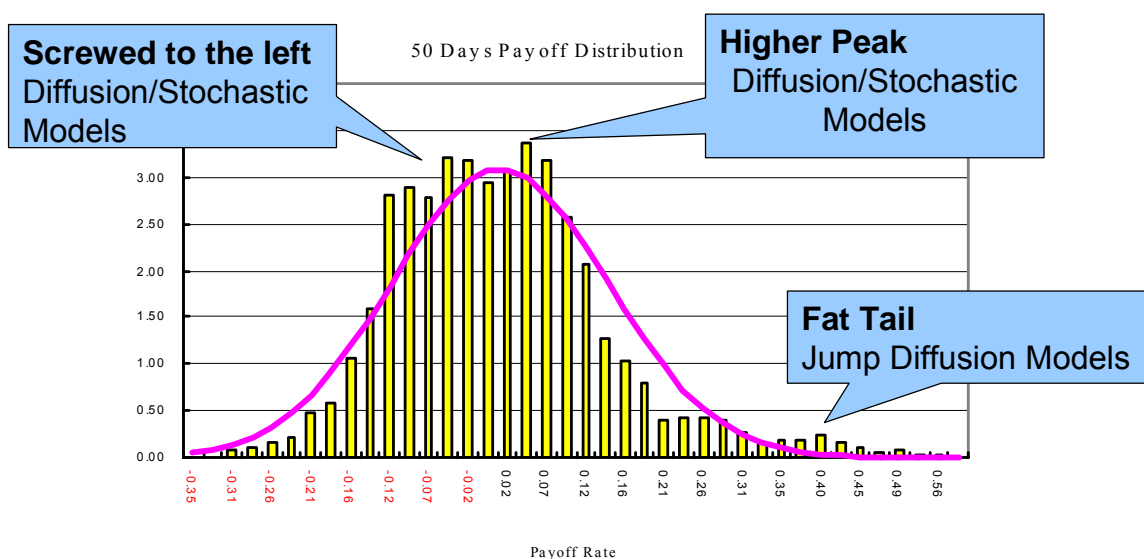
In computer science, attempts have also been made to price options using artificial intelligence models to improve options pricing performance. The most popular of these methods is the neural network approach. Unlike classical mathematical methodologies, a neural network is a non-parametric estimation technique which does not make any distributional assumptions regarding the underlying asset. Instead, this approach develops a model using sets of unknown parameters and lets the optimization routine seek the best fitting parameters to obtain the desired results. For example, Hutchinson-Lo-Poggio demonstrated that the neural network approach can be used to price S&P future options [37]. Andrew Carverhill followed this line of research and examined the best method of establishing and train a multi-layer perceptron neural network for option pricing and hedging [38]. Meissner-Kawano also trained neural networks using option prices to address the smiling effect [39] associated with options' implied volatilities. All these works demonstrate that modern computational theories can offer alternative options pricing methods. In the real world, deriving closed form formulas to price options is time consuming, which decelerates the speed of construction new derivative instruments. Moreover, user may use random number generator to produce a payoff distribution to emulate real assets or simply apply actual distributions instead. However, no solutions for pricing options with user defined distributions existed. Thus, this study focused on determining options price using user defined distributions. Moreover, the proposed model also supports time variant distributions in order to maximize the flexibility of pricing performance. This study also introduces the "real" payoff distribution obtained from a historical sample of Taiwan stock market to demonstrate the proposed model can handle none risk-neutral distributions [58].

2.4 Studies on implied volatility

In order to observe the real behaviour of an underlying asset, this study introduces a normalization method associated with implied volatility distributions and futures prices. Thus, this section briefly discusses various research topics about implied volatility and future prices. Most financial texts confer that the futures markets aggregate diverse information and expectations regarding the future prices of underlying assets, and thus provide a common reference price which is known as the price discovery function of futures. For example, three topics are commonly discussed in relation to the price discovery function of futures. The first topic deals with the lead-lag relationship and information transmission between the prices of national markets, or between different securities [40][41]. The second topic involves the discussion of volatility spillovers, since volatility is also a source of information [42][43]. The third topic relates to the phenomenon of information transmission between stock index and index futures markets [44][45][46][47].

Similar to futures markets, options markets may also provide a common reference of subsequent real volatility (RV) by calculating the implied volatilities (IV). Early research on the predictive capability of IV found that IV explains variation in future volatilities better than that in historical volatilities (HV). For example, Lantane and Rendleman [48] found that actual option valuations were better explained by actual volatility over the life of the contract than by historical volatility. Chiras and Manaster [49] also tried to compare the predictive power of IV and HV using the CBOE data, and found that IV has superior forecasting power to HV. However, subsequent studies applying time serious methodologies to study the predictive power of implied volatility have yielded mixed results. Some studies have found that IV is a poor method of forecasting subsequent RV, while other studies have found that IV is a good method of forecasting RV. For example, Canina and Figlewski [50] found virtually no relation

between the IV and subsequent RV throughout the remaining life of S&P 100 index options before maturity date. Moreover, Day and Lewis [51] and Lamoureux and Lastrapes [52] both found that GARCH associated with HV is better able to predict RV than IV. Meanwhile, other studies have found that IVs provide reasonably good information on the subsequent RVs of the underlying asset. For example, Harvey and Whaley [53] tested and rejected the hypothesis that volatility changes are unpredictable. Moreover, Fleming [54] examined the performance of the implied volatility of the S&P 100 for forecasting future stock market volatility, and found that although IV has an upward bias but it contains relevant information regarding future volatility.

3. Computational approach for pricing European options

This study uses actual distributions to demonstrate the effectiveness of applying the proposed model to price European options. For this reason, this section first discusses observations on actual distributions. Next, this section proposes a computational method of pricing European options with distribution maps saved in databases. In order to prove the effectiveness of the proposed model, this study uses high frequency time interval with one minute time ticks to compute empirical samples. High frequency examples are used to obtain large samples for verification purposes if the execution efficiency of this computation method can feasibly be applied to real world applications. The same concept can also be applied to price European options regardless of time interval.

3.1 Observations regarding actual payoff distributions

Most option pricing models use mathematical distributions. For example, the BS model assumes that underlying assets follow a geometric Brownian motion with lognormal returns. Meanwhile, other sophisticated option pricing methodologies like the stochastic volatility model apply a flexible distributional structure in which the correlation between volatility shocks and underlying stock returns controls the level of skewness, and use the volatility variation coefficient to control the kurtosis level [19][20]. However, none of these mathematical distributions can precisely describe underlying asset behaviour in the real world. To observe the real behaviour of the underlying assets, this study uses sampling data for the period 01/01/1991 to 09/03/2007 from the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX). Because most mathematical option pricing models discuss the underlying asset return distribution using lognormal related distributions (or with certain modifications), this study calculates the asset return rate as $\ln(\frac{P_t}{P_0})$ with different times to maturity where P_0 is the original price and P_t represents the price after t days. The actual distributions are compared with the normal distributions as listed in Table 2.

Table 2. The actual distribution maps compared to the normal distributions.

1. The X-Axis is the nature log asset return rate in % and the Y-Axis is the possibility value in %.
2. The histograms represent the real payoff distribution and the curve lines represent the normal distribution.
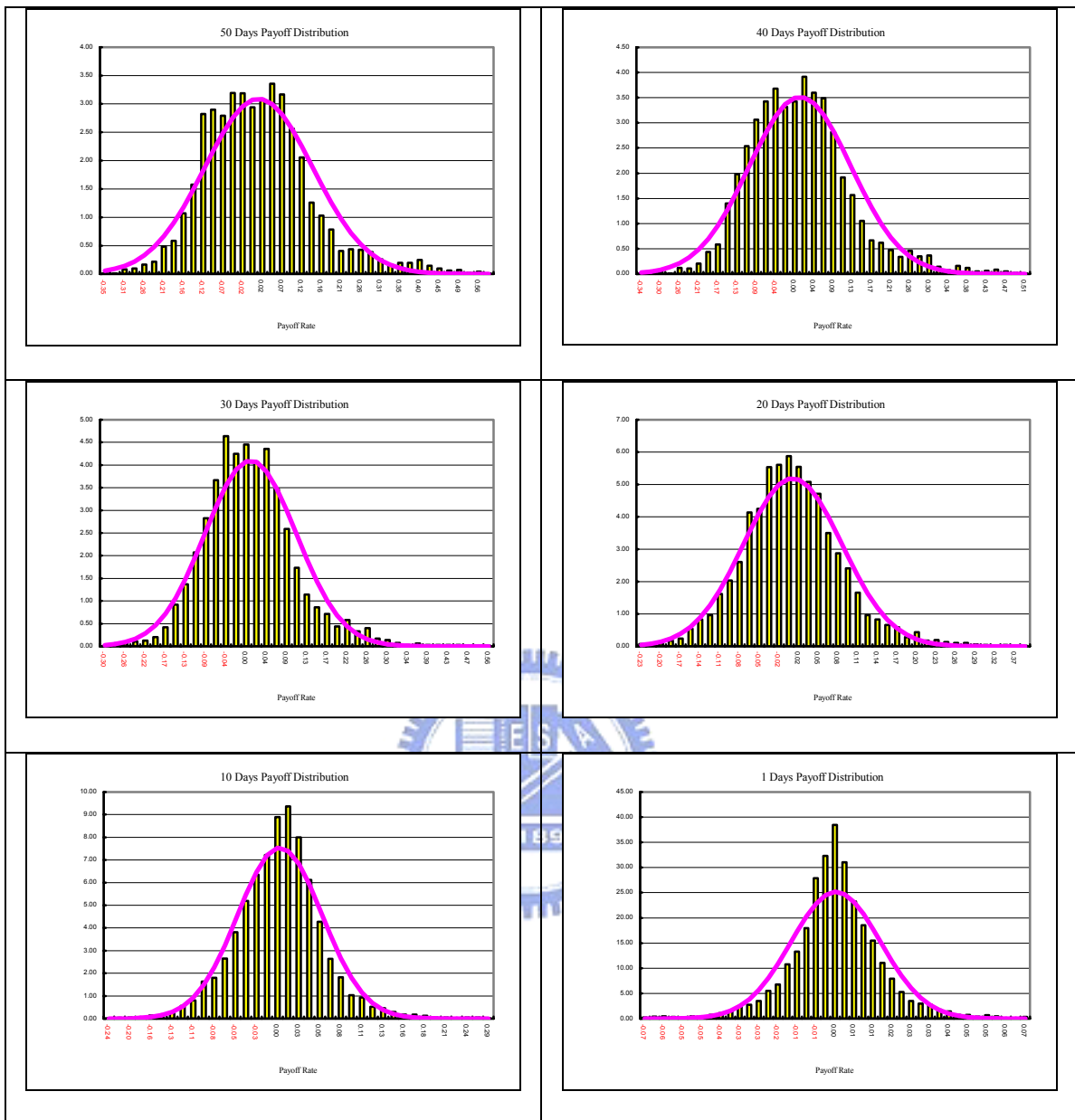
From Table 2, the real payoff distribution of the asset (TAIEX) varies with days-to-maturity. That is, the actual distribution is time variant. The most interesting finding is that the actual distribution exhibits twin-peak phenomenon in 30, 40 and 50 days to maturity distribution maps. Restated, when days to maturity exceeds 30, the real asset return rate distribution displays two peaks. This twin peak phenomenon has received little attention from academics.
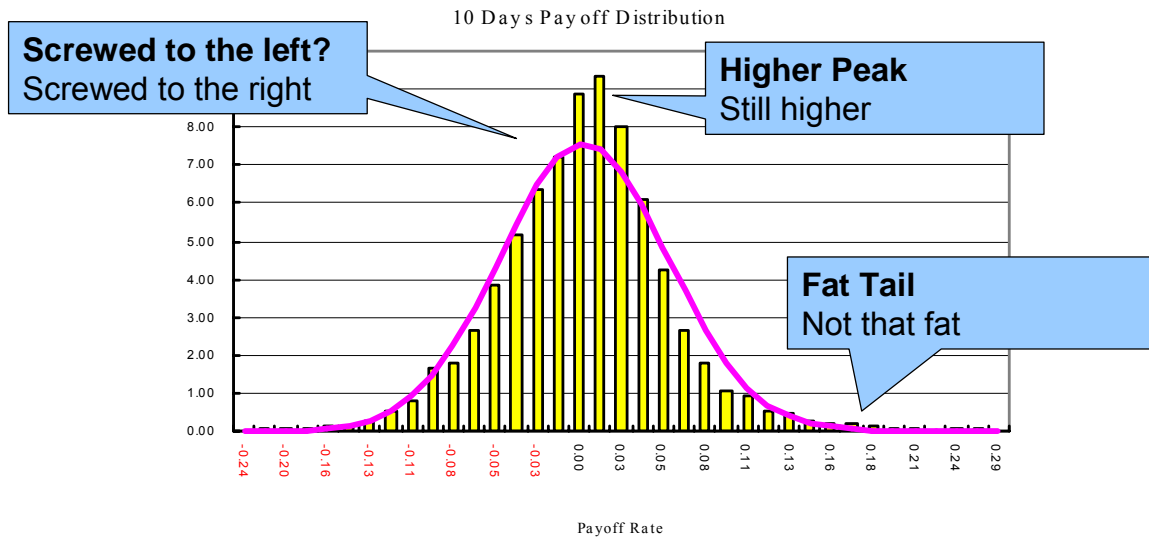
The actual distribution clearly shows that mathematical distribution approaches have difficulty obtaining precise option price (at least for the Taiwan stock market), because the actual distribution varies according to time to maturity. The time variant distribution issue limits the use of fixed mathematical distribution pattern across the entire time to maturity range because variation in time to maturity requires the option pricing model to apply time variant distributions. However, it is difficult for mathematical models to apply different distributions for different time to maturity. Furthermore, behavior may differ among assets and markets, so a mathematical model must apply different distributions to maximize its pricing performance for different assets or different markets. Another issue is that the actual payoff distributions, like the time variant distributions with twin peak phenomenon, which are difficult to be described by using mathematical distributions. This issue also limits the cross-field applications of using the traditional option pricing models.

Finally, the standard deviation of each distribution map from D1 to D250 is plotted in Figure 2.

20
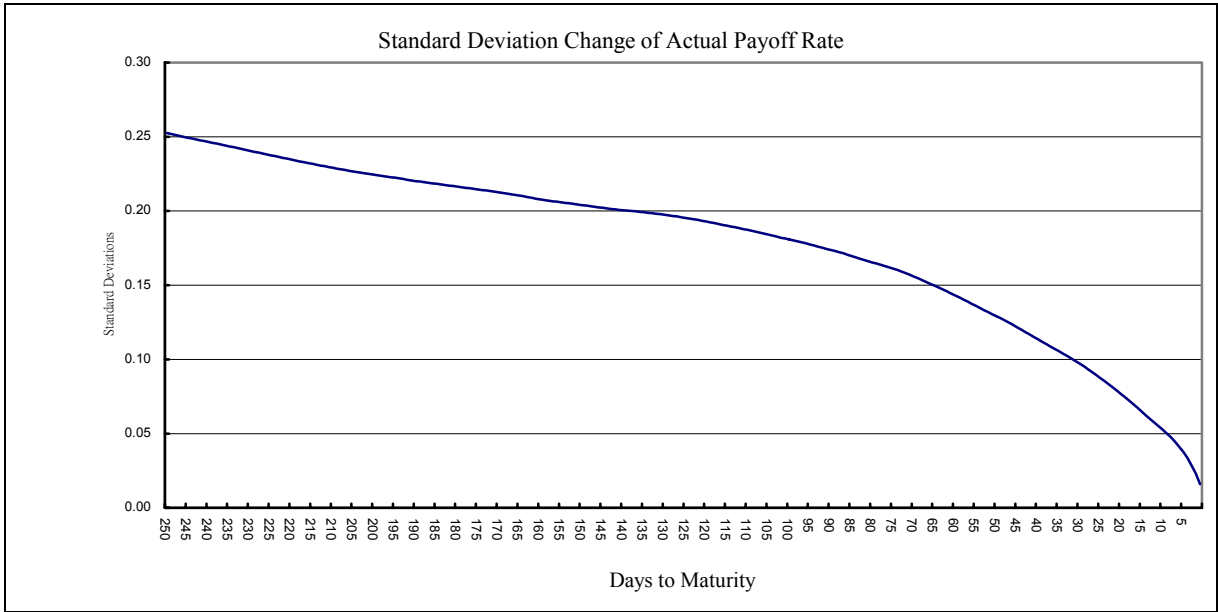
Figure 2. Standard deviations of actual distributions

3.2 Observations regarding implied volatilities

In order to observe actual payoff behaviors associated with implied volatilities, this study designs a method to demonstrate the relationship between implied volatilities and the underlying asset's final settlement value. Most researches related to market price observation among various options written on the same underlying asset used to estimate a single volatility with at-the-money or near-the-money options since they are more sensitive to volatility changes and least susceptible to the effect of the bid-ask spread [55]. Beckers [56] and Canina and Figlewski [50] suggests that the option that near the money are better predictors to future real volatility than the IVs of deep in or out of the money options. Thus, this research uses only the daily close price of near-the-money nearby options contract having same underlying with the same strike and expiry for verification.

List 1. Notations used in this paper

Notations List:

C = Call Price

P = Put Price

σ = Volatility

r = Interest Rate

S = Spot Price

K = Exercise Price

T = Time to Mature in years

t = Time to Mature in days

N() = Cumulative Standard Normal Distribution Function

$N_d(\mu, \sigma)$ = Normal Distribution Function with mean equals μ and standard deviation

equals σ

IV = Implied Volatility, yearly

V_t = Implied Volatility of *t* days.

FP = Futures Price

SP = Final Settlement value

According to Black-Scholes model, the call price *C* and put price *P* of a European option can be calculated as follows:

$$C = SN(d_1) - Ke^{-rT}N(d_2)$$ (3.1)

$$P = Ke^{-rT}N(-d_2) - SN(-d_1)$$ (3.2)

$$d_1 = \frac{\ln(\frac{S}{K}) + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$$ (3.3)

$$d_2 = \frac{\ln(\frac{S}{K}) + (r - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$$ (3.4)

This paper modified the calculation method of deriving implied volatility in order to generate an unbiased probability space. Traditionally, computing an implied volatility (IV) requires

solving (3.1) or (3.2) repeatedly with different trial values for the volatility input. However, the derived IV of a call option (3.1) is rarely equal to the value obtained from a put option (3.2) with the same exercise price and thus most IV related researches use only call or put option into discussions. However, with the relationship indicated in put-call parity, the *FP*, *C* and *P* of the same underlying with the same strike and expiry are tightly coupled with each other that any slight price change of one item will immediately cause the price moves of the other two items. Thus, in order to verify the possibility space formed by FP and IV, this study combines both put and call IVs into a single value IV, namely, the union IV (IV_u). Combining (3.1) and (3.2), the IV_u can be solved with the following formula:

$$C + P = SN(d_1) - Ke^{-rT}N(d_2) + Ke^{-rT}N(-d_2) - SN(-d_1)$$
(3.5)

Let $\mathcal{O}_c = \{C_1, C_2, ..., C_n\}$ denotes the historical near-the-money call options prices in time interval $\mathcal{T}$, $\mathcal{O}_p = \{P_1, P_2, ..., P_n\}$ denotes the historical near-the-money put options prices in $\mathcal{T}$ and $\mathcal{F} = \{FP_1, FP_2, ..., FP_n\}$ represents the futures prices in $\mathcal{T}$. For the ith historical data in $\mathcal{T}$, using (C_i, P_i, FP_i) to generate the ith sampling distribution space and the actual final settlement value of (C_i, P_i) is SP_i.

Consider the ith historical data in $\mathcal{T}$, take C_i and P_i into (3.5), the yearly IV_u of the ith day in $\mathcal{T}$ can be solved. Let $\mathcal{V} = \{V_1, V_2, ..., V_n\}$ denotes the IV_u in $\mathcal{T}$. Transfer the yearly IV_u into days to mature scale, let v_t denotes the union implied volatility of t days to mature:

$$v_t = \sqrt{\frac{t \times V_i^2}{365}}$$
(3.6)

Suppose SP_i and FP_i is the relative final settlement value and future price of the ith historical data in $\mathcal{T}$, the actual SP is located at d_i standard deviations of an $N_d(\mu, \sigma) = N(FP_i, v_t)$

distribution space:

$$d_i = \frac{\ln(\frac{SP_i}{FP_i})}{v_t}$$

(3.7)

Thus, for each historical record in $\mathscr{S}$ an associated d_i can be calculated, the collection of d_i forms a sampling space of standard deviations $\mathscr{D} = \{d_1, d_2, ..., d_n\}$. If the final settlement value follows a normal distribution of $N_d(\mu, \sigma) = N_d(FP_i, v_t)$, the distribution of $\mathscr{D}$ will be a standard normal distribution, $N_d(\mu, \sigma) = N_d(0, 1)$.
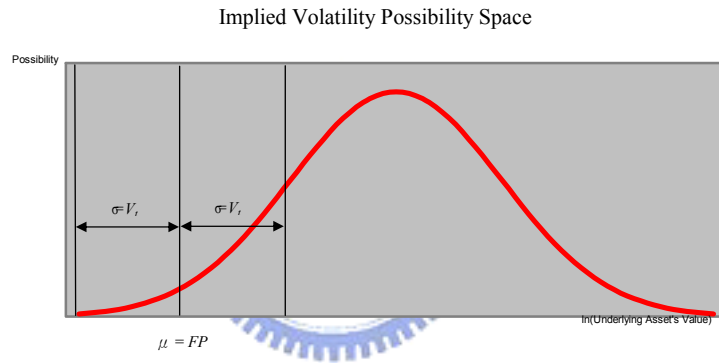
Figure 3: Market Expected Possibility Space generated by (FP, V_t)

To observe the behaviors of (FP, IV), this study uses the option and futures prices of Taiwan Stock Exchange Capitalization Weighted Stock Index (TX) from 24/12/2001 to 31/12/2006 in order to testify the results, $\mathscr{S}=[24/12/2001,31/12/2006]$.

The risk-less interest rate r applied is monthly fixed deposit interest rate collected from the Central Bank of Taiwan. The calculation results of $\mathscr{D}$ are accumulated into Figure 4 in order to compare to normal distribution $N_d(0, 1)$, the (mean, standard deviation) of $\mathscr{D}$ is (0.0755, 1.1066) respectively.
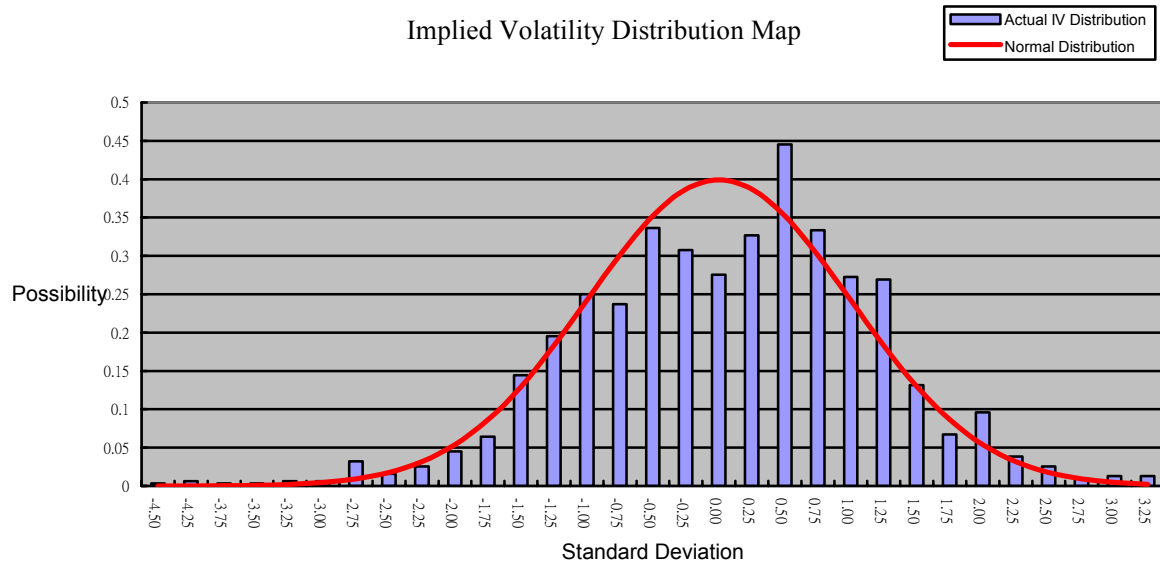
Figure 4: The actual implied volatility distribution map of (FP, IV) compared to the normal distribution,

$D(\mu,\sigma,skewness,kurtosis) = D(0.0755, 1.1066)$

Figure 4 exhibits that the distribution of $\mathcal{D}$ is roughly similar to $N_d(0,1)$ and is skewed left and peaked compared to $N_d(0.0755, 1.1066)$ but slightly flatter and skewed right compared to $N_d(0,1)$ with thicker tails. In another words, the actual final settlement values of Taiwan Stock Index futures and options roughly follow the behavior of market expectation formed by FP and IV with slightly larger volatilities. Another interesting observation is that there are two possibility peaks located at +0.5 and -0.5 standard deviations. This phenomenon indicates that the final settlement value of the underlying asset tends to bias half an implied volatility in (FP, IV) distributions.

Using the close price of stock index instead of the FP in (3.7), this study generates another possibility space formed by IV only. Let $\mathcal{D}' = \{d'_1, d'_2, ..., d'_n\}$, the (3.7) becomes:

$$d'_i = \frac{\ln(\frac{SP_i}{S_i})}{v_t}$$

(3.8)

The computed result of $\tilde{D}'$ in $\tilde{S}$ are summarized into Figure 5. Compared to normal distribution $N_d(0, 1)$, the (mean, standard deviation) of $\tilde{D}'$ is (0.046030, 1.057034) respectively.
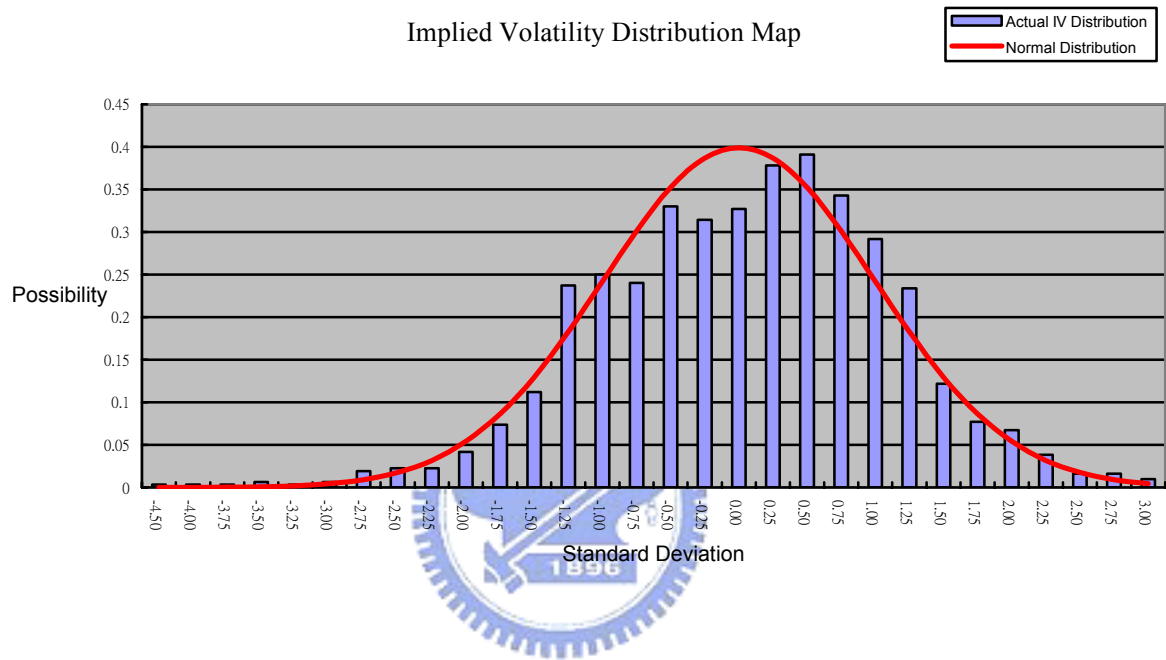
Figure 5: The market implied volatility distribution map of (S, IV) compared to the normal distribution, D'(μ,σ)

= D'(0.046030, 1.057034)

Figure 5 exhibits that the distribution of $\tilde{D}'$ is more similar to $N_d(0,1)$ than $\tilde{D}$ and is skewed left and peaked compared to $N_d(0.046030, 1.057034)$ but slightly flatter and skewed right compared to $N_d(0,1)$ with thicker tails. Comparing $\tilde{D}$ and $\tilde{D}'$, the actual final settlement values of Taiwan Stock Index futures and options roughly follow the behavior of the distribution map formed by FP and IV but even closer with the expectation of S and IV.

The cumulative possibility distributions from interval ±σ to ±3σ of normal distribution, $\tilde{D}$ and $\tilde{D}'$ are listed below:

Table 3: The cumulative possibility distributions

		±σ	±2σ	±3σ
Normal Distribution	Value	0.6827	0.9545	0.9973
Distribution of ℑ(FP, IV)	Value	0.6290	0.9303	0.9888
	Delta (%)	-7.87	-2.54	-0.85
Distribution of ℑ(S, IV)	Value	0.6611	0.9407	0.9952
	Delta (%)	-3.18	-1.41	-0.23

Table 3 shows that the cumulative possibility of ℑ is smaller than normal distribution about

7.87% in plus minus one standard deviation while 0.85% smaller in ±3σ. ℑ′ is even closer to

normal distribution in our testing example, however, both ℑ and ℑ′ clearly indicates that the

option price is slightly under estimated. Observing Figure 4, Figure 5 and Table 3, this study

concludes that the behavior of TX does not follow normal distributions. However, this

conclusion also suggests that an investor can roughly expect the final settlement value of a

stock index will locate at the futures price with a standard deviation equals to the implied

volatility derived from option price. With this conclusion, investors can expect the price

behavior of their stock positions will roughly follow the price behavior of holding both

futures and options combinations but most likely having an expectation bias of ±0.5 implied

volatilities. Another observation in this research also suggests that (S, IV) combinations are

more close to the price behaviors of a stock index than (FP, IV) combinations, that is,

applying options only may obtain better risk management performance than using both

options and futures combinations.

Observing the testing results in section 3.1 and 3.2, this study concludes that the actual

behavior of TX does not follow any mathematical distributions. More precisely, it is difficult

to approximate an underlying asset like stocks by applying mathematical distribution

combinations. With this conclusion, this study first introduces the idea to price options with actual distributions.

3.3 The computational approach

Option price is the expected value of the payoff discounted at the risk-free interest rate over the risk-neutral distribution of the underlying asset. Thus, given the price S and an agreed-upon price K for the underlying asset applicable during a certain period T, the option value can be described as follows:

$$C = E(Max(S\text{-}K, 0))$$

$$P = E(Max(K\text{-}S, 0))$$

Where C denotes the call option price, P represents the put option price, and $E(.)$ is the expected value.

In the real world the price of most assets varies continuously, and this variation is described as volatility σ. An option pricing model calculates C or P of the underlying asset under the circumstances (S, K, σ, T, r). The proposed computational model applies actual distribution maps to calculate the desired option price as illustrated in Figure 6.
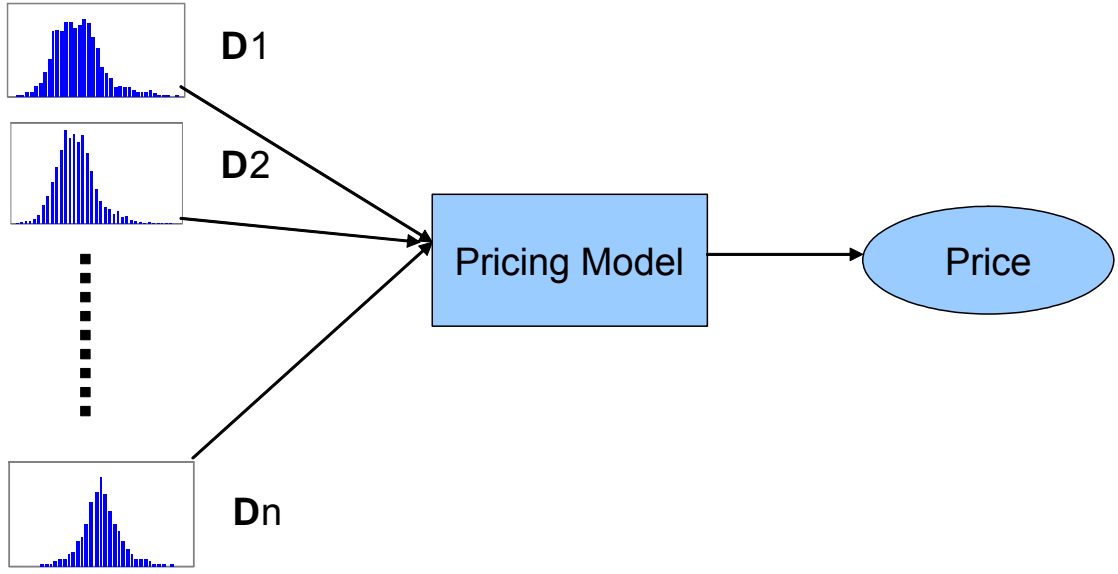
Figure 6. The core concept of the proposed model

Assume I days of sampling data, with each day containing J time ticks. Then for each sample of ith day and j^{th} time tick $X_{i,j}$, the tick payoff rate $R_{i,j}$ is

$$R_{i,j} = \begin{cases} \dfrac{X_{i+1,1} - X_{i,j}}{X_{i,j}}, \text{if the final settlement price is determined by the opening price on the final settlement day} \\ \dfrac{X_{i+1,n} - X_{i,j}}{X_{i,j}}, \text{if the final settlement price is determined by the closing price on the final settlement day} \end{cases}$$

Notably, $R_{i,j}$ can also be represented as $ln((X_{i+1,1}- X_{i,j})/X_{i,j})$ or $ln((X_{i+1,n}- X_{i,j})/X_{i,j})$ based on the assumptions of the BS model. However, the difference of applying logarithm or simple payoff rate is minor for high frequency applications. This study avoids unnecessary use of floating point functions to increase execution speed.

When using this computational approach to price options, the $R_{i,j}$ should be modified as:

$$R_{i,j} = \begin{cases} \dfrac{PV(X_{i+1,1} - X_{i,j})}{X_{i,j}}, \text{if the final settlement price is determined by the opening price on the final settlement day} \\ \dfrac{PV(X_{i+1,n} - X_{i,j})}{X_{i,j}}, \text{if the final settlement price is determined by the closing price on the final settlement day} \end{cases} \quad \text{where}$$

PV(.) denotes the present value associated with riskless interest rates. Notably, it makes

29

virtually no difference to calculate *PV(.)* for less than 50 days to maturity because the impact

on interest rates is minor for short term options.

The payoff rate can be preprocessed and stored in a database table for further use in achieving a

reasonable execution speed when calculating option prices for practical use.

Assume an option matures the next day and has strike price S, final settlement price S_t, exercise

price K and current time-tick j. Given m sampling days (which can only generate m-1 sample

entries), the call price C can be approximated as follows:

$$C(S,K,j) = E(\max(S_t - K,0)) = \frac{\sum_{i=1}^{m-1} \max(S \times (1+R_{i,j}) - K,0)}{m-1}$$

Similarly, the Put price P can be approximated as follows:

$$P(S,K,j) = E(\max(K - S_t,0)) = \frac{\sum_{i=1}^{m-1} \max(K - S \times (1+R_{i,j}),0)}{m-1}$$

Consider the riskless interest rate r with time to maturity τ, the Call/Put price can be represented

as:

$$C(S,K,j,r,\tau) = \frac{\sum_{i=1}^{m-1} \max(S \times (1+R_{i,j}) - e^{-r\tau}K,0)}{m-1}$$

(3.9)

$$P(S,K,j,r,\tau) = \frac{\sum_{i=1}^{m-1} \max(e^{-r\tau}K - S \times (1+R_{i,j}),0)}{m-1}$$

(3.10)

However, when attempting to determine the option price using (3.9) and (3.10), it quickly

becomes obvious that the calculated price does not follow the put-call parity rule because the

mean value μ of a actual distribution does not equal zero (implying the actual distribution is not

risk-neutral). Notably, arbitraging opportunities occur when the distribution is not risk-neutral.

Furthermore, the actual distribution has its own volatility which is difficult to change. For

example, if a real payoff distribution is formed based on a ten year period of sample data and

has a standard deviation σ_1, but the forecasted volatility of the target option is σ_2, then the option must be priced using a distribution with a standard deviation σ_2 rather than σ_1. If the intrinsic volatility of the actual payoff distribution cannot be transformed to fit the short term volatility, the pricing error will be too large for practical use. Given the difficulty of changing the mean value without influence the variance, this study established a computational method for adjusting both the mean value and variance of an existing distribution to obtain the desired values while maintaining a similar distribution to the original.

To obtain risk-neutral characteristics based on the actual distribution, the mean μ of the sampling data must be zero. By observing the actual distribution, if the μ changes from a positive value to zero, the occurrence probability of rightmost (larger) sampling data reduces while the leftmost (smaller) sampling data increases. Based on this phenomenon, a computational method can be developed for adjusting the mean value of the actual distributions by altering the sample occurrence possibilities.

The first step is attaching a weighting factor w_i to each sampled payoff rate $R_{i,j}$. Each w_i is assigned an original value 1.0, indicating that it has a "sampling count" of 1. The Call and Put prices thus can be represented as

$$C(S,K,j,r,\tau) = \frac{\sum_{i=1}^{m-1} \max(w_i(S \times (1 + R_{i,j}') - e^{-r\tau}K),0)}{m-1} \qquad (3.9a)$$

$$P(S,K,j,r,\tau) = \frac{\sum_{i=1}^{m-1} \max(w_i(e^{-r\tau}K - S \times (1 + R_{i,j}')),0)}{m-1} \qquad (3.10a)$$

(Note that $R_{i,j}'$ is the standard deviation transformed payoff rate in 3.15)

For each set of sampling data, the mean value μ' and standard deviation σ' can be calculated as:

$$\begin{cases} \mu' = \dfrac{\sum\limits_{i=1}^{n} R_{i,j} \times w_i}{\sum\limits_{i=1}^{n} w_i} \\[4mm] \sigma' = \sqrt{\dfrac{\sum\limits_{i=1}^{n} (R_{i,j} - \mu')^2 \times w_i}{\sum\limits_{i=1}^{n} w_i}} \end{cases}$$
for the jth tick to maturity

The second step is to adjust the weighting factors to transform the actual distribution into a risk-neutral manner. To achieve this, it is first necessary to sort the sampled payoff rates and position them on the X-axis with weighting factor 1. Assuming that the sample appearance probability changes linearly, the weighting factors can be rotated to modify the distribution, as illustrated in Figure 7. Consequently, by fixing the rotation point to $X = 0$, the weighting factors can be rotated clockwise to decrease the mean values or anti-clockwise to increase them.
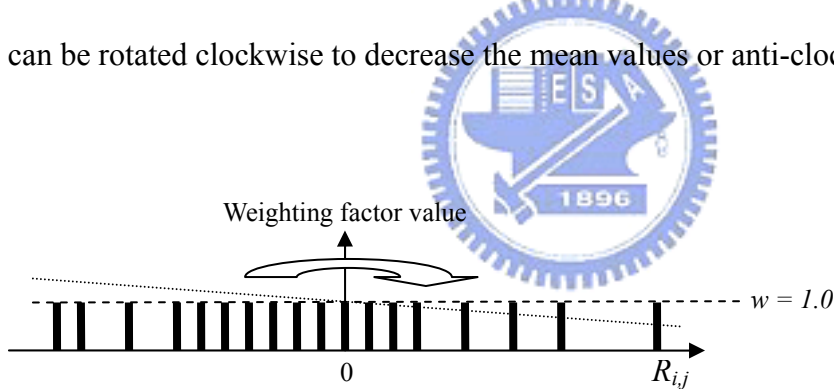
Figure 7. Rotating the factor weights clockwise decreases the mean value of the distribution

The weighting factors can be determined by solving the linear equations through the following steps:

$$X_a = \sum_{i=1}^{n} R_{i,j}\Big|_{R_{i,j} \geq 0} \qquad X_{a2} = \sum_{i=1}^{n} (R_{i,j})^2\Big|_{R_{i,j} \geq 0}$$

Let $\quad X_b = \sum_{i=1}^{n} R_{i,j}\Big|_{R_{i,j} < 0} \quad$ and $\quad X_{b2} = \sum_{i=1}^{n} (R_{i,j})^2\Big|_{R_{i,j} < 0}$ (3.11)

Let m_a denote the slope of the weighting factors for $R_{i,j} \geq 0$, while m_b represents the slope of the weighting factors for $R_{i,j} < 0$.

Solve
$$\begin{cases} X_a m_a = X_b m_b \\ \displaystyle\sum_{i=1}^{n} R_{i,j}(1 - m_a R_{i,j}) + \sum_{i=1}^{n} R_{i,j}(1 - m_b R_{i,j}) = 0 \end{cases}$$

Then
$$\begin{cases} m_a = \dfrac{(X_a + X_b)X_b}{X_b X_{a2} - X_{b2} X_a} \\ m_b = \dfrac{(X_a + X_b)X_a}{X_b X_{a2} - X_{b2} X_a} \end{cases}$$
(3.12)

Thus, the weighting factor can be transformed as follows:

$$w_i = \begin{cases} 1 - m_a R_{i,j} \Big|_{R_{ij} \geq 0} \\ 1 + m_b R_{i,j} \Big|_{R_{i,j} < 0} \end{cases}$$
(3.13)

Combining (3.12) and (3.13) yields the following weighting formula:

$$w_i = \begin{cases} 1 - \dfrac{(X_a + X_b)X_b}{X_b X_{a2} - X_{b2} X_a} R_{i,j} \Big|_{R_{i,j} \geq 0} \\ 1 + \dfrac{(X_a + X_b)X_a}{X_b X_{a2} - X_{b2} X_a} R_{i,j} \Big|_{R_{i,j} < 0} \end{cases}$$
(3.14)

This computational method can transform any distribution into a risk-neutral distribution while largely preserving the characteristics of the original, as shown in Figure 8.
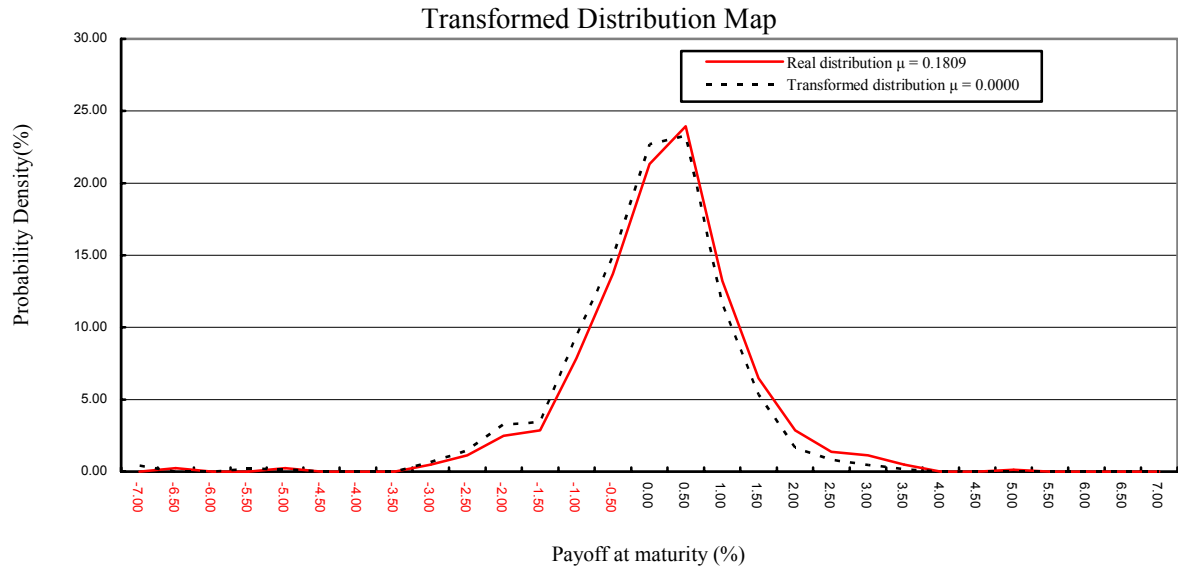
Figure. 8 The transformed distribution after rotated the weighting factors

After transforming the actual distribution into a risk-neutral distribution, the next step is to adjust its standard deviation. If the standard deviation after applying formula (3.14) is v, the desired standard deviation of the distribution map is v'; formula (3.14) then can be rewritten as:

$$w_i = \begin{cases} (1 - \dfrac{(X_a + X_b)X_b}{X_b X_{a2} - X_{b2} X_a} R_{i,j}) \Big|_{R_{i,j} \geq 0} \\ (1 + \dfrac{(X_a + X_b)X_a}{X_b X_{a2} - X_{b2} X_a} R_{i,j}) \Big|_{R_{i,j} < 0} \end{cases} \quad \text{and} \quad R_{i,j}' = \frac{v'}{v} \times R_{i,j} \tag{3.15}$$

Notably, v' must be measured using the time to maturity scale (most option pricing applications use annual volatility). Supposing t days (t is a real number) to maturity and anticipated annual volatility is σ, v' can be estimated by:

$$v' = \sqrt{\frac{t \times \sigma^2}{365}} \tag{3.16}$$

Formula (3.15) can transform the actual distribution into the desired volatility without affecting its mean value while maintaining a similar shape to the original distribution. Figure 9 shows the

transformed distribution. The option price thus can be determined via (3.9a), (3.10a), (3.11), (3.15) and (3.16).
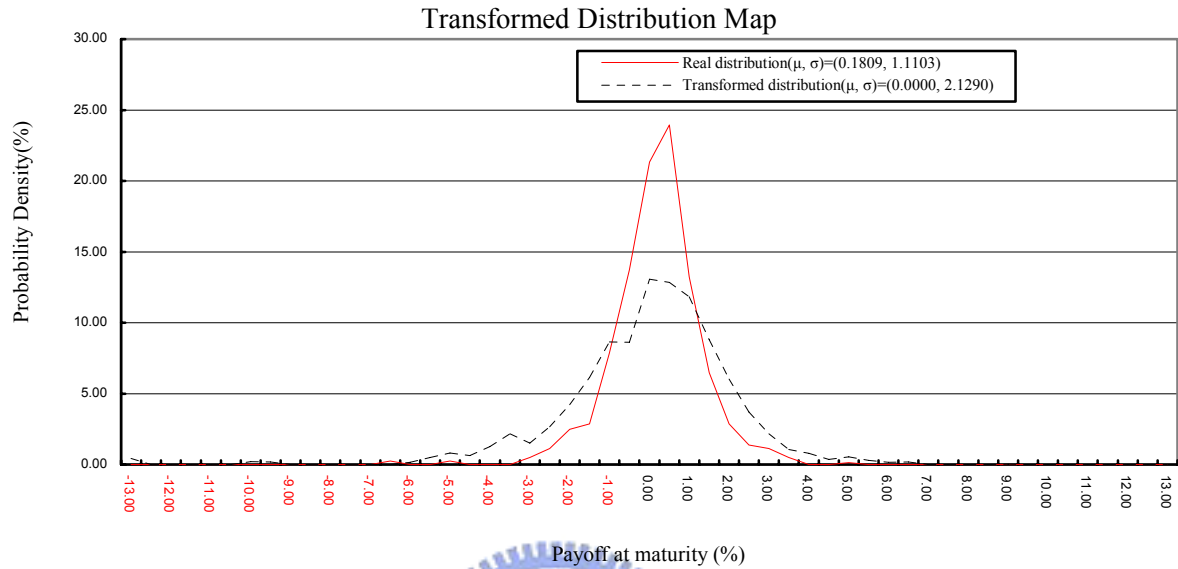
Figure 9. Transformed payoff distribution after applying formula (3.15)

With the same concept, daily option price can also be calculated by the same formulas. Here is a short example to illustrate the calculating process.

Example 1.

Today, suppose TX = 6416. There are 10 days before maturity. The risk-less interest rate = 0.02 and the current call option price on K=6400 is 76. We want to purchase a call option at K = 6400 tomorrow (days to maturity = 9).

Using the following steps to obtain the Call price with the proposed model:

- Step1: Selecting desired distribution $\heartsuit_{T-1}$ according to time to maturity T

- Step2: Using formula (3.14) to adjust selected distribution risk-neutrally

- Step3: Using binary search to obtain v'

- Step4: Applying formula (3.14) to obtain the risk-neutral distribution $\heartsuit_T$

- Step5: Using v' (in 3.15) to transform distribution $\mathcal{D}_T$ and apply (3.9a) to get the price

The calculating processes of this example are described as follows:

- Step 1: Get distribution map of $\mathcal{D}_{10}$, which can use standard SQL to select the desired

  distribution map.

SELECT PayoffRate, 1.0 as w

  FROM DistributionMap

  WHERE DayCount = 10

- Step 2: Applying (3.14) to adjust the selected distribution risk-neutrally, the weighting

  factors are listed below:

PayOffRate	w
-1.174731	1.647186
-0.700737	1.386052
-0.621213	1.342241
-0.617936	1.340435
-0.59823	1.329579
-0.571291	1.314737
-0.566551	1.312126
-0.564137	1.310796
-0.561379	1.309276

$\blacksquare$
$\blacksquare$
$\blacksquare$
$\blacksquare$
$\blacksquare$

0.5123461	0.783126
0.5461306	0.768825
0.5617606	0.762209
0.564909	0.760877
0.5651684	0.760767
0.6291424	0.733687
0.7492108	0.682862
0.8516793	0.639488

- Step 3: Using binary search to obtain v' of $\mathcal{D}_{10}$

  (applying 3.9a) on $\mathcal{D}_{10}$, *v'* = 0.1468

```
Initial Point
High (V'=1.00) =      471.497467673703400000
Low (V' = 0.01)=       17.023088615840430000
Binary Search      0.5050  Price=      240.382318074050200000
Binary Search      0.2575  Price=      126.654373708664200000
Binary Search      0.1338  Price=       70.086734082087530000
Binary Search      0.1956  Price=       98.342580788333190000
Binary Search      0.1647  Price=       84.206287523185980000
Binary Search      0.1492  Price=       77.145461345427150000
Binary Search      0.1415  Price=       73.615543158820740000
Binary Search      0.1454  Price=       75.380506713116690000
Binary Search      0.1473  Price=       76.262984974954260000
Binary Search      0.1463  Price=       75.821746101911440000
Binary Search      0.1468  Price=       76.042365600799640000
V'=         0.1468
```

```
getputcallvalue.prg
* 買賣權: 1 = 買權, 2 = 賣權
PRIVATE ALL

    IH = 1
    IL = 0.01
    CrnPri = m.選擇權價
    PH = Getprice(現價, 履約價, 買賣權, 交易日期, IH, 利率, 到期日數)
    PL = Getprice(現價, 履約價, 買賣權, 交易日期, IL, 利率, 到期日數)
? "Initial Point"
? "High (V'=1.00) = ", PH
? "Low (V' = 0.01)= ", PL
    flag = 0
    i = 0
    DO WHILE flag = 0
        x = (IH + IL)/2
        Pri = Getprice(現價, 履約價, 買賣權, 交易日期, x, 利率, 到期日數)
? "Binary Search ", x, "Price=", Pri
        IF ABS(Pri - CrnPri) < 0.1
            flag = 1
        ENDIF
        IF Pri > CrnPri THEN
            IH = x
        ELSE
            IF Pri < CrnPri THEN
                IL = x
            ENDIF
        ENDIF
        i = i + 1
        IF i >= 30 then
            RETURN x
        ENDIF
    ENDDO
? "V'=", x
    RETURN x
ENDFUNC
```

● Step 4: Applying formula (3.14) to obtain the risk-neutral distribution $\mathcal{D}_9$

Now, pricing the desired option by $\mathcal{D}_9$, select the desired distribution map:

SELECT Rat, 1.0 as w

FROM DistributionMap

WHERE DayCount = 9

Then, applying (3.14) to adjust $\mathcal{D}_9$ distribution risk-neutrally.

● Step5: Using *v'* (in 3.15) to transform distribution $\mathcal{D}_9$ and apply (3.9a) to get the price.

Calculate *v* of $\mathcal{D}_9$, *v* = 0.07;

Apply (*v', v*) = (0.1468, 0.07) into (3.15) and (3.9a), Call Price = 69.66

In practical use, if a trader or researcher wants to generate a user defined distribution for option pricing, he can use the following steps:

Step 1. Choose a random number generator:

Properly choose a random number generator for generating samples.

Step 2. Design distribution maps:

Figure out the desired distribution shape, and then use random generators to produce samples. For example, if a researcher wants to generate a twin peak distribution located at exactly 0.5 and -0.5 stand deviations. He can generate 5,000 samples with Max(N(0.5, 1), 0) and then produce another 5,000 samples with Min(N(-0.5, 1), 0). It is almost sure that the generated distribution map is not risk-neutral because the mean value of the samples does not equal to zero. However, the researcher may safely apply the proposed model to obtain unbiased option values that follows put-call parity.

Step 3. Modify distribution maps:

If the researcher does not satisfy with the produced distribution map, he may generate other sampling data and mixes them into the existed distribution maps. For example, if the researcher wants to increase the thickness of the tail at +0.9 standard deviation and -0.9 standard deviation, he may simply generate 50 samples of N(0.9, 1) and N(-0.9, 1) and mixes them with the pre-generated samples.

3.4 The pricing algorithm

The full pricing algorithm comprises two parts. The first part is the algorithm for preparing the distribution map, while the second part is the pricing algorithm.

3.4.1 The actual distribution generating algorithm

This algorithm is used to generate the actual distribution map to accelerate the calculation process. Because the actual distribution is repeatedly reused for the pricing algorithm, it is optimum to insert new sampling data into the existing distribution maps at the beginning of every trading day (or after trading hours). This algorithm requires minimal execution time if

updates are daily performed. *SettlePrice* indicates the opening or closing price for the asset (depending on whether the final settlement price is determined based on the opening or closing price on the final settlement day) on the specified date *TransactionDate*. The sampling data for the previous day are gathered in a data set *{TimeTicks, TickPrice}* that contains the time tick count and tick price of the underlying asset. The results are stored in the *DistributionMap* table with the primary index set to *(Transaction_date, Time_Ticks)*. The *Transaction_date* field represents the sampling date, the *Time_Ticks* field indicates the time tick counts of the sample, and the *Return_Rate* field stores the asset return rate.

Algorithm MakeRealDistribution

   Input: SettlePrice, TransactionDate, {TimeTicks, TickPrice} of previous trading day

   Output: DistributionMap(Transaction_date, Time_Ticks, Return_Rate)

Begin

   / Clear old data to prevent duplication */*

  DELETE FROM DistributionMap

    WHERE Transaction_date = TransactionDate

  / Insert new data */*

  For Each element pair in {TimeTicks, TickPrice}

   INSERT INTO DistributionMap (Transaction_date, Time_Ticks, Return_Rate)

    VALUES (TransactionDate, TimeTicks, SettlePrice/TickPrice)

  End For

End Algorithm

In practical use of the *DistributionMap*, end users can also write their own programs to generate

any desired mathematical distribution (or combinations) and store the generated samples into the *DistributionMap* table for the pricing algorithm to calculate the desired option price. For example, a researcher may use two lognormal distributions to simulate the twin-peak distribution as observed for the Taiwan stock market to verify whether it is worthwhile to apply two lognormal distributions to the BS model to improve the pricing performance. Researchers do not need to worry whether the two distribution combinations disobey the risk neutral characteristic before deriving sophisticated mathematical solutions. This characteristic increases the versatility of the pricing algorithm for cross field applications.

3.4.2 The pricing algorithm

This algorithm is used to price a European option with *DistributionMap* table generated by *MakeRealDistribution*. Suppose that the parameter set *(S, K, σ, T, r)* used to calculate the option price is *(SpotPrice, ExercisePrice, Volatility, TimeTicks, RisklessInterestRate)*, the pricing algorithm can be described as follows:

Algorithm GetOptionPrice

   Input: SpotPrice, ExercisePrice, Volatility, TimeTicks, RisklessInterestRate, TimeTicks

    Referenced Table: DistributionMap

    Output: CallValue, PutValue

Begin

   Define TargetRate = ExcPri/CrnPri – 1

   Define TargetMeanVaue = 0   //Suppose that the transformed distribution is //Risk-Neutral

SELECT Return_Rate, 1.0 as Weight

FROM DistributionMap

WHERE DistributionMap.Time_Ticks = TimeTicks

INTO CURSOR TmpCursor ORDER BY Return_Rate ASC

//Generate Weighting Factors

Let Cnt = record counts of TmpCursor

Summation from TmpCursor

Let $A = \Sigma Return\_Rate$ for $Return\_Rate \geq 0$

Let $B = \Sigma Return\_Rate$ for $Return\_Rate < 0$

Let $A2 = \Sigma(Return\_Rate \wedge 2)$ for $Return\_Rate \geq 0$

Let $B2 = \Sigma(Return\_Rate \wedge 2)$ for $Return\_Rate < 0$

End Summation   //Formula (3.11)

Let OriginalSD = the standard deviation of Weight in TmpCursor

Let DaysToMarurity = transfer Timeticks to days to maturity

Let TransformedVolatility = Square Root of $(DaysToMaturity*Volatility\wedge 2)/365$

//Formula (3.16)

For Each record in TmpCursor

Replace Weight With (TransformedVolatility/OriginalSD) * (1 – ((A + B) * B/(B * A2-A *

B2)) * Return_Rate)For $Return\_Rate \geq 0$

41

*Replace Weight With (TransformedVolatility/OriginalSD) * (1 + ((A + B) * A/(B * A2 - A*

*   * B2)) * Return_Rate)For Return_Rate < 0*

*   End For   //Formula (3.15)*

*   SELECT SUM(Weight * (SpotPrice * (1 + Return_Rate) - ExercisePrice))/Cnt*

*   FROM TmpCursor*

*   WHERE TmpCursor.Weight >= TargetRate*

*   INTO VARIABLE CallValue   //Formula (3.9a), processed by SQL*

*   SELECT SUM(Weight * (ExercisePrice - SpotPrice * (1 + Return_Rate)))/Cnt ;*

*   FROM TmpCur*

*   WHERE TmpCursor.Weight < TargetRate*

*   INTO VARIABLE PutValue   //Formula (3.10a), processed by SQL*

*   RETURN CallValue, PutValue*

End Algorithm

The above algorithm is carefully optimized for modern database applications involving SQL syntax and summarizing operations. The elimination of unnecessary floating point functions also increases the execution speed.

4. Empirical tests and future studies

This session first performs an empirical test and then discusses future studies of the proposed pricing model.

4.1 Empirical test

This study uses tick price data for the period from 03/01/2001 to 17/12/2003 to verify the feasibility of using the proposed computational methods to price TAIEX options using real payoff distributions. There were 270 data recorded for each sampling day, and given the sample data set contained 216,810 entries. Data for the period 03/01/2001 to 31/12/2002 were adopted as the initial distribution map, and pricing errors in high frequency transactions were verified on the last trading day of each month during 2003. The trading hours of the TAIEX run from 9:00 to 13:00. The final settlement price was taken to be the opening price of the final settlement day. The verification procedure is presented below:

Step 1: Generate the initial distribution map.

Filter out incorrect and duplicated data in the database, generate the distribution map using the *MakeRealDistribution* algorithm, and store it in a database table *DistributionMap (Transaction_date, Time_Ticks, Return_Rate)* that gives market price data on a per-minute basis between 03/01/2001 and 31/12/2002. Because the trading hours are 9:00 to 13:30, the first minute (9:01) is taken as *Time_Ticks* = 1 while the last (13:30) is *Time_Ticks* = 270. The *Return_Rate* $R_{i,j}$ equals the tick price of the TAIEX divided by the opening price for the following day:

$$R_{i,j} = \frac{X_{i+1,1}}{X_{i,j}}$$

Step 2: Determine the option price.

This study uses an out-of-sample strategy to verify the pricing performance. The nearest three in-the-money and out-of-the-money call/put option prices were then calculated and priced

using the *GetOptionPrice* algorithm for every time tick. The same option prices were also calculated using the BS model as a comparison. The riskless interest rate was the monthly fixed deposit interest rate used by the Central Bank of Taiwan.

Step 3: Estimate the pricing efficiency.

The option price is the expected value of $S_t > K$ for a call option, or $S_t < K$ for a put option at maturity. Restated, for an ideal call price $C = MAX(S_t - K, 0)$, the put price should be $P = MAX(K - S_t, 0)$. Consequently, if an individual spends C dollars to purchase a call option, they should obtain C dollars by holding the option until maturity. The returning ratios Rc and Rp were calculated for each option price to determine the pricing efficiency where the ideal value is 1.0:

$$Rc = \frac{\sum \max(S_t - K, 0)}{\sum C} \text{ for call options, and } Rp = \frac{\sum \max(K - S_t, 0)}{\sum P} \text{ for put options.}$$
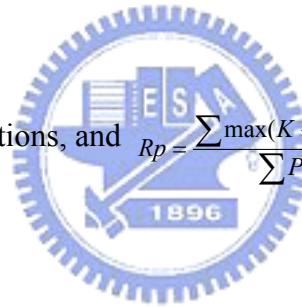
Table 4 lists the final results. According to the empirical test, the computational method outperforms the traditional BS model in pricing performance.

Table 4. Pricing error.

	Computational Method	Black-Scholes Method
Call Option, Rc	0.9290	0.9037
Pricing Error	7.10%	9.63%
Put Option, Rp	0.9874	0.9081
Pricing Error	1.26%	9.19%

Besides the pricing performance test, the execution speed was tested using Microsoft Visual FoxPro. The computational option model examined in this study is sufficiently efficient to price

1,000 option prices in 16 seconds (approximately 0.02 seconds each) where the distribution map contains 216,810 sample data, and is run on a 1GB RAM Intel Pentium4 2.6 GHz CPU personal computer system. All analytical results indicate that this computational method provides good pricing performance and efficient execution speeds when run on modern personal computer systems.

4.2 Future studies

Theoretically, pricing options must apply risk-neutral concepts in order to obtain unbiased prices. However, there always exist the needs to price options by user's personal speculations. This session briefly discusses future studies about determining option pricing via both risk-neutral and none risk-neutral view points.

Most option pricing theories suggest the use of risk-neutral distributions to price options. The reason is simple because those models assume that the underlying asset's price is unpredictable. In another words, if everyone can predict the future price, there will be no one lost money. Because the future price is unpredictable, most pricing models describe the underlying asset's behavior as "Random Walk". According to random walk hypothesis, the future price of an underlying asset is independent to its historical price while the future price is dependent only to the current price. For example, if one person tosses a fair coin that he wins 10 dollars if upper side appears or losses 10 dollars if opposite side occurs. If that person owns 100 dollars, the next state he will own either 90 dollars or 110 dollars. That is, the future is unpredictable but the next state is dependant to the current state. For that reason, traditional option pricing theories suggest the use of risk-neutral distributions for option pricing. However, it also lefts a mysterious question: which volatility is correct? Even most option pricing models apply risk-neutral distributions; however, investors have to decide the right

volatility value used for pricing. Thus, investors are bewildered of choosing the correct volatility value for risk-neutral pricing while it is also difficult to select the "right" volatility because it also varies consistently as illustrated in Figure 10.
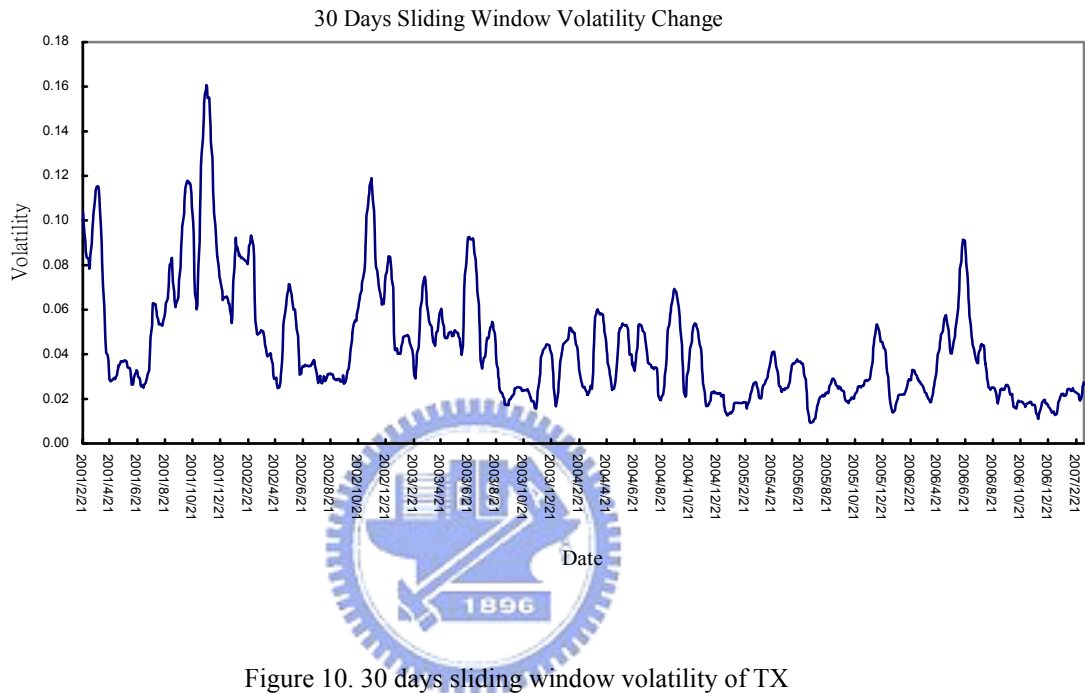

Figure 10. 30 days sliding window volatility of TX

In order to obtain the right volatility used in option pricing models, the most famous methodologies may be ARCH [63] or GARCH [64] approaches. Those methodologies have been proved that they can predict future volatility with an acceptable level [40][65]. For future studies on the proposed computational model, it may combine the use of ARCH or GARCH models to predict future volatilities in order to obtain more accurate prices. Also, artificial intelligence (AI) related models may be applied to forecast future volatilities for obtaining more precise prices with the proposed model.

Additionally, the proposed model uses actual distributions instead of mathematical distributions. Thus, it is also feasible to use "proper" periods of sampling data to dynamically construct the desired distribution maps. With this approach, it is possible to use pattern

matching techniques to determine which period of past samples best fits current conditions. And then, uses only properly fitted samples to build up the distribution maps for calculating. To sum up, it is possible to risk-neutrally optimize the proposed model by optimizing forecasted volatilities or sampling periods.

For example, if current volatility change looks like the blue area illustrated in Figure 11, users may use pattern matching technique or other methodologies to determine the most similar areas in the past (as in red periods). Using only similar periods as the sampling data, the SQL may be look like:

SELECT * FROM DistributionMap

WHERE DistributionMap.SamplingDate IN RedPeriod AND DistributionMap.DayCount =
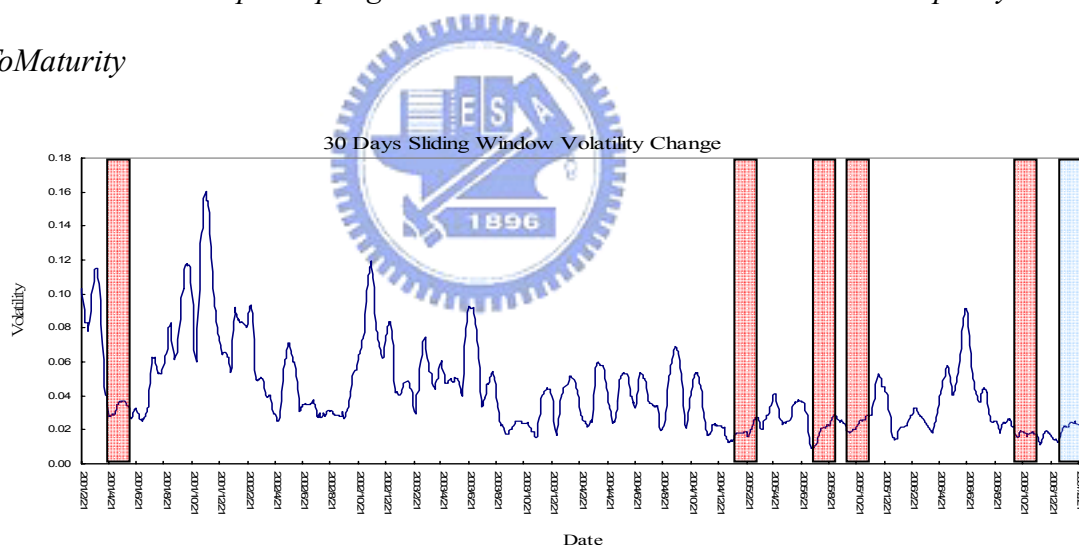
DaysToMaturity

Figure 11. Selecting proper sampling period according to volatility change

Although it is vital for theoretical option pricing models to obey risk-neutral perspectives, it is also equally important for individual investor to have his personal view to the future behavior of the underlying asset. However, if the investor has his own perspective, the derived option prices will not behave risk-neutrally. That is, the put-call parity will be violated if the investor applies his personal views. Figure 12 demonstrates the TX index values from 1/1/2001 to 9/3/2007. It is also clear that the underlying asset's price behaves differently from volatility
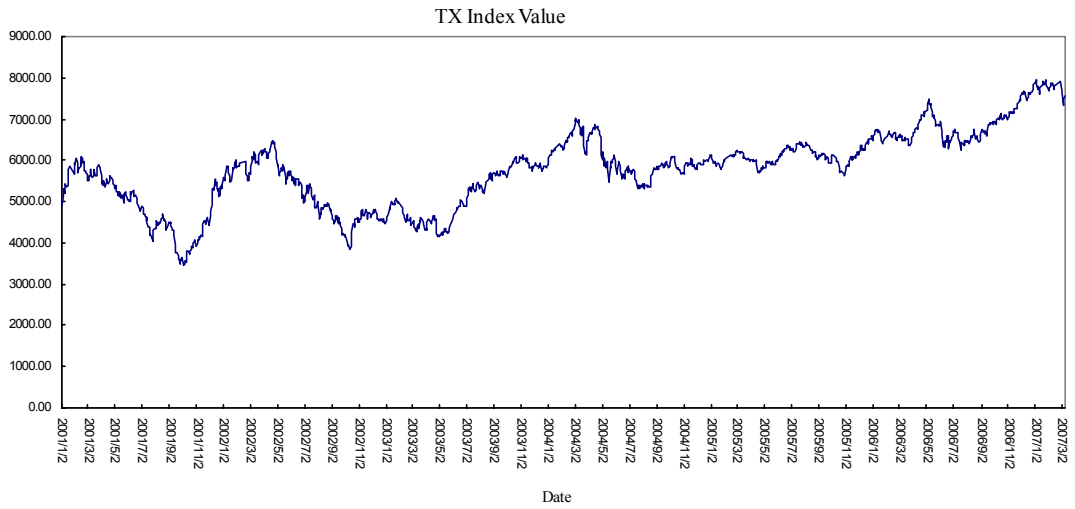
values.

Figure 12. TX index values

If an investor has his personal viewpoints to the future, it is worth to use pattern matching techniques to decide which interval of the historical period may reflect the future prices. And then, applies the selected samples as the distribution map and turn off the weighting factor rotating mechanism to price the options. However, this methodology cannot provide risk-neutral prices.
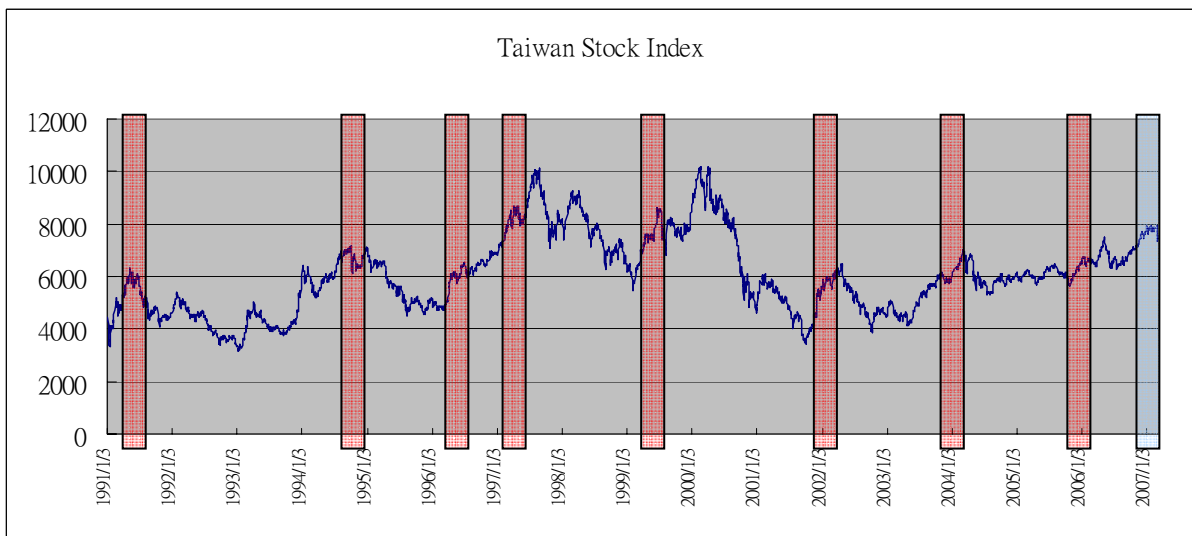

Figure 13. Selecting proper sampling period according to current pricesc

For example, current price is marked in blue color as illustrated in Figure 13, users may use pattern matching techniques to determine the most similar periods of the past as the desired distribution map. The SQL command may look like this:

SELECT * FROM DistributionMap

WHERE DistributionMap.SamplingDate IN RedPeriod AND DistributionMap.DayCount = DaysToMaturity

It is also worth mention that applying piecewise linear regression (PLR) methodologies may be applicable for pattern matching. Because PLR uses limited regression lines to describe a certain system characteristics, it is ideal to match different patterns by minimizing mean square errors between different set of PLRs.

Here is a simple example to demonstrate the applicable usage of PLR. The observations on actual payoff distributions have been discussed in the previous paragraphs. However, there still lacks detailed analysis. In order to provide in-depth information about the actual distributions, this study applies piecewise linear regression to generalize discrete distributions gathered from historical samples.

When observing actual distributions, the distribution is displayed with X-axis represents payoff rate and Y-axis indicates frequency ratio as illustrated in Figure 14.
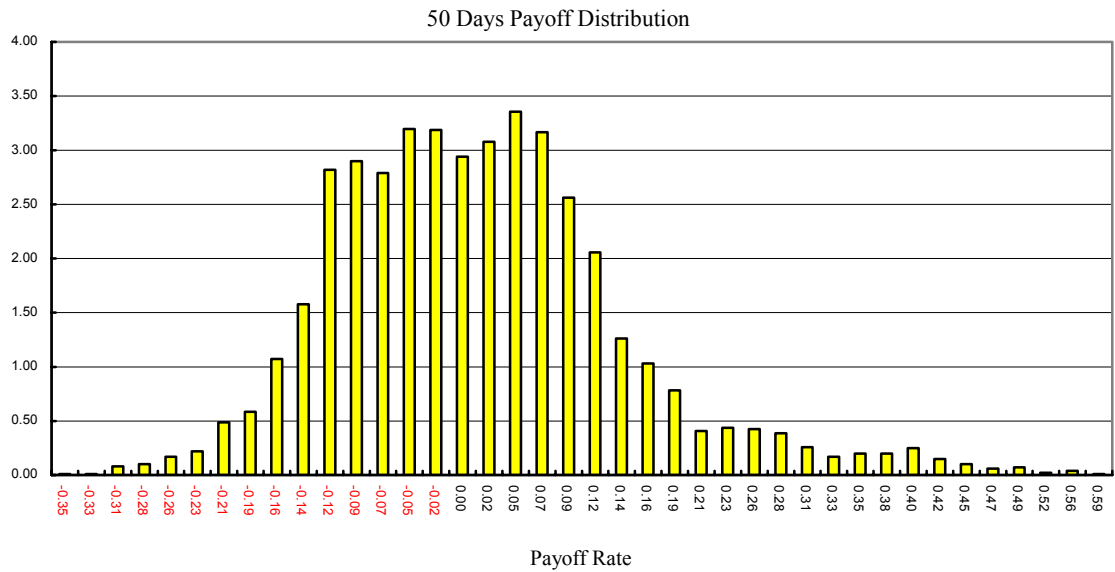
50 Days Payoff Distribution

Payoff Rate

Figure 14. 50 days payoff distribution with 40 intervals

In order to generate Figure 14, a researcher must first determine the intervals to cumulate sample counts. In this case, the interval counts equals 40. The reason to apply 40 intervals is because this study first testified several interval counts and then chose a smoother looking one. However, this method can only illustrate a rough shape of a certain distribution. The subtle behavior of such distribution is enclosed in histograms. For example, the same distribution map of Figure 14 will become into Figure 15 if applying 120 intervals.
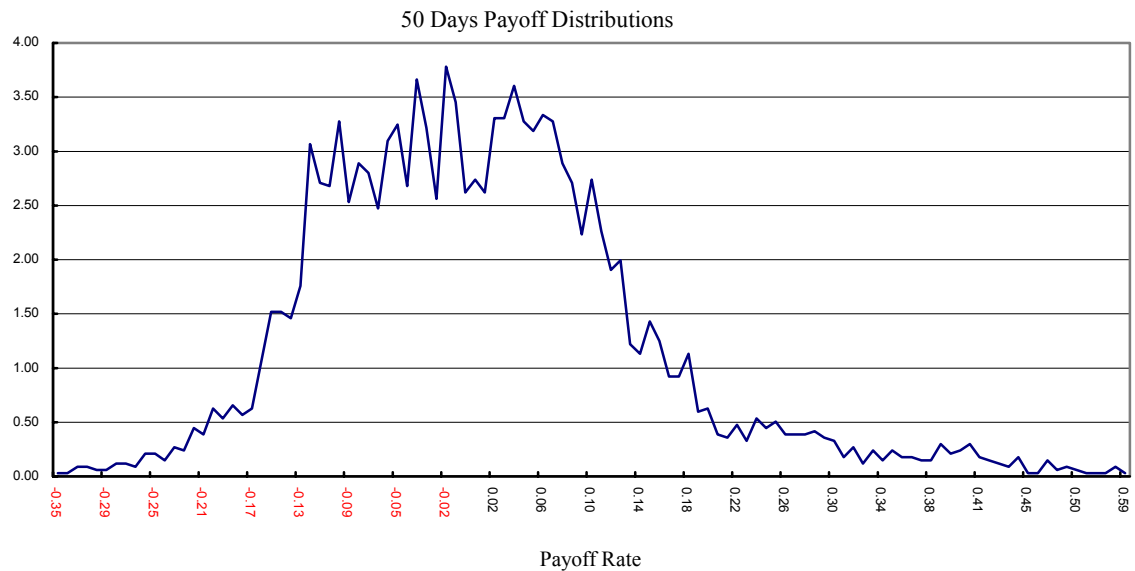
50 Days Payoff Distributions

Payoff Rate

Figure 15. 50 days payoff distribution with 120 intervals

By observing Figure 15, it is obvious that it requires applying certain methodologies to describe a distribution in a more meaningful way. Among various possible methodologies, this study demonstrates the use of PLR methodology. The term "piecewise" is to propose different lines over different intervals of a system with scattered (or polynomial) set of X-Y data. For a piecewise system, using linear regression lines for representing different intervals is called piecewise linear regression as illustrated in Figure 16. Although it is sophisticated to optimize the regression line (for example, Tanaka [59][60] and improved by Yu [61][62]), this uses simple piecewise linear regression (without break-point optimization) for a preliminary analysis.
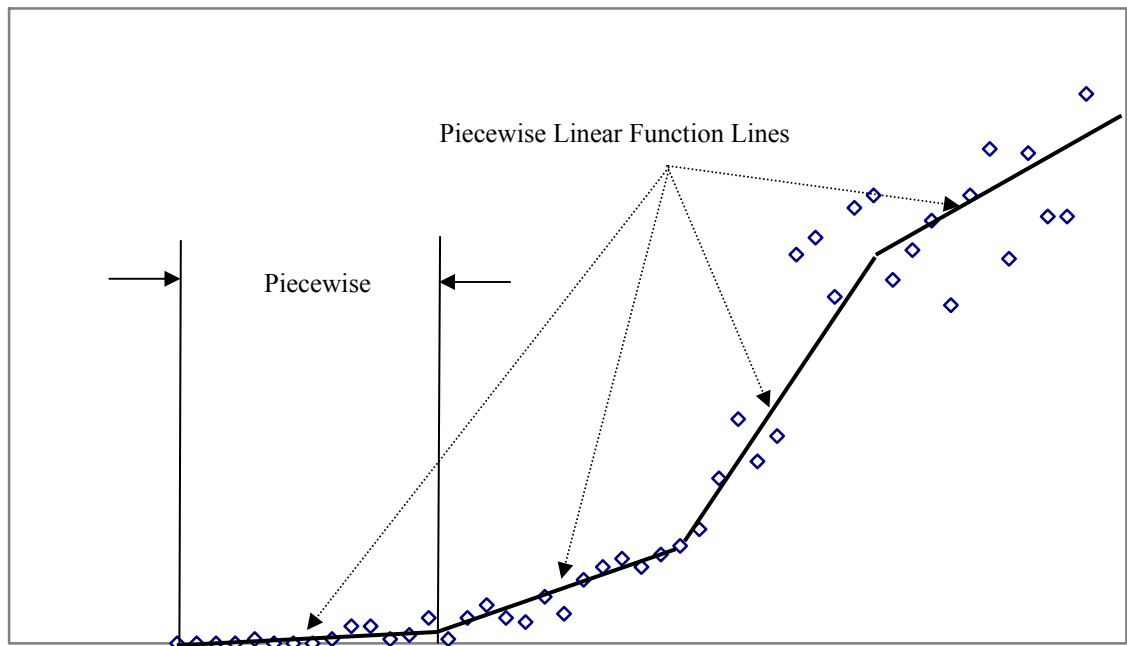
Figure 16. Piecewise linear function concept

Here is the question: what shape does all payoff distribution change related to days to maturity? In order to observe all distribution maps in the same scale, the payoff rates are transferred into counts of standard deviations: $R' = \dfrac{R}{Std(\Omega)}$ where $\Omega \in \{\Omega_1, \Omega_2, \dots \Omega_n\}$.

Note that the standard deviation of every transformed distribution becomes 1.0 so that it will be easier to observe the frequency peak value of each distribution. Now, it is obvious that if all the distributions have the same standard deviation then the distributions with higher peak value will be narrow and sharp around their mean value.

The formulas used for linear regression is:

$Y' = a + bX$

where Y' represents the estimated Y value (frequency count), X is deviation count, a is intercept and b is slope.

Then, $\quad b = \dfrac{\sum (X - \overline{X})(Y - \overline{Y})}{\sum (X - \overline{X})^2}$

$a = \overline{Y} - b\overline{X}$

This study uses 17 piecewise linear regression lines to approximate the actual distribution. That is, each line contains 8 intervals. A sample piecewise linear regression approximation is listed in Figure 15.
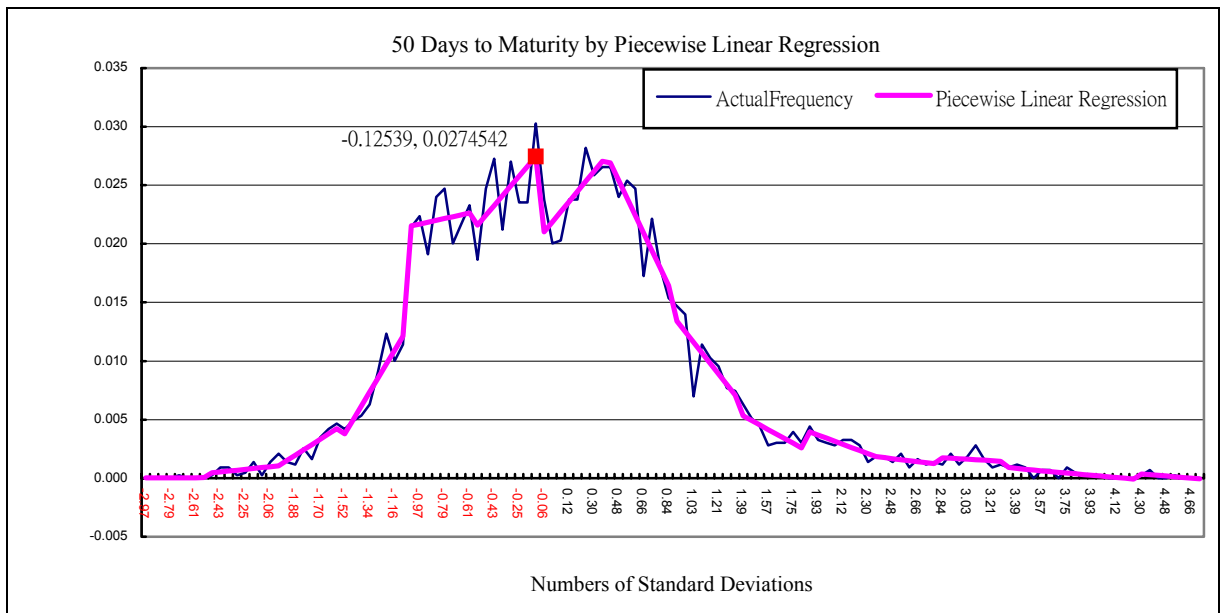

Figure 17. A piecewise linear regression sample

In Figure 17, the frequency of $\heartsuit_{50}$ peaks at (-0.12539, 0.0274542). Collects every peak value of each day to maturity distribution map, the results can be summarized in Figure 16. Note that the number pairs in Figure 16 denote (Days to Maturity, Highest Frequency). For example, *(4,0.04462502)* means that the highest frequency is 0.04462502 in the payoff distribution of 4 days to maturity ($\heartsuit_4$).
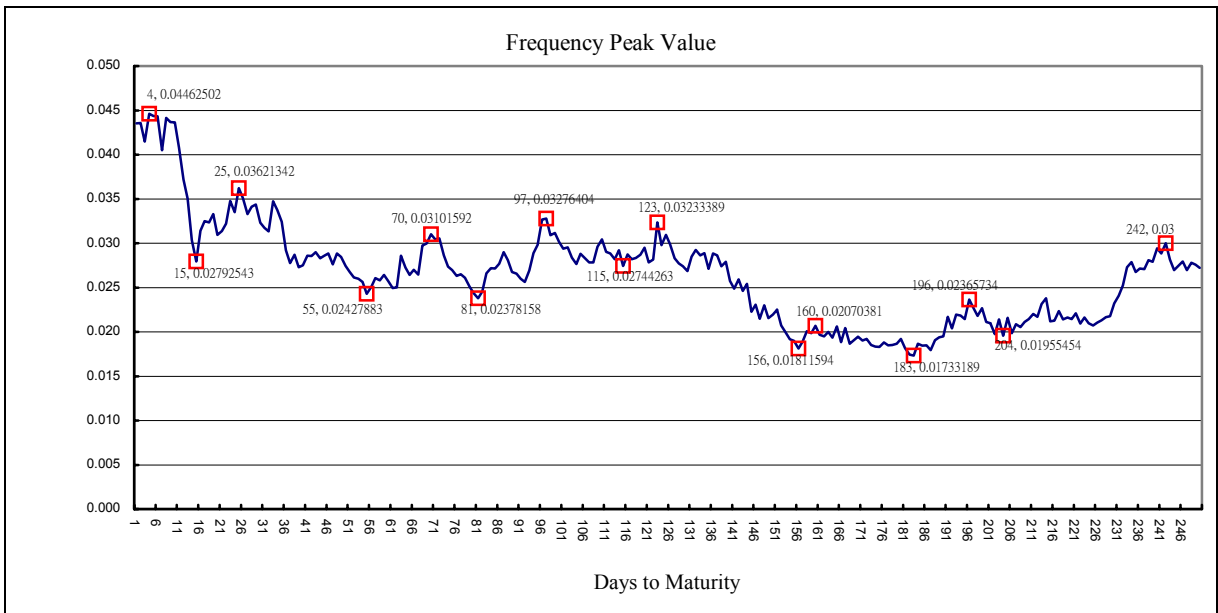
Figure 18. Frequency peak values of different days to maturity

Observing Figure 18, it is obvious that the frequency peak value of actual distributions varies with time to maturity. However, in the real world (at least in Taiwan stock market), short term options always have larger transaction amount. Long term options with days to maturity less than one month generally have less liquidity. For that reason, investors tend to apply short term options (less than one month) for hedging purpose. According to Figure 16, it is obvious that the frequency peak value changes according to days to maturity. A distribution with standard deviation equals 1.0 with higher peak values means that such distribution is sharp around central points. The highest peak occurs around 4 to 5 days to maturity, that is, the final week is most important. The lowest peak occurs around 15 days or half a month before maturity. Another interesting finding is that the actual payoff distribution is flattest at half a year to maturity (183 days).

5. Conclusions

It is important to obtain fair option prices because it is useful to deliver a theoretical fair price of a certain option to provide traders with a rational basis that can be used to judge whether actual option prices in the marketplace are reasonable, and to help investors determine where to place bids and offers. Furthermore, cross field application such as Real Option Analysis must apply option pricing models. As a result, almost all expected value related applications will benefit from option pricing models.

Most modern option pricing models apply mathematical distributions to approximate underlying asset behavior and attempt to calculate the desired option price using close form formulas. However, limitations on current option pricing models also decline the use of mathematical option pricing models. For example, the limitations may be a) too many assumptions, b) mismatch between real payoff distributions and mathematical distributions, and c) lack of flexibility & limited application ranges.

In order to construct a new option pricing model, this study first observes actual payoff distributions from the real world. The empirical evidence based on observation of the actual payoff distribution suggests that the actual distribution of a stock index is time variant and cannot be described using mathematical distributions, meaning the approach of most options pricing models is ineffective. To optimize the pricing performance, this study first introduces a computational model for pricing European options via time variant distributions and then demonstrates its practical feasibility using actual payoff distributions.

This study solves two key issues in applying user defined distribution to options pricing. First, this study uses weighting factors to adjust the mean value of a none-risk-neutral distribution to zero while maintaining its distribution characteristics in accordance with the put-call parity rule.

Second, this study scales the distribution to adjust its standard deviation to meet the needs associated with applying dynamic volatility to practical problems. Solving these two issues makes this computational model highly suitable for cross field applications where mathematical distribution cannot be used to obtain feasible solutions, particularly for situations involving time variant distributions.

Although the proposed computational method is practical for real world application, room still exists for improvement. First, the weighting factor rotating method used to adjust the value of the distribution means can be enhanced. This study assumes linearly changing weighting factors. Nonlinear modification methodologies require further study. Second, this study uses a simple method based on adjusting standard deviation that may not be able to deal with complex applications. Third, the computational method must be simplified before it can be applied to execution speed critical applications.

Reference

[1] E.K. Clemons (1991) Evaluating strategic investment in information system, Communications of the ACM 34(1) : 22-36.

[2] B.L. Dos Santos (1991) Justifying Investment in new information technologies, Journal of management information systems 7(4) : 71-89.

[3] F. Black and M. Scholes (1973) The pricing of options and corporate liabilities. Journal of Political Economy 81(3): 637–59.

[4] M. Benaroch (2002) Managing information technology investment risk: a real options perspective, Journal of management information systems 19(2) : 43 - 84.

[5] W.K. Chen (2001) Options: Theory, Practice and Application, Best-wise corp., 383-401

[6] J.O. Katz (2005) Advanced Option Pricing Models, Mc Graw Hill Corp., 52-56

[7] Stoll, Hans R. (1969) The relationship between put and call option prices, Journal of Finance 23: 801-824.

[8] Cox J.C. and A. Ross (1976) The valuation of options for alternative stochastic processes. Journal of Financial Economics 3: 145-166

[9] M. Joshi (2003) The Concepts and Practice of Mathematical Finance, Cambridge publications, 17-18

[10] R.S. Johnson, C. Giaccotto (1995) Options and Futures, West publishing company, 207-208

[11] Whaley, E. Robert (1982) Valuation of American Call Options on Dividend-Paying Stocks: Empirical Tests. Journal of Financial Economics,10: 60-67

[12] R. Merton. (1973) Theory of rational option pricing, Bell Journal of Economics 4: 141-183.

[13] K. Amin, R. Jarrow (1992) Pricing Options on Risky Assets in a Stochastic Interest Rate Economy. Mathematical Finance 2: 217-237.

[14] D.S. Bates (1996) Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options, Review of Financial Studies 9: 69–107.

[15] D.S. Bates (1991) The Crash of '87: Was It Expected? The Evidence from Options Markets. The Journal of Finance 46: 1009-1044.

[16] Madan, B. Dilip, P. Carr and Eric C. Chang (1998) The Variance Gamma Process and Options Pricing. European Finance Review 2(1): 79-105.

[17] M. Rubinstein (1994) Implied binomial trees. Journal of Finance 49 : 771--818.

[18] A.S Yacine, A.W. Lo (1996) Nonparametric estimation of state-price densities implicit in financial asset prices. Journal of Finance 52: 499-548.

[19] D. Bates (1996) Jumps and stochastic volatility: exchange rate processes implicit in Deutsch mark options. Review of financial studies 9 : 69--107.

[20] Scott, O. Louis (1997) Pricing Stock Options in a Jump-Diffusion Model with Stochastic Volatility and Interest Rates: Applications of Fourier Inversion Methods. Mathematical Finance 7: 345-358.

[21] B. Mandelbrot (1963) The Variation of Certain Speculative Prices, Journal of Business, 36, 394-419.

[22] L.C.G. Rogers (1997) Arbitrage from fractional Brownian motion, Mathematical Finance, Vol. 7, No. 1, 95–105.

[23] R.C. Blattberg, N.J. Gonedes (1974) A comparison of the stable and student distributions as statistical models for stock prices, Journal of Business 47, 244-80.

[24] B. Nielsen, O. E. Stephard. (2001) Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics (with discussion), J. Royal Stat. Soc. 63(2), ,pp.167-241

[25] P.K. Clark (1973) A subordinated stochastic process model with finite variance for speculative prices, Econometrica 41, pp.135-155.

[27] C.C. Heyde (2000) A risky asset model with strong dependence through fractal activity time, Journal of Applied Probability 36, pp.1234-1239.

[28] Hull, J., A. White (1987) The pricing of options on asset with stochastic volatilities, Journal of Finance 42, , pp.281-300.

[29] R. Engle (1995) ARCH, Selected Readings., Oxford University Press, Oxford, U.K.

[30] J.P. Fouque, , G. Papanicolaou, K. R. Sircar (2000) Derivatives in Financial Market with Stochastic Volatility, Cambridge University Press, Cambridge, U.K.

[31] D. Davydov, V. Linetsky (2001) The valuation and hedging of path-dependent options under the CEV process, Management Science 47, pp.949-956.

[32] S.L. Heston (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options, The review of Financial studies, 6:2, pp.327-343.

[33] D. Duffie, J. Pan, K. Singleton (2000) Transform analysis and asset pricing for affine jump-diffusions,. Ecomometrica 68(3), pp.1343-1376.

[34] S.G. Kou (2002) A jump-diffusion model for option pricing, Management Science 48:8, pp.1086-1101.

[35] E. Derman, I. Kani (1994) Riding on a smile, RISK 7, pp.32-39.

[36] B. Dupire (2000) Pricing with a smile, RISK 7, pp.18-20.

[37] J.M. Hutchinson, A.W. Lo, and T. Poggio (1994) A nonparametric approach to Pricing and Hedging Derivatives Securities Via Learning Networks. Journal of Finance 49: 851-889.

[38] A. Carverhill, T.H.F. Cheuk (2003) Alternative Neural Network Approach for Option Pricing and Hedging. Available at SSRN: http://ssrn.com/abstract=480562 or DOI: 10.2139/ssrn.480562

[39] G. Meissner, N. Kawano (2001) Capturing the volatility smile of options on high-tech stocks: A combined GARCH-Neural network approach, Journal of economics and finance 25: 276-292.

[40] H.R. Stoll, R.E. Whaley (1990).The dynamics of stock index and stock index futures returns.   Journal of Financial and Quantitative Analysis, 25, 441-468.

[41] K. Chan (1992). A further analysis of the lead-lag relationship between the cash market and stock index futures market. Review of Financial Studies, 5, 123-152.

[42] K.R. French, R. Roll (1986). Stock return variances: The arrival of information and the reaction of traders. Journal of Financial Economics, 17, 5-26.

[43] S. Ross (1989) Information and volatility: The no-arbitrage martingale approach to timing and resolution irrelevancy. Journal of Finance, 44, 1-17.

[44] I.G. Kawaller, P.D. Koch, T.W. Koch (1987) The Temporal Price Relationship between S&P 500 Futures and the S&P 500 index. Journal of Finance, 42:1309-1329.

[45] K. Chan, K.C. Chan, G.A. Karolyi (1991) Intraday volatility in the stock index and stock index futures markets.   Review of Financial Studies, 4, 657-684.

[46] G. Koutmos, M. Tucker (1996). Temporal relationships and dynamic interactions between spot and futures stock markets.Journal of Futures Markets, 16, 55-69.

[47] W.R. So, Y. Tse (2004) Price Discovery in the HANG SENG INDEX MARKETS: index, futures, and the tracker fund. The Journal of Futures Markets. Hoboken, Sep 2004.Vol.24, Iss. 9; pg. 887

[48] H. Latane, R. Rendleman (1976) Standard deviation of stock price ratios implied in

option prices. Journal of Finance 31,369-381.

[49] D.P. Chiras, S. Manaster (1978) The information content of option prices and a test of market efficiency. Journal of Financial Economics 6 (2/3), 213-234.

[50] L. Canina, S. Figlewski (1993). The informational content of implied volatility. Review of Financial Studies 6, 659-681.

[51] T. Day, C. Lewis (1992) Stock market volatility and the information content of stock index options. Journal of Econometrics 52, 267-287.

[52] C.G. Lamoureux, W. Lastrapes (1993) Forecasting stock return variance: towards understanding stochastic implied volatility. Review of Financial Studies 6, 293-326.

[53] C.R. Harvey, R.E. Whaley (1992) Market volatility prediction and the efficiency of the S&P 100 index option market, Journal of Financial Economics 31, 43-73.

[54] J. Fleming (1998) The quality of market volatility forecasts implied by S&P 100 index option prices   Journal of Empirical Finance vol. 5 pp. 317-345.

[55] Whaley, E. Robert (1982) Valuation of American Call Options on Dividend-Paying Stocks: Empirical Tests. Journal of Financial Economics,10: 29-58.

[56] S. Beckers (1981) Standard Deviations Implied in Option Prices as Predictors of Future Stock Price Variability, Journal of Banking and Finance, Vol. 5, No. 3, 1981, pp. 363-381.

[58] C.C. Sheng, H.Y. Chiu, A.P. Chen Using Computational Methodology to Price European Options with Actual Payoff Distributions, Soft Computing, Online published February 2007

[59] H. Tanaka, S. Uejima, K. Asai (1982) Linear regression analysis with fuzzy regression model, IEEE Trans. Systems Man Cybernet, 12, 903-907

[60] H. Tanaka, J. Watada (1988) Possibilistic linear systems and their application to the linear regression model, Fuzzy Sets and Systems, 27, 275-289

[61] J.R. Yu, G.H. Tzeng, and H.L. Li (1999) A General Piecewise Necessity Regression

Analysis Based on Linear Programming, Fuzzy Set and Systems, 105, 429-436

[62] J.R. Yu, G.H. Tzeng, and H.L. Li (2001) General Fuzzy Piecewise Regression Analysis with Automatic Change-Point Detection, Fuzzy Set and Systems, 119, 247-257

[63] R.F. Engle (1982) Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation, Econometrica, 50, 987-1007

[64] T. Bollerslev (1986) Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 31, 307-327.

[65] I.G. Kawaller, P.D. Koch, and T.W. Koch (1987) The Temporal Price Relationship Between S&P 500 Futures and the S&P 500 Index," Journal of Finance, 42, 1309-1329