INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
ON MULTINOMIAL MODELS OF SOME FINANCIAL INSTRUMENTS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Jianping Zhou, PH.D.

* * * * *

The Ohio State University

1996

Dissertation Committee:

Bostwick Wyman
Lijia Guo
Zhiwu Chen

Approved by

Bostwick Wyman

Adviser
The Department of Mathematics
To My Family
Acknowledgements

I would like to express my appreciation to many people. Yiming Xiao, Professor Peter March and professor Neil Falkner for answering many stochastic process related questions. Professor Rene Stulz for his wonderful and inspiring lecture on advanced finance theory. My greatest appreciation goes to the following people: Professor Wyman for his integrity, principles, guidance, vision, encouragement and financial support over every stage of my dissertation, to Dr. Lijia Guo for her constant encouragement and numerous discussions and guidance on many topics of mathematics and finance, to Professor Zhiwu Chen for reviewing my research work and giving authoritative comments. Lastly, to my wife for her emotional, financial support and encouragement without which I would never have gone this far, and to my daughter for giving me joy and happiness which make all my efforts worthwhile.
VITA

September 9, 1962 ................................. Born - JinHua, China

1982 ...................................................... B.S., Applied Mathematics,
Hunan University,
ChangSha, China.

1982-85 .................................................. Consultant,
ChangZen Electronics Company.,
ZhunYi, China.

1988 ...................................................... M.S., Institute of Mathematics,
Academia Sinica, Beijing, China.

1988-1989 ............................................... Research Assistant,
University of Science and Technology
of China,
HeFei, China.

1996 ...................................................... M.A, Department of Economics,
The Ohio State University,
Columbus, Ohio.

1990-present .......................................... Teaching Assistant,
Department of Mathematics,
The Ohio State University,
Columbus, Ohio.

Fields of Study

Major Field: Mathematics
# Table of Contents

DEDICATON ................................................................. ii

ACKNOWLEDGEMENTS ....................................................... iii

VITA ....................................................................................... iv

CHAPTER 

I  INTRODUCTION ................................................................. 1

1.1 Binomial Models of Some Financial Instruments ......................... 1
  1.1.1 Binomial Option Pricing Models ........................................ 1
  1.1.2 Ho-Lee’s Binomial Model of Term Structure of Interest Rates .... 7

1.2 Issues Investigated in this dissertation ........................................ 10
  1.2.1 Optimal Discretizations of Random Variables and Multinomial Option Pricing Formulas .................. 10
  1.2.2 Multinomial Models of Term Structure of Interest Rates ........ 12

1.3 Historical Development of Relevant Research ............................... 13

II  OPTIMAL DISCRETIZATION OF RANDOM VARIABLES AND OPTION PRICING FORMULAS ......... 17

2.1 Introduction ........................................................................ 17

2.2 Optimal Finite State Approximations to Random Variables with Infinite Many States ..................... 19

2.3 Optimal N-nomial Discretizations of Wiener Processes ................ 30

2.4 Optimal N-nomial Discretizations of Poisson Counting Processes ... 36

2.5 Stock Price Dynamics and Its Discrete Approximations ................. 44
  2.5.1 Continuous Model ....................................................... 44
  2.5.2 Discontinuous Model ................................................... 45
2.6 Two Limit Theorems .............................................................. 50
2.7 Implementation Procedures .................................................. 56
  2.7.1 Continuous Case ................................................................. 56
  2.7.2 Discontinuous Case ............................................................ 57
2.8 Conclusion ............................................................................. 60

III ON TERM STRUCTURE OF INTEREST RATES: A MULTINOMIAL
  GENERALIZATION OF HO-LEE MODEL ........................................ 61
  3.1 Introduction ........................................................................ 61
  3.2 Non-Arbitrage $N$-Nomial Bond Vector Movements ............... 64
    3.2.1 The Basic Assumptions .................................................. 64
    3.2.2 The $N$-Nomial Lattice .................................................... 65
    3.2.3 The Non-Arbitrage Condition .......................................... 66
  3.3 The Path-Independent $N$-nomial Bond Vector Movements ....... 69
    3.3.1 The Path Independence Condition .................................. 69
    3.3.2 The "Perturbation Functions" .......................................... 70
    3.3.3 Solutions of the "Perturbation Functions" ......................... 72
  3.4 The $N$-nomial Model .......................................................... 75
  3.5 Implications of the $N$-nomial Ho-Lee Model ......................... 77
    3.5.1 The Short Rate ............................................................... 77
    3.5.2 Local Expectations Hypothesis and the Term Premium ........ 79
  3.6 Pricing of Interest Contingent Claims .................................... 79
  3.7 Continuous Time Limit of the $N$-nomial Ho-Lee Model .......... 82
  3.8 Conclusion ........................................................................... 86

IV A NEW TERM STRUCTURE MODEL ............................................ 88
  4.1 Introduction ........................................................................ 88
  4.2 Construction of a Non-arbitrage Term Structure Vector Tree .... 89
  4.3 Constraints on the Parameters .............................................. 93
  4.4 Implementation Procedure .................................................. 96
  4.5 Conclusion ........................................................................... 97

V CONCLUSIONS ........................................................................... 99
  5.1 Summary ............................................................................. 99
  5.2 Future Research Topics ....................................................... 101

BIBLIOGRAPHY ............................................................................ 103
1.1 Binomial Models of Some Financial Instruments

In this section, we will first describe some classical binomial models of option pricing formulas and their continuous time limit, then we will describe Ho-Lee's binomial model of term structure of interest rates. These binomial models and their continuous limits provide foundations and motivations to our work in later chapters.

1.1.1 Binomial Option Pricing Models

The binomial option pricing formula which was originated by Cox, Ross and Rubinstein [23] and the continuous-time Black-Scholes formula [9] are two pioneering results in the option pricing theory. Their importance lies not only in the fact that they were derived under arbitrage theory which is the foundation of modern finance, but also in the fact that they are extremely simple to implement. Much of later development of option pricing theory follows the same methodology.
An option is a security which gives its owner the right to trade in a fixed number of shares of a specified common stock at a fixed price at any time on or before a given date. The act of making this transaction is referred to as exercising the option. The fixed price is termed the striking price, and the given date is called the expiration date. A call option gives the right to buy the shares; while a put option gives the right to sell the shares. If an option can only be exercised at a specified date (the maturity date), it is called an European option; if it can be exercised any time before a specified date, then it is called an American option. Throughout this and later chapters, we will assume that the options we deal with are all European options.

The binomial option pricing is based on the following assumption on the stock price: it follows a multiplicative binomial process over discrete periods. The rate of return on the stock over each period can have two possible values: $u - 1$ with probability $q$, or $d - 1$ with probability $1 - q$. Thus if the the current stock price is $S$, the stock price at the end of the period will be either $uS$ or $dS$. We can represent this movement with the following diagram:

\[
\begin{align*}
S \begin{cases} 
  uS & \text{with probability } q \\
  dS & \text{with probability } 1 - q 
\end{cases}
\end{align*}
\]

We also assume that the interest rate is a constant. Individuals may borrow or lend as much as they wish at this rate. To focus on the issues, we will assume that there are no taxes, transaction costs, or margin requirements. Hence, individuals are allowed to sell short any security and receive full use of the proceeds. Let $r$ denote one plus the riskless interest rate over one period, we require that $u > r > d$. If these inequalities
did not hold, there would be profitable riskless arbitrage opportunities involving only
the stock and riskless borrowing and lending.

To see how to value a call option, we start with the simplest situation: the expiration
date is just one period away. Let $C$ be the current value of the call, $C_u$ be its value
at the end of the period if the stock price goes to $uS$, and $C_d$ be its value at the
end of the period if the stock price goes to $dS$. Since there is now only one period
remaining in the life of the call, we know that the terms of its contract and a rational
exercise policy implies that $C = \max[0, dS - K]$. Therefore,

$$
C_u = \max[0, uS - k] \text{ with probability } q.
$$

$$
C_d = \max[0, dS - K] \text{ with probability } 1 - q.
$$

Suppose we form a portfolio containing $\Delta$ shares of stock and the dollar amount $B$
in the riskless bonds. This will cost $\Delta S + B$. At the end of the period, the value of
this portfolio will be

$$
\Delta S + B = \left\{ \begin{array}{ll}
\Delta uS + rB & \text{with probability } q, \\
\Delta dS + rB & \text{with probability } 1 - q.
\end{array} \right.
$$

Since we can select $\Delta$ and $B$ in any way we wish, suppose we choose them to equate
the end-of-period values of the portfolio and the call for each possible outcome. This
requires that

$$
\Delta uS + rB = C_u
$$

$$
\Delta dS + rB = C_d
$$

Solving these equations, we find
\[ \Delta = \frac{C_u - C_d}{(u - d)S}, \quad B = \frac{uC_d - dC_u}{(u - d)r} \]  

(1.1)

With \( \Delta \) and \( B \) chosen this way, we call this the hedging portfolio.

If there is no riskless arbitrage opportunity, the current value of the call, \( C \), cannot be less than the current value of the hedging portfolio, \( \Delta S + B \). If it were, we could make a riskless profit with no net investment by buying the call and selling the portfolio. For European call, the price \( C \) cannot be worth more, since then we would have a riskless arbitrage opportunity by reversing our procedure and selling the call and buying the portfolio. Note that for American call, the person who bought the call we sold has the right to exercise it immediately, hence the above argument should be modified. But as we noted earlier, we are focusing on European options. Hence for an European call, the price \( C \) is given as follows:

\[
C = \Delta S + B \\
= \frac{C_u - C_d}{(u - d)} + \frac{uC_d - dC_u}{(u - d)r} \\
= [(\frac{r-d}{u-d})C_u + (\frac{u-r}{u-d})C_d]/r
\]

(1.2)

Eq.(1.2) can be simplified by defining

\[
p = \frac{r - d}{u - d} \quad \text{and} \quad 1 - p = \frac{u - r}{u - d}
\]

so that we can write

\[
C = [pC_u + (1 - p)C_d]/r
\]

(1.3)

When we consider a call with any \( n \) \((n > 2)\) periods remaining before its expiration date. Since the stock price in the next \( n \) periods form a recombining tree, a backward
induction (see Chapter 3, Lemma 3 and its proof) can be used to obtain the call price:

\[ C = \sum_{j=0}^{n} \left( \frac{n!}{j!(n-j)!} p^j (1 - p)^{n-j} \max[0, \{u^j d^{n-j} S - K\}] \right) / r^n \]  

(1.4)

Let \( a \) denote the minimum number of upward moves which the stock must make over the next \( n \) periods for the call to finish in-the-money. Thus \( a \) will be the smallest non-negative integer such that \( u^a d^{n-a} S > K \) from which we could write \( a \) as the smallest non-negative integer greater than \( \log(K/Sd^n)/\log(u/d) \).

For all \( j < a \),

\[ \max[0, u^a d^{n-a} S - K] = 0, \]

and for all \( j \geq a \),

\[ \max[0, ad^{n-a} S - K] = ad^{n-a} S - K. \]

Therefore (1.4) becomes

\[ C = \sum_{j=a}^{n} \left( \frac{n!}{j!(n-j)!} p^j (1 - p)^{n-j} \max[0, \{u^j d^{n-j} S - K\}] \right) / r^n \]  

(1.5)

By breaking up \( C \) into two terms, we can write

\[ C = \sum_{j=a}^{n} \left( \frac{n!}{j!(n-j)!} p^j (1 - p)^{n-j} \max[0, \{u^j d^{n-j} S - K\}] \right) / r^n \]  

(1.6)

Now, the later bracketed expression is the complementary binomial distribution function \( \Psi[a; n, p'] \). The first bracketed expression can also be interpreted as a complementary binomial distribution function \( \Psi[a; n, p'] \), where

\[ p' = (u/r)p \quad \text{and} \quad 1 - p' = (d/r)(1 - p). \]
$p'$ is a probability, since $0 < p' < 1$. Now we can write (1.5) as

**Binomial Option Pricing Formula:**

$$C = S\Psi[a; n, p'] - Kr^{-n}\Psi[a; n, p],$$

where

$$p = (r - d)/(u - d), \quad \text{and} \quad p' = (u/r)p,$$

(1.7)

$a$ = the smallest non-negative integer

greater than $log(K/Sd^n)/log(u/d)$

If $a > n$, $C = 0$.

By taking limit of (1.7) with $u = e^{\sigma \sqrt{t/n}}$ and $d = e^{-\sigma \sqrt{t/n}}$, we obtain the Black-Scholes formula

**Black-Scholes Option Pricing Formula:**

$$C = SN(x) - Ke^{-rt}N(x - \sigma \sqrt{t}),$$

(1.8)

where

$$x = \frac{\log(S/K) + (r + \sigma^2/2)t}{\sigma \sqrt{t}}.$$ 

Now we choose $u = u$ and $d = e^{\xi(t/n)}$ in (1.7) and defining

$$\Phi[x, y] = \sum_{i=x}^{\infty} \frac{e^{-y} y^i}{i!},$$
then the limit option pricing formula corresponds to a continuous-time jump process
formula,

\[
C = S\Phi[x, y] - Ke^{-rT}\Phi[x, y/u],
\]

where

\[
y = (\log r - \xi)ut/(u - 1),
\]

and

\[
x = \text{the smallest non-negative integer greater than } (\log(K/S) - \xi t)/\log u.
\]

1.1.2 Ho-Lee's Binomial Model of Term Structure of Interest Rates

Ho-Lee [42] originated a binomial term structure model. It is based on the following assumptions:

(A1) markets are frictionless. There are no taxes and no transaction costs, and all securities are perfectly divisible.

(A2) The market clears at discrete points in time, which are separated in regular intervals. Each period is taken as a unit of time for simplicity. A discount bond of maturity \( T \) is defined to be a bond that pays $1 at the end of the \( T \)-th period, with no other payments to its holder.
(A3) The bond market is complete. There exists a discount bond for each maturity \( n \) \((n = 0, 1, 2, \cdots)\).

(A4) At each time there are finite number of states of the world. For state \( i \), the equilibrium price of the discount bond of maturity \( T \) is denoted by \( P_i^{(n)}(T) \). Note that \( P_i^{(n)}(\cdot) \) is a function that relates the price of a discount bond to its maturity. This function is called the \textit{discount function}. The discount function completely describes the term structure of interest rates of the \( i \)th state at time \( n \).

The discount function \( P_i^{(n)}(\cdot) \) must satisfy several conditions. It must be positive since the function represents assets' values, and the following two equations hold:

\[
P_i^{(n)}(0) = 1 \quad \text{for all } i, n
\]  

(1.10)

and

\[
\lim_{T \to \infty} P_i^{(n)}(T) = 0 \quad \text{for all } i, n.
\]  

(1.11)

With the above assumptions, the binomial movements of the term structures are described by two \textit{perturbation functions}, \( h(T) \) and \( h^*(T) \) and the following equations:

\[
P_i^{(n+1)}(T) = \frac{P_i^{(n)}(T + 1)}{P_i^{(n)}(1)} h(T) \quad \text{in the upstate,}
\]

(1.12)

and

\[
P_i^{(n+1)}(T) = \frac{P_i^{(n)}(T + 1)}{P_i^{(n)}(1)} h^*(T) \quad \text{in the downstate,}
\]

(1.13)

The perturbation functions specify the deviations of the discount functions from the implied forward function. For \( T > 0 \), \( h(T) \) is always greater than 1 and \( h^*(T) \) is
always less than 1. But for $T = 0$, equation (1.10), (1.11) and (1.13) imply that

$$h(0) = h^*(0) = 1 \quad (1.14)$$

Under path-independent conditions (detailed definition see chapter 4), Ho and Lee in [42] solved the perturbation functions explicitly with two parameters, i.e.,

$$h(T) = \frac{1}{\pi + (1 - \pi)\delta^T} \quad (1.15)$$

$$h^*(T) = \frac{\delta^T}{\pi + (1 - \pi)\delta^T} \quad (1.16)$$

Hence the discount functions can be found explicitly using the two variables $\pi$ and $\delta$.

At any time and state $(n, i)$, by constructing a risk free portfolio consisting of bond and the contingent claim on the bond, Ho and Lee [42] proved the following pricing formula of interest rate contingent claims:

**Proposition 1 (Risk-Neutral Pricing formula):** Consider any interest rate contingent claims $C(n, i)$ that can be bought and sold in a frictionless market environment described by assumption (A1)-(A4). If no arbitrage profit is to be realized in holding any portfolio of the contingent claim and the discount bonds, the following equations must hold:

$$C(n, i) = \pi\{C(n + 1, i + 1) + X(n + 1, i + 1)\} \quad (1.17)$$

$$+ (1 - \pi)\{C(n + 1, i) + X(n + 1, i)\}\right\}P_t^n(1)$$

where $X(n, i)$ is the payoff of the contingent claim at time $n$ and state $i$. 
1.2 Issues Investigated in this dissertation

In the previous section, we have introduced the classical binomial option pricing formulas. In Chapter 2, we will consider multinomial versions of the Black-Scholes formula and Merton's Jump diffusion option pricing formula (which is more general than formula (1.9)). In fact, such generalizations satisfy some optimality conditions. To do that, we first derive a general result about how to optimally discretize any infinite-many-state random variable, then apply it to normal distributions and Poisson distributions to obtain corresponding option pricing formula. These multinomial options pricing formulas are then shown to converge to Black-Scholes formula and Merton's formula. In Chapter 3, we generalize Ho-Lee's Binomial model of term structure of interest rates to multinomial models. Chapter 4 gives a new non-arbitrage binomial term structure model. Though the multinomial version has been worked out, it is not presented here to keep our focus.

1.2.1 Optimal Discretizations of Random Variables and Multinomial Option Pricing Formulas

Since continuous-time models of the stock price usually involves either Wiener processes or Poisson processes, so to get multinomial models of option pricing formulas with satisfy some optimality conditions, it is natural to consider constructing optimal finite states approximations to normal distributions and Poisson distributions, This leads to a more general question about how to optimally discretize any given random variable. One of our contributions in this dissertation is that we completely solved this
problem. We measure accuracy of an finite state approximation to a given random variable (typically with infinite state) in terms of moments, i.e., we define an optimal approximation to be such that it has the highest order of equal moments with the random variable to be approximated. This kind of approximation is justified by the well known fact that for any random variable which has analytic moment generating function (a series), its distribution is completely determined by its moments. Most random variables such as normally distributed ones and Poisson distributed ones all are completely determined by their moments. With the above definition, the optimal n-state approximation to a given random variable is defined to be the one which is uniquely determined by imposing as many constraints of equal moments to those of the given random variable as possible. Such constraints on moments result in a system of equations which is nonlinear in n state variables but linear in n probability variables. Such a system of equations is shown to have a unique solution and the solution is completely characterized by a polynomial of degree n whose coefficients are uniquely determined by some given moments.

By applying the above result to normal distributions and Poisson distributions, we get n-nomial approximations to the Black-Scholes option pricing formula and Merton's Jump diffusion option pricing formula. Such approximations are then shown to converge back to their continuous-time formulas respectively.
1.2.2 Multinomial Models of Term Structure of Interest Rates

In chapter 3, we present a $N$-nomial generalization of Ho-Lee’s model (we may call it $N$-nomial Ho-Lee model) under the framework, i.e., it is arbitrage free, path independent and consistent with the initial term structure of interest rates. It turns out that such a generalization has similar “perturbation functions” as in the original Ho-Lee model. This model is the first $N$-nomial model in literature that has path-independence property. We also obtained a non-arbitrage pricing formula for interest rate sensitive contingent claims in a general $N$-nomial environment. Applying this formula to “$N$-nomial Ho-Lee model” results in a path independent $N$-nomial pricing formula for the contingent claims, generalizing the binomial one of Ho-Lee’s. Continuous limit of the model is also considered. It turns out that the limit found by Heath-Jarrow-Merton [39] for the binomial Ho-Lee model is a special case of the limit we found.

In Chapter 4, we present a new non-arbitrage model of term structure of interest rates. Though it could be generalized to multinomial case, we only present the binomial one to focus on the distinct features of this model. The distinct feature of this model is that it has infinitely many parameters and the implied risk neutral probabilities are time dependent (time independent in the Ho-Lee model). Such a sequence of parameters provide flexibilities in controlling the term structure movements. For example, under some easily described conditions on the parameters, the interest rates are guaranteed to be positive for all time intervals.
1.3 Historical Development of Relevant Research

In general, the price of contingent claim on an asset which satisfies a stochastic differential equations can be described by ordinary partial differential equations with special boundary conditions. Black and Scholes [9] was the first in trying this approach. Specifically, with stock price movement described by a geometric Brownian motion with drift, then used continuous arbitrage argument to derive an ordinary differential equation for the option price. By successfully solving the partial differential equation, they obtained their famous Black-Scholes formula for option pricing for stocks. Cox, Ross and Rubinstein [23] and Rendleman and Bartter [66] established the binomial option pricing formula for stocks. If one use continuous time approach to value option, the corresponding ordinary differential equation obtained through arbitrage argument may be not solvable explicitly. Hence numerical approach have to be used. Typically one uses finite difference method (see Brennan-Shwartz [16], Courtadon [24] and Hull-White [48], Wiggins [73]). Other approaches are numerical integration (see Chen [18]), Monte-Carlo simulation as in Boyle [10] and lattice approach as in Boyle [11], Boyle-Evnine-Gibbs [12] and He [38], Diz [26], Heath-Jarrow-Morton [39] and Amin [3]). The papers on lattice approach are most relevant to the our research. Among those papers only the binomial discretization in Amin [3] satisfies optimality conditions on moments. Part of the results in this paper can be viewed as a very natural generalization of Amin [3] to the multi-nominal case. There is another very relevant and comparable paper Madan-Milne-Shefrin [60] which considered multinomial ap-
proximation to the Black-Scholes formula and Merton's formula, but their approach does not seem easily implementable, furthermore they did not consider optimality condition in any sense. The main results in chapter 2 provide some optimal ways for choosing parameters in discretizing various continuous model of financial instruments.

Many financial instruments in the fixed income market are interest rate contingent claims, e.g., interest rate options, callable bonds, floating rate notes, etc. So term structure modeling is essential in fixed income analysis. much academic literature has been devoted to this problem. One earlier attempt is that of Pye [64]. He assumed that the interest rates move according to a (Markov) transition probabilities matrix, and then used the expectation hypothesis to price the expected cash flow of the asset-in his case, a callable bond (Pye [65]). In later development, There equilibrium models have been the focus. Cox, Ingersoll and Ross (CIR) [21] assumed that the short rate follows a mean-reverting process. By further assuming that all interest rate contingent claims are priced contingent on only the short rate, using a continuous arbitrage argument they derived an equilibrium pricing model. Brennan and Schwartz [15] extended the CIR model to incorporate both short and long rates and studied the pricing of a broad range of contingent claims ([14], [17]). Other two-factor or even multifactor extensions of CIR models are given by Bakshi and Chen [5], Chen and Scott ([19], [20]), Kraus and Smith [58], and Longstaff and Schwartz [59]. Constantinides [25], Jamshidian [52] and Wang [72] are some other relevant papers along this line. Ho and Lee [42] pioneered another approach. They take the observed term
structure of interest rates as given and derive the feasible, arbitrage-free subsequent movements in the term structure of interest rates. Their binomial process permits the entire term structure to move up or down by a maturity-dependent constant at each date. Ho-Lee's assumptions of a constant-parameter process for the movements of the term structure have the rather unfortunate consequence of leading, alternatively, to extremely high rates of interest along one "branch" (of the binomial tree) and negative interest rates along the other. An extension of Ho-Lee's binomial model to trinomial model was carried out by Robert and Ehud [67]. Their empirical study shows that a trinomial, rather than a binomial, model of the term structure fits the data in a much more superior fashion. Results in Chapter 3 may be considered a further extension along this line. Robert-Ehud's trinomial model is not path independent which, among other things, causes the number of nodes in the corresponding tree to grow exponentially, resulting much high computing cost. In Chapter 3, we consider the most general form of multinomial model of term structure of interest rates which satisfies the path-independent condition. The results shows that a generalized Ho-Lee model has its multinomial version which has all features of the original Ho-Lee model. Due to the limitations of Ho-Lee model, Black, Derman and Toy (BDT)[7] proposed another binomial model of term structure of interest rates. Their model can fit not only the initial term structure but also the initial volatility curve. The drawback of their model is that no explicit formula is given, rather a procedure on how to fit the initial term structure and the initial volatility is provided. Implementing BDT model in most cases involves trial and error process which is extremely inefficient.
Jamshidian [53] provided a forward induction method to implement BDT model and other related models. His method proved to be effective and fast.

In Chapter 4, a new model of term structure of interest rates is proposed. This model has seems to has some good aspects of both HL model and BDT model. It has explicit formula for the term structure movements, yet its parameters can be controlled to yield nonnegative interest rates for all intervals.
CHAPTER II

OPTIMAL DISCRETIZATION OF RANDOM VARIABLES AND OPTION PRICING FORMULAS

2.1 Introduction

In this chapter, we first construct optimal finite state approximations to any given infinite state random variable. Each such approximation is shown to be completely described by a polynomial which in certain cases is independent of the variance of the random variable to be approximated. We apply the results obtained to construct optimal discrete time finite state approximations to typical random processes used in finance, i.e., Wiener process and Poisson process. As a result of such discretizations, we obtain easily implementable multinomial pricing formula for option prices under the Black-Scholes model [9] which is obtained under continuous asset return assumption and Merton's model [62], has discontinuous asset returns assumption. To justify such discretizations of stock returns, hence the option prices, we show that the limits of the option prices obtained are exactly Black-Scholes's formula and Merton's formula under Black-Scholes' model and Merton's model respectively. In the whole chapter,
an optimal n-state approximation to a given random variable is defined to be the one which is uniquely determined by imposing as many constraints of equal moments to those of the given random variable as possible. The existence and its descriptions of such optimal approximation to a given random variable are given in section 1, 2 and 3. The limit theorems given in section 6 are valid for a general class of discretizations of the processes involved and only require conditions on the first and second order moments, whereas the discrete time finite state approximations we constructed for some special random processes such as Wiener process and Poisson process satisfy conditions on higher moments. So practically, such kind of discretization procedure results in algorithms for computing option prices which could be considered as most efficient.

Optimality of approximations to option prices in any sense has not been considered before in the literature. For a discrete time and finite state approximation to make sense, one must at least have the continuous time infinite state model at hand. For the two models mentioned above, the option price formulas are already explicitly given. The Black-Scholes's formula is easy to implement itself. Any discrete approximation only serves us in that we could incorporate other instruments such as dividends into the process. For Merton's formula, due to its complexity, implementing it is not a trivial task. Boyle [10] briefly mentioned on how to use Monte Carlo stimulation to implement Merton's Model. Our approach results in a approximation to Merton's formula similar to the classical binomial option pricing formula given by Cox-Ross-Rubinstein [23] or Rendleman-Barter [66] to Black-Scholes's formula. There is one
difference though, we are not concerned with non arbitrage in one period, instead we are only concerned with approximating the continuous time result as fast as possible. As a result, even in the binomial case, the option pricing formula we obtained for the Black-Scholes' model is different in parameters from the classical one obtained through nonarbitrage approach. The difference can be explain as follows: The classical approach only imposes conditions on the first and second moments to let the non-arbitrage argument go through, while our approach imposes an additional condition on the third moment of the approximation. Both approach satisfy the conditions of our limit theorem so both converge to the Black-Scholes's formula.

2.2 Optimal Finite State Approximations to Random Variables with Infinite Many States

In this section, we consider the following problem: given an integer $n > 1$, find the "optimal" $n$-state random variable approximation of a given random variable with infinite many states. Of course, the meaning of "optimal" should be well explained. We choose to measure the accuracy of such an approximation by moments, i.e., the optimality of such an approximation means that the two random variables have equal moments to the highest possible order.
Let $F(x)$ be the distribution function of a random variable. For a given positive integer $m$, let

$$J_m = \int_{-\infty}^{\infty} \left( \sum_{k=0}^{m} t_k(x-a)^k \right)^2 dF(x) = \sum_{j=0}^{m} \sum_{k=0}^{m} t_k t_j m_{k+j}(a)$$

where $a$ is any constant real number and $m_{i}(a)$ is the $i$-th moment of the random variable centered at $a$.

**Lemma 1** $J_m$ is a positive definite quadratic form if one of the following conditions holds:

(i). the distribution has density function $p(x)$ which is positive and continuous at $m+1$ or more distinct points.

(ii). $F(x)$ has jumps at $m+1$ or more distinct points.

In particular, for a normal distribution or a poisson distribution, the corresponding $J_m$ is positive definite for any $m$.

**Proof:** Clearly $J_m$ is nonnegative for all $t_0, t_1, \ldots, t_m$. Let

$$h_m(t_0, t_1, \ldots, t_m, x) = \sum_{k=0}^{m} t_k(x-a)^k$$

then

$$J_m = \int_{-\infty}^{\infty} h_m^2(t_0, t_1, \ldots, t_m, x) dF(x)$$

Suppose that $p(x)$ is continuous at $x_0$ and $p(x_0) > 0$ or $F(x)$ has a jump at $x_0$, then if $J_m = 0$ would imply that $h_m(t_0, t_1, \ldots, t_m, x_0) = 0$. But we have $m+1$ such points
and $h_m(t_0, t_1, \ldots, t_m, x)$ has degree at most $m$, the polynomial must be identically zero, i.e., $t_0, t_1, \ldots, t_m$ must be all zero. So for $t_0, t_1, \ldots, t_m$ not all zero, $J_m$ must be positive definite. For normal and poisson distributions, the former has continuous positive density function and the later has a distribution function with infinite many jumps, hence the conclusion follows. Q.E.D.

In what follows, let $\xi$ be the target random variable to be approximated with moments denoted by $\{m_i\}_{i=0}^\infty$; and we always assume that $\xi$ satisfies one of the conditions in the above lemma. Let $\psi_{\delta,n}$ be a random variable with $n$ non-zero states $z_1, z_2, \ldots, z_n$ and their associated probabilities $p_1, p_2, \ldots, p_n$ and $\delta = 1$ or $0$ depending on whether zero is allowed as an extra state or not. If $\delta = 1$, then zero state probability is $p_0$. The moments of $\psi_{\delta,n}$ are denoted by $\{\mu_i\}_{i=0}^\infty$.

Let $\sigma_s, 1 \leq s \leq n$ be the $s$-th order elementary symmetric polynomials in $z_1, z_2, \cdots, z_n$, i.e.

$$\sigma_s = \sum_{i_1, i_2, \cdots, i_s} z_{i_1} z_{i_2} \cdots z_{i_s},$$

where the summation indices $i_1, i_2, \cdots, i_s$ are over mutually different integers in the range between 1 to $n$.

For $i \geq 0$, we define the following

$$u_{i,n} = [m_i, m_{i+1}, \cdots, m_{n+i-1}]',$
\[
M_{i,n} = \begin{pmatrix}
m_i & m_{i+1} & \cdots & m_{n+i-1} \\
m_{i+1} & m_{i+2} & \cdots & m_{i+n} \\
\vdots & \vdots & \ddots & \vdots \\
m_{i+n-1} & m_{i+n} & \cdots & m_{2n+i-2}
\end{pmatrix},
\]

\[
W_{i,n} = \begin{pmatrix}
z_1^i & z_1^{i+1} & \cdots & z_1^{i+n-1} \\
z_2^i & z_2^{i+1} & \cdots & z_2^{i+n-1} \\
\vdots & \vdots & \ddots & \vdots \\
z_n^i & z_n^{i+1} & \cdots & z_n^{i+n-1}
\end{pmatrix},
\]

\[
A = \begin{pmatrix}
0 & 0 & \cdots & 0 & -a_n \\
1 & 0 & \cdots & 0 & -a_{n-1} \\
0 & 1 & \cdots & 0 & -a_{n-2} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 & -a_1
\end{pmatrix},
\]

where \(a_i = (-1)^i \sigma_i\) for \(1 \leq i \leq n\).

**Theorem 1** For a given \(n\), if \(M_{1,n}\) is nonsingular, then there is a unique random variable \(\psi_{\delta,n}\) with \(n\) nonzero states \(z_i(1 \leq i \leq n)\) satisfying the following equations

\[
\mu_i = m_i \text{ for } 0 \leq i \leq 2n + \delta - 1; \quad (2.1)
\]

Furthermore the \(z_i\)'s are the \(n\) distinct nonzero (positive if \(M_{1,n}\) is positive definite) real roots of the following polynomial

\[
f_{\delta,n,\xi}(z) = z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_n,
\]

where

\[
[a_n, \cdots, a_2, a_1]' = -M_{\delta,n}^{-1} v_{n+\delta,n}, \quad (2.2)
\]

\[
[p_1, p_2, \cdots, p_n] = v_{\delta,n} W_{\delta,n}^{-1}, \quad (2.3)
\]

In any case, the \(p_i\)'s are positive, in particular, if \(\delta = 1\), then \(p_0 = 1 - \sum_{i=0}^{n} p_i > 0\).
In what follows, we prove two lemmas which are needed in proving Theorem 1.

**Lemma 2** Let \( z_i, \sigma_i, 1 \leq i \leq n, A \) and \( W_i(i \geq 0) \) be as defined previously, then we have the following identities

\[
W_{i+1,n}^{-1}W_{i+1,n} = A
\]

and

\[
W_{1,n}^{-1}(1, 1, \cdots, 1)' = (-\frac{a_{n-1}}{a_n}, -\frac{a_{n-2}}{a_n}, \cdots, -\frac{a_1}{a_n}, -\frac{1}{a_n})'.
\]

**Proof:** Since for each \( z_i \), we have

\[
z_t^n + a_1z_t^{n-1} + \cdots + a_{n-1}z_t + a_n = 0,
\]

multiplying both sides of (2.6) by \( z_i \), we get

\[
z_t^{n+i} + a_1z_t^{n+i-1} + \cdots + a_{n-1}z_t^{i+1} + a_nz_i = 0
\]

let \( C_j = [z_{i1}^1, z_{i2}^2, \cdots, z_{in}^n]' \) for any integer \( j \geq 0 \), then by the above equations, we have

\[
C_{n+i} + a_1C_{n+i-1} + \cdots + a_{n-1}C_{i+1} + a_nC_i = 0,
\]

and from which it follows that \( W_{i+1,n} = W_{i,n}A \) which is equivalent to (2.4). Clearly, (2.7) is true for any \( i \geq 0 \), letting \( i = 0 \), we get

\[
(1, 1, \cdots, 1)' = -\frac{1}{a_n}(C_n + a_1C_{n-1} + \cdots + a_{n-1}C_1).
\]

Let \( \Delta_1 = \det(W_{1,n}) \) and \( W_{1,n}^* = (c_{s,t})_{n \times n} \), be the companion matrix of \( W_{1,n} \), then

\[
W_{1,n}^{-1} = \Delta_1^{-1}W_{1,n}^*.
\]
but for $1 \leq s \leq n$, the sum of $s$-th row of $W_{1,n}$, is simply the determinant of the matrix obtained from $W_{1,n}$ by replacing $s$-th column by $(1, 1, \cdots, 1)$, hence, by (2.8), it is the product of $-\frac{a_{n-s}}{a_n}$ and $\Delta_1$, (2.5) then follows from (2.9). Q.E.D.

Lemma 3 For any $n \times n$ symmetric matrices $P$ (nonsingular) and $Q$, if $P$ or $Q$ is definite (positive or negative), then all eigenvalues of $P^{-1}Q$ are real and $P^{-1}Q$ is similar to a diagonal matrix. Furthermore, if $P$ and $Q$ are both positive definite or both negative definite then all eigenvalues of $P^{-1}Q$ are positive real numbers.

Proof: Let $\lambda$ be an eigenvalue of $P^{-1}Q$ and $v$ is one of its non-zero eigenvector, then

$$(\lambda E - P^{-1}Q)v = 0$$

which is equivalent to

$$\lambda P v = Q v.$$ 

Let $\overline{v}$ be the complex conjugate of $v$, multiplying both sides of the above equation by $\overline{v}^t$ yields

$$\lambda \overline{v}^t P v = \overline{v}^t Q v.$$ 

Clearly, both $\overline{v}^t P v$ and $\overline{v}^t Q v$ are real and either $\overline{v}^t P v$ or $\overline{v}^t Q v$ is nonzero by assumption, the above equation implies that $\lambda$ is real. If $P$ and $Q$ are both positive definite or both negative definite, then both $\overline{v}^t Q v$ and $\overline{v}^t P v$ are both positive or both negative real numbers, thus $\lambda$ must be a positive real number.
Now assume $P$ is definite, if $P^{-1}Q$ is not diagonalizable, then the Jordan form of $P^{-1}Q$ is not diagonal, hence there exist an eigenvalue $\lambda$ of $P^{-1}Q$ and two nonzero vectors $v_1$ and $v_2$ such that

$$P^{-1}Qv_1 = \lambda v_1$$

$$P^{-1}Qv_2 = v_1 + \lambda v_2,$$

or equivalently

$$(\lambda P - Q)v_1 = 0$$

$$(\lambda P - Q)v_2 = Pv_1. \tag{2.10}$$

Left multiplying (2.11) by $v_1'$ yields

$$v_1'(\lambda P - Q)v_2 = v_1'Pv_1.$$

Since $\lambda$ is real and $\lambda P - Q$ is real and symmetric, left hand side of the last equation is zero by (2.10), but the right hand side of the equation is nonzero by the definitiveness of $P$, this is a contradiction which shows that $P^{-1}Q$ must be diagonalizable. Finally, if $Q$ is definite, then the above shows $Q^{-1}P$ is diagonalizable, so is its inverse $P^{-1}Q$. Q.E.D.

**Proof of Theorem 2**: The proof proceeds as follows: we first show that under condition (2.1), the $z_i$'s are roots of the polynomial $f_{\delta,n,\xi}(z)$ and the probabilities are given by (2.3); we then prove that the polynomial given in the theorem does have $n$ distinct nonzero real roots and (2.3) gives a probability vector i.e., $q_i > 0$ for
1 \leq i \leq k \text{ and } p_0 > 0 \text{ if } \delta = 1, \text{ and the corresponding random variable with nonzero states } z_i, 1 \leq i \leq n \text{ satisfies equations (2.1).}

We first write equations in (2.1) with \( i \geq \delta \) into the following equivalent form

\[(p_1, p_2, \cdots, p_n) W_{i,n} = v_{i,n}', \text{ for } \delta \leq i \leq n + \delta,\]  

(2.12)

from which it follows

\[v_{i,n}' W_{i,n}^{-1} W_{i+1,n} = v_{i+1,n}', \text{ for } \delta \leq i \leq n + \delta - 1.\]  

(2.13)

By Lemma 2, the above equations become

\[v_{i,n}' A = v_{i+1,n}', \text{ for } \delta \leq i \leq n + \delta - 1.\]  

(2.14)

Rewriting it into the following form

\[M_{\delta,n} A = M_{\delta+1,n},\]  

(2.15)

and comparing the last column on both sides, we get (2.2). (2.3) follows from (2.12) with \( i = \delta \). It is easy to check that \( f_{\delta,n,\xi}(z) \) is exactly the characteristic polynomial of the matrix \( A \). By lemma 1, \( M_{2i,n} \) is positive definite for any \( i \geq 0 \), so one of \( P = M_{\delta,n} \) or \( Q = M_{\delta+1,n} \) is positive definite, then the nonsingularity of \( M_{1,n} \) ensures that both of them are nonsingular. By apply Lemma 3, we conclude that all roots of \( f_{\delta,n,\xi}(z) \) are nonzero real numbers and are positive if \( M_{1,n} \) is positive definite. We now show that \( f_{\delta,n,\xi}(z) \) does not have repeated roots. Again by lemma 3, \( A \) is diagonalizable, if \( \lambda \) is a multiple root of \( f_{\delta,n,\xi}(z) \), it implies that

\[\text{rank}(\lambda E - A) \leq n - 2.\]
From (2.15) and the assumption that \( M_{\delta,n} \) has full rank, we have

\[
\text{rank}(\lambda M_{\delta,n} - M_{\delta+1,n}) \leq n - 2.
\]

or

\[
\text{rank}(\lambda v_{\delta,n} - v_{\delta+1,n}, \lambda v_{\delta+1,n} - v_{\delta+2,n}, \ldots, \lambda v_{\delta+n-1,n} - v_{\delta+n,n}) \leq n - 2,
\]

This implies that the vectors \( \lambda v_{\delta+i,n} - v_{\delta+i+1,n}, 0 \leq i \leq n - 2 \) are linearly dependent over real numbers, i.e. there exist real numbers \( c_0, c_1, \ldots, c_{n-2} \), not all zero, such that

\[
\sum_{i=0}^{n-2} c_i (\lambda v_{\delta+i,n} - v_{\delta+i+1,n}) = 0,
\]

or

\[
\lambda c_0 v_{\delta,n} + \sum_{i=1}^{n-2} (\lambda c_i - c_{i-1}) v_{\delta+i} - c_{n-2} v_{n-1} = 0.
\]

Since \( v_{\delta+i,n}, 0 \leq i \leq n - 1 \) are just the column vectors of \( M_{\delta,n} \) which is of full rank, they must be linearly independent over real numbers, the last equation implies that

\[
\lambda c_0 = 0, \lambda c_i - c_{i-1} = 0, 1 \leq i \leq n - 2, -c_{n-2} = 0.
\]

Since \( \lambda \neq 0 \), we must have \( c_0 = c_1 = \cdots = c_{n-2} = 0 \). This is a contradiction, so \( f_{\delta,n,\xi}(z) \) has \( n \) distinct real roots.

By tracing backward from (2.15) to (2.12), we see that the properties of \( f_{\delta,n,\xi}(z) \) lead to (2.12) which in turn implies the following

\[
W_{\delta,n}' \text{diag}(p_1, p_2, \cdots, p_n) W_{\delta,n} = M_{2\delta,n},
\]

hence \( \text{diag}(p_1, p_2, \cdots, p_n) \) is similar to a positive definite matrix \( M_{2\delta,n} \), therefore its eigenvalues \( p_1, p_2, \cdots, p_n \) are positive. If \( \delta = 0 \), the first equation in (2.12) gives \( \sum_{i=1}^{n} p_i = 1 \). If \( \delta = 1 \), by (2.12) and Lemma 2, we have the following
\[ p_0 = 1 - \sum_{i=1}^{n} p_i \]
\[ = 1 - (p_1, p_2, \cdots, p_n)(1, 1, \cdots, 1)' \]
\[ = 1 - v_{1,n}' W_{1,n}^{-1}(1, 1, \cdots, 1)' \]
\[ = 1 - v_{1,n}' \left( -\frac{a_{n-1}}{a_n}, -\frac{a_{n-2}}{a_n}, \cdots, -\frac{a_1}{a_n}, -\frac{1}{a_k} \right)' \]
\[ = \frac{m_0 a_n + m_1 a_{n-1} + \cdots + m_{n-1} a_1 + m_n}{a_n} \]
\[ = \frac{1}{a_n} (m_0, m_1, \cdots, m_{n-1}, m_n)(a_n, a_{n-1}, \cdots, a_1, 1)' \]

Now we write (2.2) into the following form,
\[ (M_{1,n}, v_{n+1,n})(a_n, a_{n-1}, \cdots, a_1, 1)' = (0, 0, \cdots, 0)' \]
from which and the last equation in (2.16) we have
\[ (a_n, a_{n-1}, \cdots, a_1, 1) M_{0,n+1}(a_n, a_{n-1}, \cdots, a_1, 1)' = a_n^2 p_0. \]

Thus \( p_0 \) must be positive, since \( M_{0,n+1} \) is positive definite by Lemma 1. Q.E.D.

We now focus on two important classes of random variables, i.e., those with symmetric density functions and those which only attain nonnegative values. Normal distributions and Poisson distributions are typical examples.

For a random variable with symmetric density function, we may naturally assume that \( \psi_{\delta,n} \) also has symmetric states. we will change notations slightly as follows: let \( \psi_n \) be a random variable with \( n \) symmetric states which are denoted by \( \pm x_1, \pm x_2, \cdots, \pm x_k \) if \( n = 2k \) or \( 0, \pm x_1, \pm x_2, \cdots, \pm x_k \) if \( n = 2k + 1 \), where in each case, the \( z_i \)'s are
positive and ±z_i has probabilities p_{±i}. For a random variable with symmetric density function, the moments m_i = 0 if i is odd, so we let E_i = m_2i. We then modify the definitions of σ_i, v_{i,k}, M_{i,k}, W_{i,k} and A as follows: each z_i is replaced by y_i = x_i^2 in σ_i and W_{i,k}, each m_i is replaced by E_i in v_{i,k} and M_{i,k}. Finally A is still defined in terms of the σ_i's and δ = 1 or 0 depending on n is odd or even.

With such notations, we have the following theorem:

**Theorem 2** For any given n > 1, let ξ be a random variable with symmetric density function which satisfies one of the conditions in lemma 1 for some m > n. Then there is a unique random variable Ψ_n which has n symmetric states as described above and satisfies the following equations

\[ \mu_i = m_i \text{ for } 0 \leq i \leq 2n - 1. \] (2.17)

Furthermore, y_i = x_i^2, 1 ≤ i ≤ k are the k distinct positive real roots of the following polynomial

\[ f_{n,ξ}(y) = y^k + a_1 y^{k-1} + a_2 y^{k-2} + \cdots + a_k, \]

where

\[ [a_k, \cdots, a_2, a_1]^T = -M_{k, k}^{-1} v_{k+δ,k}, \] (2.18)

and p_{−i} = p_i, 1 \leq i \leq k

\[ [q_1, q_2, \cdots, q_k] = v_{δ,k} W_{δ,k}^{-1}, \] (2.19)

where \( q_i = p_i + p_{−i} \), and if δ = 1, then \( q_0 = p_0 = 1 - \sum_{i=1}^{k} p_i. \)
Proof: The fact that \( m_i = 0 \) for \( i = 1, 3, \ldots, 2n - 1 \) implies that \( p_{-i} = p_i, 1 \leq i \leq k \).

Let \( m = k - 1 \) and \( a = 0 \) in lemma 1; and let

\[
h_{k-1}(t_0, t_1, \ldots, t_{k-1}, x) = \sum_{j=1}^{k+1-1} t_{j-i} x^{i+2(j-i)}
\]

then the corresponding positive definite quadratic form has matrix \( M_{i,k} \), hence \( M_{i,k} \) is positive definite for every \( i \geq 0 \). Now consider the equations in (2.17) for which \( i = 2\delta, 2\delta + 2, \ldots, 2n - 2 \), applying the same argument as in the proof of theorem 1 with modified \( \sigma_i \)'s, \( v_i,k \)'s, \( M_{i,k} \)'s and \( A \) discussed above, we get a unique random variable \( \psi_{\delta,k} \) with \( k \) positive states \( y_i \) and probabilities \( q_i \), where \( q_i = p_{-i} + p_{+i} \) and \( q_0 = p_0 \) if \( \delta = 1 \). The positiveness of the states follows from lemma 3 and the fact that \( M_{\delta,k} \) and \( M_{\delta+1,k} \) are now both positive definite. Clearly this \( \psi_{\delta,k} \) is exactly \( \psi_n \). Q.E.D.

For the second class of random variables, the note that \( M_{1,n} \) is always positive definite, hence the roots of \( f_{n,k}(z) \) are all positive. To show this we only need to show that the quadratic form obtained by letting, in lemma 1, \( a = 0 \), \( m = n - 1 \) and \( h(t_0, t_1, \ldots, t_{n-1}) \) is replaced by \( x h_{n-1}(t_0, t_1, \ldots, t_{n-1}) \) is positive definite. But this is obvious, since now \( x \) only take nonnegative values.

2.3 Optimal N-nomial Discretizations of Wiener Processes

A Wiener process starting at 0 is the random process \( w_t, 0 \leq t < \infty \) with the following properties:
I. $w_0 = 0$.

II. For any $0 \leq t_0 < t_1 < \cdots < t_m$, the random variables $w_{t_1} - w_{t_0}, w_{t_2} - w_{t_1}, \ldots, w_{t_m} - w_{t_{m-1}}$ are independent.

III. The random variable $w_t - w_s, 0 \leq s \leq t$ has a normal distribution with mean 0 and variance $t - s$, i.e., $N(0, \sqrt{t - s})$.

Let $t_i = \frac{t_i}{m}, 0 \leq i \leq m$, then $\{w_{t_i}\}_{i=0}^m$ is a discrete time simulation of the Wiener process $w_t$ on the interval $[0, T]$, where $w_{t_i}$ has normal distribution $N(0, \sqrt{\frac{t_i}{m}})$. Consider the following $n$-nomial simulation of the discrete time Wiener process: let $\psi_n$ be the unique $n$-state random variable in theorem 2 with $\xi$ distributed as $N(0, \sqrt{\frac{t}{m}})$, and let $\psi_1^{(1)}, \psi_2^{(2)}, \ldots, \psi_m^{(m)}$ be $m$ independent random variables having the same distribution as $\psi_m$, and $\Psi_m^{(i)} = \psi_1^{(i)} + \psi_2^{(i)} + \cdots + \psi_m^{(i)}$, for $1 \leq i \leq m$, then uniformly on $i$, $\Psi_m^{(i)}$ is an optimal approximation to $w_{t_i}$, i.e., to the normal distribution $N(0, \sqrt{\frac{t_i}{m}})$, as shown below:

**Proposition 2** For any $i, 1 \leq i \leq m$, $\Psi_m^{(i)}$ and $w_{t_i}$, where $t_i = \frac{t_i}{m}$, have equal moments of order up to $2n - 1$.

**Proof:** $w_{t_i}$ has distribution $N(0, \sqrt{\frac{t_i}{m}})$, hence by addition property of normal distributions, it is the sum of $i$ mutually independent copies of $N(0, \sqrt{\frac{t}{m}})$. The j-th moment of any sum of $i$ independent random variables only depends on the moments
of its summands which are of order lower or equal to \( j \). Now both \( \Psi_m^{(i)} \) and \( N(0, \sqrt{\frac{2}{m}}) \) are sums of \( i \) independent random variables and by Theorem 2, and any two summands, one from each sum, have the equal moments up to \((2n - 1)\)-th order, so the two sum must have equal moments up to order \( 2n - 1 \). Q.E.D.

If \( \xi \) is normally distributed as \( N(0, \sigma) \), then all the moments are determined by \( \sigma \).

We denote the \( n \)-state random variable and the corresponding polynomial in theorem 2 by \( \psi_n(\sigma) \) and \( f_{n,\sigma}(y) \).

**Proposition 3** \( \psi_n(\sigma) = \sigma \psi_n(1) \) and \( f_{n,\sigma}(y) = \sigma^n f_{n,1}(y) \).

**Proof:** In this case, we have

\[
E_i = m_{2i} = (2i - 1)!! \sigma^{2i}
\]

(2.20)

If we make a substitution \( x_i \rightarrow \sigma x'_i \) in the equations (2.17), the resulting equations become independent of \( \sigma \), hence the proposition follows. Q.E.D.

This proposition makes it easy to find the \( \psi_n(\sigma) \) explicitly for small \( n \), we only need to find \( \psi_n(1) \). Here are some examples:

For \( n = 2, k = 1, \delta = 0, M_{\delta,1} = (1), M_{\delta+1,1} = (1), v_{\delta,1} = [1] \), hence

\[
A = M_{\delta,1}^{-1} M_{\delta+1,1} = (1)
\]

\[
f_{2,1}(x) = x - 1,
\]
\[ [q_1] = v_{\delta,1} W_{\delta,1}^{-1} = [1], \]

so the states and probabilities of \( \psi_2(1) \) are

\[ (\pm 1, \frac{1}{2}) \quad (2.21) \]

For \( n = 3, k = 1, \delta = 1, M_{\delta,1} = (1), M_{\delta+1,1} = (3), v_{\delta,1} = [1], \) hence

\[ A = M_{\delta,1}^{-1} M_{\delta+1,1} = (3) \]

\[ f_{3,1}(x) = x - 3, \]

\[ q_1 = v_{\delta,1} W_{\delta,1}^{-1} = \left[ \frac{1}{3} \right] \]

and the states and probabilities of \( \psi_3(1) \) are

\[ (\pm \sqrt{3}, \frac{1}{6}), (0, \frac{2}{3}) \quad (2.22) \]

For \( n = 4, k = 2, \delta = 0, \)

\[
M_{\delta,2} = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}, \quad M_{\delta+1,2} = \begin{pmatrix} 1 & 3 \\ 3 & 15 \end{pmatrix}
\]

and \( v_{\delta,2} = [1, 1], \) hence

\[ A = M_{\delta,2}^{-1} M_{\delta+1,2} = \begin{pmatrix} 0 & -3 \\ 1 & 6 \end{pmatrix} \]

\[ f_{4,1}(x) = x^2 - 6x + 3, \]

\[ [q_1, q_2] = v_{\delta,2} W_{\delta,2}^{-1} = \left[ \frac{3 + \sqrt{6}}{6}, \frac{3 - \sqrt{6}}{6} \right] \]

and the states and probabilities of \( \psi_4(\sigma) \) are

\[ (\pm \sqrt{3 - \sqrt{6}}, \frac{3 + \sqrt{6}}{6}), (\pm \sqrt{3 + \sqrt{6}}, \frac{3 - \sqrt{6}}{6}) \quad (2.23) \]
For $n = 5$, $k = 2$, $\delta = 1$,

$$M_{\delta,2} = \begin{pmatrix} 1 & 3 \\ 3 & 15 \end{pmatrix}, \quad M_{\delta+1,2} = \begin{pmatrix} 3 & 15 \\ 15 & 105 \end{pmatrix}$$

and $v_{\delta,2} = [1, 3]$, hence

$$A = M_{\delta,2}^{-1} M_{\delta+1,2} = \begin{pmatrix} 0 & -15 \\ 1 & 10 \end{pmatrix}$$

$$f_{5,1}(x) = x^2 - 10x + 15,$$

$$[q_1, q_2] = v_{\delta,2} W_{\delta,2}^{-1} = \left[ \frac{7 + 2\sqrt{5} - 2\sqrt{10}}{30}, \frac{7 - 2\sqrt{5}}{30} \right]$$

and the states and probabilities of $\psi_5(\sigma)$ are

$$(\pm \sqrt{\frac{5 - \sqrt{10}}{10}}, \frac{7 + 2\sqrt{10}}{30}, 0, \frac{8}{15}), (\pm \sqrt{\frac{5 + \sqrt{10}}{10}}, \frac{7 - 2\sqrt{10}}{30}) \quad (2.24)$$

For $n = 6, 7, 8, 9$, theoretically, we can also find the exact states and their probabilities of $\psi_n(\sigma)$, because we may apply the root formulas for polynomials of 3rd and 4th degree, but they will involve quite complicated arithmetic expressions, so we calculate the states and their probabilities numerically for $n > 5$. This can be easily accomplished by using such softwares as Maple, Matlab, or Mathematica etc..

For $n = 6$,

$$f_{6,1}(x) = x^3 - 15x^2 + 45x - 15$$

$$[q_1, q_2, q_3] = [0.8176569388, 0.1772314923, 0.00511156881]$$

$$[x_1^2, x_2^2, x_3^2] = [0.3803270184, 3.568985497, 11.05068748]$$

$$[x_1, x_2, x_3] = [0.6167065902, 1.889175878, 3.324257433]$$
For $n = 7$,

\[ f_{7,1}(x) = x^3 - 21x^2 + 105x - 105 \]

\[ [q_0, q_1, q_2, q_3] = [0.4571428572, 0.4802463571, 0.06151424800, 0.001096537707] \]

\[ [x_1^2, x_2^2, x_3^2] = [1.332651815, 5.601550108, 14.06579808] \]

\[ [x_1, x_2, x_3] = [1.154405395, 2.366759411, 3.750439718] \]

For $n = 8$,

\[ f_{8,1}(x) = x^4 - 28x^3 + 210x^2 - 420x + 105 \]

\[ [q_1, q_2, q_3, q_4] = [0.7460245147, 0.2344798153, 0.01927044030, 0.000225229063] \]

\[ [x_1^2, x_2^2, x_3^2, x_4^2] = [0.2906070430, 2.678194576, 7.853927003, 17.17727138] \]

\[ [x_1, x_2, x_3] = [0.5390798113, 1.636519042, 2.802485861, 4.144547186] \]

For $n = 9$,

\[ f_{9,1}(x) = x^4 - 36x^3 + 378x^2 - 1260x + 945 \]

\[ [q_0, q_1, q_2, q_3, q_4] = \\]

\[ [0.4063492055, 0.4881950071, 0.9983281307, 0.00557828264, 0.00004469168] \]

\[ [x_1^2, x_2^2, x_3^2, x_4^2] = [1.047052153, 4.313297527, 10.27477509, 20.36487523] \]

\[ [x_1, x_2, x_3, x_4] = [1.023255664, 2.076847979, 3.205429002, 4.512745864] \]
2.4 Optimal N-nomial Discretizations of Poisson Counting Processes

A Poisson counting process for times $t \geq t_0$ is by definition a counting process \( \{N_t; t \geq t_0\} \) with the following three properties:

(i). \( \Pr [N_{t_0}=0] = 1; \)

(ii). for $t_0 \leq s \leq t$, the increment $N_{s,t} = N_t - N_s$ is a Poisson distributed with parameter $\Lambda_t - \Lambda_s$,

\[
\Pr [N_{s,t} = n] = (n!)^{-1}(\Lambda_t - \Lambda_s)^n exp[-(\Lambda_t - \Lambda_s)]
\]

for $n = 0, 1, 2, \ldots$, where $\Lambda_t$ is a nonnegative, nondecreasing function of $t$;

(iii). \( \{N_t; t \geq t_0\} \) has independent increments.

With the condition of absolute continuity on $\Lambda_t$, we can write

\[
\Lambda_t = \int_{t_0}^{t} \lambda_y dy
\]

and we call $\lambda_t$ the intensity function of the process. If $\lambda$ is constant, the process is called homogeneous Poisson process, otherwise it is called a inhomogeneous Poisson Process.

A marked point process is a point process with an auxiliary variable, called a mark, associated with each point (or event). The mark is usually used to identify random quantities associated with the point it accompanies.
Consider a process in which the points or events are the sudden jumps of the price of a stock. Presumably the jumps occur discretely, so the process is a point process. For each jump, the magnitude of the jump can be considered as a random variable. In our terminology, this random variable is the mark associated with the jump. Thus the sudden jumps of stock price can be reasonably modeled with a marked point process. The mark space in which the mark takes value, in our treatment, will be assumed to be a denumerable collection of real numbers or the real line or part of it. we will also assume that the marks are independent and are of the same distribution.

Let \(\{N_t; t \geq 0\}\) be a counting process indicating the total number of points (events) in the interval \([t_0, t)\) regardless of their mark, then the mark accumulator process \(\{x_t, t \geq t_0\}\) is defined by

\[
x_t = \sum_{i=0}^{N_t} u_i
\]

where we identify \(u_0 = 0\), thus \(x_t\) is the sum of all marks on all points occurring in \([t_0, t)\). In this paper we are mainly interested in a special kind of marked point process, called compound Poisson process. It is a marked point process which satisfy the following properties:

(i). \(\{N_t; t \geq t_0\}\) is an inhomogeneous Poisson counting process with intensity function \(\lambda_t\) for \(t \geq t_0\);
(ii). \( \{ u_t \} \) is a sequence of mutually independent, identically distributed random variables which are also independent of \( \{ N_t; t \geq t_0 \} \)

We use the same terminology for the associated mark accumulator process \( \{ x_t; t \geq 0 \} \). The independence in (ii) implies that \( \{ x_t; t \geq t_0 \} \) has independent increments.

A Poisson counting process has unit jumps and can be viewed as a special case of a compound Poisson process when the mark space consists of a single point \( \mathcal{U} = \{ 1 \} \).

In this work, we will mainly use a result on representations for a compound process in terms of sums of independent Poisson process. These representations are useful not only for the insight they provide into the structure of the compound Poisson process but also in many applications where this process is used as a model. In particular, such representations provide some insights into Merton’s model of discontinuous stock price movement, which leads to effective implementation procedures for option prices on the stock if the mark space is simple.

We have the following representation for a compound Poisson process having a denumerable mark space.

**Lemma 4** Let \( \{ x_t; t \geq t_0 \} \) be a compound Poisson process with a denumerable mark space \( \mathcal{U} = \{ U_1, U_2, \ldots \} \). Points occur as a Poisson process with intensity \( \lambda_t \), and the probability \( U_k \) occurs as a mark on a particular point is \( p_k \) independently of the marks occurring on other points. Let \( \{ N_T(U_k); t \geq t_0 \} \) for \( k = 1, 2, \ldots \), are mutually
independent Poisson counting process for which the k-th process has intensity $p_k \lambda_t$ for $t \geq t_0$.

In terms of these, we have the representation of $x_t$ as

$$x_t = \sum_{k=1}^{\infty} U_k N_t(U_k)$$

for $t \geq t_0$.

Let $t_i = \frac{i \ell T}{m}, 1 \leq i \leq m$, then $\{N_{t_i}\}_{i=0}^m$ is a discrete time simulation of the Poisson counting process on the interval $[0, T]$, where $N_{t_i}$ is Poisson distributed with parameter $\Lambda_{t_i}$. Consider the following n-nomial simulation of above discrete time Poisson counting process: let $\psi^{(i)}_{\delta,n}(\lambda_i)$ be the unique $n + \delta$-state random variable given by theorem 1 with $\xi = N_{t_i} - N_{t_{i-1}}$ which is Poisson distributed with parameter $\lambda_i = \Lambda_{t_i} - \Lambda_{t_{i-1}}$; let $\Psi_m^{(i)} = \sum_{j=0}^{m} \psi^{(i)}_{\delta,n}(\lambda_i)$ for $1 \leq i \leq m$, then uniformly on $i$, $\Psi_m^{(i)}$ is an optimal approximation to $N_{t_i}$ in the following sense:

**Proposition 4** For any $i, 1 \leq i \leq m$, $\Psi_m^{(i)}$ and $N_{t_i}$ have equal moments of order up to $2n + \delta - 1$.

Note that if $\Lambda_t = \lambda(t + t_0)$, then the $N_i$'s are identically distributed, so are the $\psi^{(i)}_{\delta,n}$'s.

The proof of the above proposition is identical to that of proposition 1.

Let $\psi_{\delta,n,\lambda}$ and $f_{\delta,n,\lambda}(x)$ be the $n$-state random variable and its associated polynomial given by theorem 1 with $\xi$ Poisson distributed with parameter $\lambda$. In what follows, we first find $\psi_{\delta,n,\lambda}$ explicitly for $n = 2$, then we will discuss some general properties of the polynomial $f_{\delta,n,\lambda}(x)$ which will be used later.
The first few moments of $\xi$ are

$$m_0 = 1, \ m_1 = \lambda, \ m_2 = \lambda + \lambda^2, \ m_3 = \lambda + 3\lambda^2 + \lambda^3, \ m_4 = \lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4$$

hence

$$M_{0,2} = \begin{pmatrix} 1 & \lambda \\ \lambda & \lambda + \lambda^2 \end{pmatrix}$$

$$M_{1,2} = \begin{pmatrix} \lambda & \lambda + \lambda^2 \\ \lambda + \lambda^2 & \lambda + 3\lambda^2 + \lambda^3 \end{pmatrix}$$

and

$$v_{0,2} = [1, \lambda]', \ v_{1,2} = [\lambda, \lambda + \lambda^2]'$$

$$v_{2,2} = [\lambda + \lambda^2, \lambda + 3\lambda^2 + \lambda^3]'$$

$$v_{3,2} = [\lambda + 3\lambda^2 + \lambda^3, \lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4]'$$

For $n = 2$ and $\delta = 0$, it can be solved easily by theorem 1 that

$$f_{0,2,\lambda}(x) = x^2 - (2\lambda + 1)x + \lambda^2$$

and that the states and their associated probabilities of $\psi_{0,2,\lambda}$ are

$$[x_1, x_2] = \left[ \frac{2\lambda + 1 - \sqrt{4\lambda + 1}}{2}, \frac{2\lambda + 1 + \sqrt{4\lambda + 1}}{2} \right]$$

$$[p_1, p_2] = \left[ \frac{\sqrt{4\lambda + 1} + 1}{2\sqrt{4\lambda + 1}}, \frac{\sqrt{4\lambda + 1} - 1}{2\sqrt{4\lambda + 1}} \right]$$

For $n = 2$ and $\delta = 1$,

$$f_{1,2,\lambda}(x) = x^2 - (2\lambda + 3)x + (\lambda^2 + 2\lambda + 2)$$

and the states and their probabilities of $\psi_{1,2}$ are

$$[x_0, x_1, x_2] = \left[ 0, \frac{2\lambda + 3 - \sqrt{4\lambda + 1}}{2}, \frac{2\lambda + 3 + \sqrt{4\lambda + 1}}{2} \right]$$
\[ p_0 = \frac{8}{(2\lambda + 3 - \sqrt{4\lambda + 1})(2\lambda + 3 + \sqrt{4\lambda + 1})}, \]

\[ p_1 = \frac{\lambda(\sqrt{4\lambda + 1} + 1)}{(2\lambda + 3 - \sqrt{4\lambda + 1})\sqrt{4\lambda + 1}}, \]  

(2.28)

\[ p_2 = \frac{\lambda(\sqrt{4\lambda + 1} - 1)}{(2\lambda + 3 + \sqrt{4\lambda + 1})\sqrt{4\lambda + 1}}. \]

**Lemma 5** As in theorem 1, let the \( z_{i,\lambda} = z_{i,\delta,\mu,\lambda} \)'s be roots of \( f_{\delta,\mu,\lambda}(z) \) and \( p_{i,\lambda} = p_{i,\delta,\mu,\lambda} \)'s are given by (2.3), then we have

\[
\sum_{i=1}^{n} p_{i,\lambda}(z_{i,\lambda} - \lambda)^2(z_{i,\lambda} - \lambda - 1)^2 = 3\lambda^2
\]

(2.29)

and for arbitrary small \( \epsilon > 0, \)

\[
m \sum_{|z_{i,\lambda} - 1| > \epsilon} p_{i,\lambda} = 0 \text{ as } m \to \infty
\]

(2.30)

**Proof:**

\[
\sum_{i=1}^{n} p_{i,\lambda}z_{i,\lambda}^2(z_{i,\lambda} - \lambda)^2
\]

\[
= \sum_{i=1}^{n} p_{i,\lambda}(z_{i,\lambda}^4 - (4\lambda + 2)z_{i,\lambda}^3 + (6\lambda^2 + 6\lambda + 1)z_{i,\lambda}^2 - (4\lambda^3 + 6\lambda^2 + 2\lambda)z_{i,\lambda} + (\lambda^4 + 2\lambda^3 + \lambda^2))
\]

\[
= m_4 - (4\lambda + 2)m_3 + (6\lambda^2 + 6\lambda + 1)m_2 - (4\lambda^3 + 6\lambda^2 + 2\lambda)m_1 + (\lambda^4 + 2\lambda^3 + \lambda^2))
\]

\[
= 3\lambda^2
\]

this proves (2.29). For sufficient large \( m, \) \( |z_{i,\lambda} - 1| > \epsilon \) implies \( |z_{i,\lambda} - \lambda - 1| > \epsilon, \)

thus it follows from (2.29) that
\[
m \sum_{|z_i, \frac{\lambda}{m} - 1| > \varepsilon}^m p_i \left( z_i, \frac{\lambda}{m} - \frac{\lambda}{m} \right)^2 < 3m \left( \frac{\lambda}{m} \right)^2 t^{-1} \rightarrow 0 \text{ as } m \rightarrow 0.
\]

Q.E.D.

By (2.2) in theorem 1, the coefficients of \( f_{\delta, n, \lambda}(x) \), i.e., the \( a_i \)'s satisfy the following equation

\[
[v_{\delta, n}, v_{\delta + 1, n}, \ldots, v_{n + 1, n}] [a_n, a_{n-1}, \ldots, a_1, -1]' = 0
\]

(2.31)

Let

\[
\prod_{i=1}^{i} (x - t) = x^i + c_{i,1} x^{i-1} + \cdots + c_{i,i-1} x + c_{i,i}
\]

for \( i = 1, 2, 3, \ldots \)

and define the matrix \( Q_s \) as follows:

\[
Q_s = \begin{pmatrix}
1 & 0 & 0 & 0 & \cdots & 0 & 0 \\
c_{11} & 1 & 0 & 0 & \cdots & 0 & 0 \\
c_{2,2} & c_{2,1} & 1 & 0 & \cdots & 0 & 0 \\
c_{3,3} & c_{3,2} & c_{3,1} & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
c_{s,s} & c_{s,s-1} & c_{s,s-2} & c_{s,s-3} & \cdots & c_{s,1} & 1
\end{pmatrix}
\]

Lemma 6

\[
Q_{n-1}[v_{1,n}, v_{2,n}, \ldots, v_{n+1,n}] =
\begin{pmatrix}
\lambda + O(\lambda^2) & \lambda + O(\lambda^2) & \lambda + O(\lambda^2) & \cdots & \lambda + O(\lambda^2) \\
\lambda^2 + O(\lambda^3) & 2\lambda^2 + 0(\lambda^3) & 2^2\lambda^2 + 0(\lambda^3) & \cdots & 2^n\lambda^2 + 0(\lambda^3) \\
\lambda^3 + O(\lambda^4) & 3\lambda^3 + 0(\lambda^3) & 3^2\lambda^3 + 0(\lambda^4) & \cdots & 3^n\lambda^3 + 0(\lambda^4) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\lambda^n + O(\lambda^{n+1}) & n\lambda^n + 0(\lambda^{n+1}) & n^2\lambda^n + 0(\lambda^{n+1}) & \cdots & n^n\lambda^n + 0(\lambda^{n+1})
\end{pmatrix}
\]

(2.32)
Proof: The moment generating function of a Poisson counting process is

\[ e^{\lambda(e^t-1)} = \sum_{i=0}^{\infty} \frac{(-1)^i m_i t^i}{i!} \]

Direct Taylor expansion gives

\[ m_i = \sum_{l=1}^{i} \frac{\lambda^l}{l!} \sum_{j_{i_1}, j_{i_2}, \ldots, j_l \geq 1} \frac{t^{l_i}}{i_1! i_2! \cdots i_l!} \]

Hence

\[ i=h+k \]

The innermost sum is zero if \( t < k \) by the definition of the \( c_{k,j} \)'s. For \( t = k + 1 \) and the sum over \( s \) yields \((k + 1)! (k + 1)^i\), hence

\[ \sum_{i=h}^{i=h+k} c_{k,h+k-i} m_i = \sum_{l=1}^{i} \frac{\lambda^l}{l!} \sum_{s=0}^{l-1} (-1)^s C_l^{l-s} (t - s)^i \]

This is exactly what the lemma claims. Q.E.D.

Lemma 7

\[ f_{\delta,n,0}(x) = \prod_{i=\delta}^{n+k-1} (x - i) \]

Proof: For \( \delta = 1 \), multiplying both sides of (2.31) by \( \text{diag}((\lambda^{-1}, \lambda^{-2}, \ldots, \lambda^{-n})) Q_{n-1} \), using lemma 6, and letting \( \lambda \to 0 \) to yield

\[
\begin{pmatrix}
1 & 1 & 1 & \cdots & 1 \\
1 & 2 & 2^2 & \cdots & 2^n \\
1 & 3 & 3^2 & \cdots & 3^n \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & n & n^2 & \cdots & n^n \\
\end{pmatrix}
\begin{pmatrix}
a_n(0) \\
a_{n-1}(0) \\
a_{n-2}(0) \\
\vdots \\
a_0(0) \\
\end{pmatrix}
= 0
\]
where $a_0(0) = 1$. Now it is now clear that $f_{1,n,0}(x)$ has roots $z_{i,0} = i$ for $i = 1, 2, \ldots, n$.

For $\delta = 0$, the first equation in the system (2.31) is

\[ [m_0, m_1, \ldots, m_n][a_n, a_{n-1}, \ldots, a_1, -1]' = 0. \]  

(2.34)

Since $m_i = \lambda + o(\lambda^2)$ for $i \geq 1$ and $m_0 = 1$, letting $\lambda \to 0$ in (2.34) to get $a_n(0) = 0$,

hence $f_{0,n,0}(x) = x f_{1,n-1,0}(x)$. Therefore $f_{0,n,0}(x)$ has roots $z_{i,0} = 0, 1, \ldots, n - 1$.

Q.E.D.

2.5 Stock Price Dynamics and Its Discrete Approximations

2.5.1 Continuous Model

Under continuous assumption, the dynamics of a stock price $S_t$ is typically modeled as a geometric Brownian motion, i.e.,

\[ \frac{dS_t}{S_t} = \mu dt + \sigma dw_t \]

where $w_t$ is a standard Wiener process, $\mu$ and $\sigma$ are assumed to be constants. Such a model implies that $S_t$ follows a lognormal distribution, i.e.

\[ \ln \frac{S_T}{S} \sim N((\mu - \frac{\sigma^2}{2})T, \sigma \sqrt{T}) \]  

(2.35)

where $S = S_0$ is deterministic. From which the Black-Scholes formula can be easily derived using risk-neutral valuation.
Write (2.35) as

$$\ln \frac{S_T}{S} \sim (\mu - \frac{\sigma^2}{2})T + N(0, \sigma \sqrt{T}).$$

Let $\psi_n = \psi_n(1)$ is as in proposition 2, and let $\Psi_{n,m} = \sigma \sqrt{T} \sum_{i=1}^{m} \psi_{n,i}$, where $\psi_{n,i}$'s are $m$ independent duplicates of $\psi_n$. By proposition 1, $\Psi_{n,m}$ has equal moments with $N(0, \sigma \sqrt{T})$ up to $(2n - 1)$-th order. Using $\Psi_{n,m}$ as an approximation to $N(0, \sigma \sqrt{T})$, then by (2.35), the stock price at time $T$ is approximated by a discrete random variable $S_T = Se^{(\mu - \frac{\sigma^2}{2})T + \Psi_m}$, hence a call option price on the stock is, by definition,

$$P_{n,m} = r_m^{-m} \sum_{i_1=\delta}^{m} \cdots \sum_{i_n=\delta}^{m} \frac{m!}{i_1!i_2! \cdots i_n!} \prod_{j=1}^{n} p_j^{i_j} C_{i_1, i_2, \ldots, i_n}$$

(2.36)

where

$$C_{i_1, i_2, \ldots, i_n} = Max(0, Se^{(\mu - \frac{\sigma^2}{2})T + \sigma \sqrt{T} \sum_{i=1}^{m} i_i^2 - K}),$$

$i_1, i_2, \ldots, i_n$ in the summation are nonnegative integers, $r_m = (1 + \frac{rT}{m})$, where $r$ is the riskless interest rate, and the $z_i$'s and the $p_i$'s are just the symmetrical states and probabilities of $\psi_n$ for which $z_0 = 0$ if $\delta = 1$. Note that these $p_t$'s and $z_t$'s are independent of $T$ and $\sigma$. We will show later that the limit of $P_m$ as $m \to \infty$ exists and equals the Black-Scholes's formula when we let $\mu = r$.

### 2.5.2 Discontinuous Model

The total change in the stock price, in this case, is assumed to be the composition of two types of changes: (1) The 'normal' vibrations in price, which is caused by new information which evolves gradually and whose impact only cause marginal changes
in the stock's value. This component is modeled by a standard geometric Brownian motion with a constant variance per unit time and it has a continuous path. (2) The 'abnormal' vibrations in price are due to the arrival of important new information about the stock that has more than a marginal effect on price. By its very nature, important information arrives only at discrete point in time. This component is modeled by a 'jump' process reflecting the non-marginal impact of the information.

Based on the above observations, Merton(1976) proposed a model for discontinuous stock price returns. We formulate his model in terms of a compound Poisson process as follows:

$$\frac{dS_t}{S_t} = (\mu - \lambda \tau) dt + \sigma dw_t + dx_t, \quad (2.37)$$

where, $\mu$ is the instantaneous expected return on the stock; $\sigma^2$ is the instantaneous variance of the return, conditional on no arrivals of important new information (i.e., the Poisson event does not occur); $dw_t$ is a standard Wiener process; $x_t$ is an independent compound Poisson process with mark $u$ and finally, $\lambda \tau$ is the instantaneous expected value of $x_t$, i.e., $\tau \equiv E(u)$ and $\lambda$ is the instantaneous arrival time of the underlying Poisson process.

Before considering equation (2.37) in detail, we first state the following special case of generalized Ito’s lemma:
Lemma 8 (Ito's lemma) Assume $F(x, y, z)$ is a continuous function and has continuous second derivative with respect to $x$, $y$ and $z$, then we have the following:

$$
\frac{dF}{dt}(t, w_t, x_t) = \left( \frac{\partial F}{\partial t}(t, w_t, x_t) + \frac{1}{2} \frac{\partial^2 F}{\partial x^2}(t, w_t, x_t) \right) dt
+ \frac{\partial F}{\partial x}(t, w_t, x_t) dw_t + (F(t, w_t, x_t) - F(t, w_t, x_{t-})).
$$

Assume that $\mu$, $\lambda$ and $\sigma$ are constants, then by applying the above Ito's lemma, one can check that the solution to (2.37) is

$$
S_t = S \exp[(\mu - \lambda t - \frac{\sigma^2}{2})t + \sigma z_t + q_t]
$$

(2.38)

or

$$
\ln \frac{S_t}{S} = (\mu - \lambda t - \frac{\sigma^2}{2})t + \sigma z_t + q_t
$$

(2.39)

where $z_t$ is a normal random variable with a zero mean and variance equal to $t$ and $q_t$ is a compound Poisson process with parameter $\lambda t$ and mark $v = -\ln(1 - u)$. In what follows, we denote by $v(n)$ the sum of $n$ independent random variables $v_i (1 \leq i \leq n)$ each of which has the same distribution as $v$.

From now on, we assume that the mark space of $v$ is finite, i.e., $v$ has state set $\mathcal{U} = \{U_1, U_2, \ldots, U_s\}$ with associated probability vector $\{p_1, p_2, \ldots, p_s\}$.

Apply lemma 4, we have $q_t = \sum_{i=1}^s U_i P_i(p_i\lambda t)$ where $\{P_i(p_i\lambda t); 1 \leq i \leq s\}$ are independent Poisson counting processes for which the $i$-th process has intensity $p_i \lambda t$.

Then (2.39) becomes, at time $T$,

$$
\ln \frac{S_T}{S} = (\mu - \lambda T - \frac{\sigma^2}{2})T + \sigma z_T + \sum_{i=1}^s U_i P_i(p_i\lambda T)
$$

(2.40)
For each $i$, and let $t_k = \frac{kT}{m}$, we construct $(n+\delta)$-nomial approximations to $P_i(p_i\lambda t_k)$ as follows: let $\Phi_{i,k,\delta,n,m} = \sum_{j=1}^{k} \phi_{i,\delta,n,m}^{(j)}$ where the $\phi_{i,\delta,n,m}^{(j)}$'s are $k$ independent duplicates of $\phi_{i,\delta,n,m} = \psi_{\delta,n,m}^{(\epsilon \lambda T/m)}$ which is obtained by theorem 1 with $\xi = P_i(p_i\lambda T/m)$. Then proposition 3 implies that uniformly on $k$ with $1 \leq k \leq m$, $q_{it_k}$ and the discrete random variable $\Phi_{k,\delta,n,m} = \sum_{i=1}^{m} U_i \Phi_{i,k,\delta,n,m}$ have equal moments up to $(2n+\delta-1)$-th order. So we let $\Phi_{\delta,n,m} = \Phi_{m,\delta,n,m}$ be the discrete random variable approximation for $q_T$, and let $P_{si,n,m}$ be the discrete random variable approximation constructed for $\sigma w_T$ in the continuous case. Finally we use $\Phi_{n_1,m} + \Phi_{\delta,n_2,m}$ to approximate $\sigma w_T + q_T$, then by (2.40), the stock price is well approximated, by

$$S_T = S_0 e^{(\mu - \lambda T - \frac{\sigma^2}{2})T + \Phi_{\delta,n,m} + \psi_{n,m}}$$  \hspace{1cm} (2.41)

Under this approximation, the price of a call option on the stock can be obtained immediately by taking expectation of $\max[S_T - K, 0]$. We will show in a later section that, if we replace $\mu$ by $r$, then, as $m \to \infty$, the option price obtained this way converges to the Merton's formula in the case $\tau = 0$.

We now derive a formula for the expectation of $\max(S_t - K, 0)]$ under risk-neutral evaluation (i.e., replacing $\mu$ by $r$, the riskless interest rate).

Let $g(x)$ and $p(x)$ be the distribution function and density function of the random variable $\psi_t$ which has distribution $(\mu - \lambda T - \frac{\sigma^2}{2})t + \sigma z_t$ or equivalently $N((\mu - \lambda T - \frac{\sigma^2}{2})t, \sigma)$, and let $h(x)$ and $H(x)$ be the distribution functions of $q_t$ and $\psi_t + q_t$, respectively.
Then we have the following:

\[ \frac{dS_t}{S} = e^{\mu_t + \eta_t} \]

\[ h(x) = Pr[q_t < x] = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} Pr[v(k) < x] \]

and

\[ H(x) = \int_{-\infty}^{\infty} g(x - \xi) d\eta(\xi) \]

\[ = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} \int_{-\infty}^{\infty} g(x - \xi) dPr[v(k) < \xi] \]

**Theorem 3** Under risk neutral valuation, i.e., replacing \( \mu \) by the riskless interest rate \( r \), the option price on the stock is

\[ C(S, t, K, r) = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t(1+r)} [\xi_k W(Se^{v(k)}, t, K, \sigma^2, r - \lambda t)] \]

**Proof:**

\[ C(S, t, K, r) \]

\[ = E[e^{-rt} \max[S(t) - K, 0]] \]

\[ = e^{-rt} \int_{x > \ln \frac{K}{S}} (Se^{x} - K) dH(x) \]

\[ = e^{-rt} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} \int_{x > \ln \frac{K}{S}} (Se^{x} - K) (\int_{-\infty}^{\infty} p(x - \xi) dPr[v(k) < \xi]) dx \]

\[ = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t(1+r)} \int_{-\infty}^{\infty} (e^{-r - \lambda t}) \int_{x > \ln \frac{K}{S}} (Se^{x} - K) p(x - \xi) dx dPr[v(k) < \xi] \]
making a substitution \( y = x - \xi \) in the inner most integral, then

\[
C(S, t, K, r) = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t(1+r)} \int_{-\infty}^{\infty} (e^{-(r-\lambda r)t} \int_{y>\ln S_k \xi} (Se^\xi e^y - K)p(y)dy) dPr[v(k) < \xi]
\]

\[
= \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t(1+r)} \int_{-\infty}^{\infty} W(Se^\xi, t, K, \sigma^2, r - \lambda r) dPr[v(k) < \xi]
\]

\[
= \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t(1+r)} [\xi_k W(Se^{\xi(k)}, t, K, \sigma^2, r - \lambda r)]
\]

Q.E.D.

Note that the above formula is the same as the one given by Merton(1976) when \( \mu = r \) and \( \tau = 0 \).

### 2.6 Two Limit Theorems

**Lemma 9** Given an integer \( n(n > 1) \), let \( r, \mu, \sigma \) and \( T \) be any positive real numbers, suppose we have a sequence of random variables \( \{\zeta_m | m = 1, 2, \cdots\} \) satisfying the following conditions:

1. Each \( \zeta_m \) has \( n \)-states and the \( n \) states, together with their probabilities are \( z_{m,i}, p_{m,i}, \ 1 \leq i \leq n \).

2. \( mE(\zeta_m) \rightarrow (\mu - \frac{\sigma^2}{2})T \) as \( m \rightarrow \infty \).

3. \( mD^2(\zeta_m) \rightarrow \sigma^2 T \) as \( m \rightarrow \infty \).
(1.4) \( z_m^{(i)} = O\left(\frac{1}{m^\alpha}\right) \) uniformly on \( i \), for some \( \alpha > \frac{1}{3} \).

Let \( \Gamma_m = \zeta_m^{(1)} + \zeta_m^{(2)} + \cdots + \zeta_m^{(n)} \) be the sum of \( m \) independent random variables having the same distribution as \( \zeta_m \), then \( \Gamma_m \) converges to \( N((\mu - \frac{\sigma^2}{2})T, \sigma T) \) in distribution.

**Proof:** By Lindeberg-Feller's theorem [see theorem 1 in section 49 of Gnedenko, 1962, p. 321], conditions (1.1) - (1.4) guarantee that the sequence of distribution functions of \( \frac{\Gamma_m - (\mu - \frac{\sigma^2}{2})T}{\sigma \sqrt{T}} \), \( m = 1, 2, 3, \cdots \) converge to the distribution function of \( N(0, 1) \), hence the sequence of distribution functions of the \( \Gamma_m \)'s converges to the distribution function of \( N((\mu - \frac{\sigma^2}{2})T, \sigma \sqrt{T}) \). Q.E.D.

**Theorem 4** Let

\[
P_m = r_m^m \int_{\ln(S/K)}^{\infty} (Se^x - K)dF_m(x)
\]

where \( r_m = (1 + \frac{rT}{m}) \) and \( F_m(x) \) is the distribution function of \( \Gamma_m \), then we have

\[
\lim_{m \to \infty} P_m = SN(d_1) - Ke^{-rT}N(d_2) \quad \text{(Black - Scholes formula)},
\]

where

\[
d_1 = \frac{\ln(S/K) + (r + \frac{\sigma^2}{2})T}{\sigma \sqrt{T}},
\]

\[
d_2 = \frac{\ln(S/K) + (r - \frac{\sigma^2}{2})T}{\sigma \sqrt{T}} = d_1 - \sigma \sqrt{T}.
\]

Note that we can write \( P_m \) in discrete form which is convenient for computations. We will discuss this in section 7.
Lemma 10 Suppose \( h(x) \) is continuous function, \( g(x) > 0 \), and \( \frac{|h(x)|}{g(x)} \to 0 \) as \( |x| \to \infty \). If \( F_m(x) \Rightarrow F(x) \) (weak convergence) and
\[
\limsup_{m \to \infty} \int g(x) dF_m(x) < \infty,
\]
then
\[
\int h(x) dF_m(x) \to \int h(x) dF(x).
\]
This is exercise 2.12 in [29] on page 74.

Proof of Theorem 4:. For each \( m > 1 \), let \( \eta_m = e^{Tm} \) and \( G_m(x) \) be the distribution function of \( \eta_m \), then the sequence \( \{G_m(x)\}_{m=1}^{\infty} \) converges to the lognormal distribution function with parameters \( (\mu - \frac{\sigma^2}{2})T \) and \( \sigma \sqrt{T} \), i.e.
\[
G_m(x) \Rightarrow G(x) = \int_0^x \frac{1}{\sqrt{2\pi T}} e^{-\frac{(u-(\mu-\frac{\sigma^2}{2})T)^2}{2\sigma^2 T}} \, du.
\]
Consider
\[
E(\eta_m^2) = \int_{-\infty}^{\infty} x^2 dG_m(x), \text{ note that } G_m(x) = 0 \text{ for } x < 0
\]
by the independence of the \( \zeta_m^{(i)} \)'s and conditions (1.1)-(1.4), we have the following:
\[
E(\eta_m^2)
= (E(e^{2\zeta_m}))^m = e^{2mE\zeta_m} (E(e^{2(\zeta_m - E\zeta_m)}))^m
= e^{2mE\zeta_m} \left( \sum_{i=1}^n e^{2(z_{m,i} - E(\zeta_m))p_{m,i}} \right)^m
= e^{2mE\zeta_m} \left( \sum_{i=1}^n p_{m,i} (1 + 2(z_{m,i} - E(\zeta_m))) + 2(z_{m,i} - E(\zeta_m))^2 + O(m^{-3\alpha}) \right)^m
= e^{2mE\zeta_m} \left( 1 + 2D^2(\zeta_m) + O(m^{-3\alpha}) \right)^m
\]
Write $P_m$ in the following form, by making the substitution $x \to e^x$

$$P_m = e^{-mE\xi_m} \left( 1 + \frac{2mD^2(\xi_m) + O(m^{-3a-1})}{m} \right)^m \to e^{2(\mu - \frac{\sigma^2}{2})T + 2\sigma^2 T} \text{ as } m \to \infty.$$
(1.3) \( mD^{2}(\gamma_{m}) \to \lambda t \) as \( m \to \infty \).

(1.4) \( m \sum_{|\gamma_{m,i} - 1| > \epsilon} p_{m,i}(\gamma_{m,i} - E(\gamma_{m}))^{2} \to 0 \) for any \( \epsilon > 0 \).

Let \( \Gamma_{m} = \gamma_{m}^{(1)} + \gamma_{m}^{(2)} + \cdots + \gamma_{m}^{(m)} \) be the sum of \( m \) independent random variables each of which having the same distribution as \( \gamma_{m} \), then \( \Gamma_{m} \) converges in distribution to \( p_{t} \).

**Proof:** This is only a special case of a known result [see theorem 2 in section 49 of Gnedenko, 1962, p. 324]. Q.E.D.

**Theorem 5** Let \( \Psi_{n_{1},m} \) and \( \Phi_{n_{2},m} \) be as in section 5, and \( F_{m}(x) \) be the distribution function of \( (r - \lambda t - \frac{\sigma^{2}}{2})T + \Phi_{n_{1},m} + \Psi_{n_{2},m} \) and \( r, S \) and \( K \) be real positive numbers. Define

\[
C_{n_{1},n_{2},m}(S, T, K, r) = (1 + \frac{rT}{m})^{-m} \int_{x > \ln \frac{K}{S}} (Se^{x} - K) dF_{m}(x)
\]

then \( C_{n_{1},n_{2},m}(S, T, K, r) \to C(S, T, K, r) \) where \( C(S, T, K, r) \) is described by Theorem 3.

**Proof:** Let \( G_{m}(x) \) be the distribution of \( e^{\alpha + \Phi_{n_{1},m} + \Psi_{n_{2},m}} \) where \( \alpha = (r - \lambda t - \frac{\sigma^{2}}{2})T \), then

\[
\int_{-\infty}^{\infty} x^{2} dG_{m}(x) = E(e^{2\alpha + 2\Phi_{n_{1},m} + 2\sum_{i=1}^{s} U_{i} \Phi_{i,n_{2},m}}) = e^{2\alpha} E(e^{2\Psi_{n_{1},m}}) \prod_{i=1}^{s} E(e^{2U_{i} \Phi_{i,n_{2},m}}) = e^{2\alpha}(E(e^{2\sqrt{\frac{T}{m}} \Psi_{n_{1}}}))^{m} \prod_{i=1}^{s}(E(e^{2U_{i} \Phi_{i,n_{2}}(\frac{E_{i} \lambda T}{m})}))^{m}
\]
Similar to the proof of theorem 4, we have

\[(E(e^{2\sigma \sqrt{\frac{T}{m}}\psi_n}))^m \rightarrow e^{2\sigma^2 T}, \text{ as } m \rightarrow \infty.\]

We also have

\[E(e^{2U_1\phi_{\delta,n}(\frac{p_{\lambda T}}{m})})\]

\[= \sum_{j=0}^{n} \frac{(\frac{2U_1z_{j,\delta,n,p_{\lambda T}}}{k!})}{\sum_{j=0}^{n} (2U_1z_{j,\delta,n,p_{\lambda T}})^k} \]

\[= \sum_{k=0}^{\infty} \frac{(2U_1)^k}{k!} \sum_{j=0}^{n} (\frac{z_{j,\delta,n,p_{\lambda T}}}{m})^k \]

\[= \sum_{k=0}^{\infty} \frac{(2U_1)^k}{k!} m_k \quad (2.42)\]

where \(m_k\), with some abuse of symbol, is the \(k\)-th moment of the random variable \(\phi_{\delta,n}(\frac{p_{\lambda T}}{m})\). By lemma 7, \(z_{j,\delta,n,p_{\lambda T}} < n + 1\) for sufficiently large \(m\), hence

\[m_k = \sum_{j=0}^{n} (\frac{z_{j,\delta,n,p_{\lambda T}}}{m})^k < (n + 1)^k m_1 = (n + 1)^k \left(\frac{p_{\lambda T}}{m}\right), \text{ for } i \geq 1.\]

Therefore, it follows from (2.42) that

\[(E(e^{2U_1\phi_{\delta,n}(\frac{p_{\lambda T}}{m})}))^m < (1 + \sum_{k=1}^{\infty} (\frac{2U_1(n + 1)^k p_{\lambda T}}{m})^k) m \]

\[< e^{\frac{p_{\lambda T}}{m}(e^{2U_1(n + 1)^k}) - 1)}\]

Therefore

\[\int_{-\infty}^{\infty} x^2 dG_m(x) < \infty\]

By lemma 10, we have

\[\lim_{m \rightarrow \infty} \left(1 + \frac{rT}{m}\right)^{-m} \int_{0}^{\infty} (Sx - K)dG_m(x) = e^{-rT} \int_{0}^{\infty} (Sx - K)dG(x) \quad (2.43)\]

where \(G(x)\) is the distribution function of the random variable \(e^{a+\sigma \omega T + q}T\). After making the substitution \(x \rightarrow e^x\), the expression under the limit on the left hand side...
of (2.43) is the exactly the $C_{n_1,n_2,m}(s,t,k,r)$, and the right hand side of (2.43) is $C(s,t,k,r)$ by definition. Q.E.D.

2.7 Implementation Procedures

2.7.1 Continuous Case

In the limit theorem 4, $P_m$ can be written in discrete forms and (2.36) is a special one which satisfies some optimality conditions and has parameters $p_i$’s and $z_i$’s in the forms given in section 4. Unless there is a special need for bigger $n$, we may always choose the parameters to be in one of forms in (2.21) – (2.24).

For $n = 2$, using (2.21) in (2.36), we have

$$P_{2,m} = (2r_m)^{-m} \sum_{i_1 + i_2 = m} \frac{m!}{i_1!i_2!} C_{i_1,i_2}$$

(2.44)

where

$$C_{i_1,i_2} = Max(0, \frac{e}{2})^{i_1} \frac{1}{6}^{i_2} C_{i_0,i_1,i_2}$$

For $n = 3$, using (2.22) in (2.36), we have

$$P_{3,m} = r_m^{-m} \sum_{i_0 + i_1 + i_2 = m} \frac{m!}{i_0!i_1!i_2!} \left( \frac{2}{3} \right)^{i_0} \left( \frac{1}{6} \right)^{i_1} \frac{1}{6}^{i_2} C_{i_0,i_1,i_2}$$

(2.45)

where

$$C_{i_0,i_1,i_2} = Max(0, \frac{e}{2})^{i_1} \frac{1}{6}^{i_2} C_{i_0,i_1,i_2}$$
For $n = 4$,

$$P_{4,m} = r_m^{-m} \sum_{i_0+i_1+i_2+i_3+i_4=m} \frac{m!}{i_0!i_1!i_2!i_3!i_4!} q_1^{i_1+i_2} q_2^{i_3+i_4} C_{i_0,i_1,i_2,i_3,i_4}$$  (2.46)

where $q_1 = \frac{3+\sqrt{6}}{6}$, $q_2 = \frac{3+\sqrt{6}}{6}$, and

$$C_{i_0,i_1,i_2,i_3,i_4} = \text{Max}(0, S e^{\frac{e^2}{2} T} + \sigma \sqrt{\frac{T}{m}}(z_1(t_1-t_2) + z_2(t_3-t_4)) - K).$$

with $z_1 = \sqrt{3} - \sqrt{6}$ and $z_2 = \sqrt{3} + \sqrt{6}$.

For $n = 5$,

$$P_{5,m} = r_m^{-m} \sum_{i_0+i_1+i_2+i_3+i_4=i_0+i_1+i_2+i_3+i_4=m} \frac{m!}{i_0!i_1!i_2!i_3!i_4!} q_1^{i_0+i_1+i_2} q_2^{i_3+i_4} C_{i_0,i_1,i_2,i_3,i_4}$$  (2.47)

where $q_0 = \frac{8}{15}$, $q_1 = \frac{7+2\sqrt{10}}{60}$, $q_2 = \frac{7-2\sqrt{10}}{60}$, and

$$C_{i_0,i_1,i_2,i_3,i_4} = \text{Max}(0, S e^{\frac{e^2}{2} T} + \sigma \sqrt{\frac{T}{m}}(z_1(t_1-t_2) + z_2(t_3-t_4)) - K).$$

with $z_1 = \sqrt{5} - \sqrt{10}$ and $z_2 = \sqrt{5} + \sqrt{10}$.

Similar formulas $P_{n,m}$ with $n > 5$ can be obtained explicitly. But they do not look as pretty as the ones given above, since the parameters can only be calculated numerically.

### 2.7.2 Discontinuous Case

In theorem 5, let $n_1 = 2, 3, 4$ or $5$, $n_2 = 2$ and $\delta = 0$ or $1$, we can make $C_{n_1,\delta,n_2,m}(S, T, K, r)$ explicit. Note that we can choose bigger $n_1$ with numerical parameters, but for any $n_2 > 2$, it is impossible to make the formula explicit. In this sense we are fortunate
to have the only choice for \( n_2 \) for which \( C_{n_1, \delta, n_2, m}(S, T, K, r) \) can be made explicit for implementation. For simplicity of exposition, we will make the following assumptions: the mark of the corresponding compound Poisson process \( q_t \) in (2.38) or (2.39) has only two states, say \( U_1 \) and \( U_2 \) and their associated probabilities are \( p(U_1) \) and \( p(U_2) \) respectively. For \( n_1 = 2 \) and \( \delta = 0 \), it follows from (2.40) (with \( \mu \) replaced by \( r \)), (2.25) and (2.26) that the option price with strike price \( K \) on the stock with current price \( S \) is

\[
C_{2,0,2,m}(S, T, K, r) = (2r_m)^{-m} \sum_{i_1+i_2=m} \sum_{j_1+j_2=m} \sum_{k_1+k_2=m} f(i_1, j_1, k_1)q(i_1, j_1, k_1)C_{i_1,j_1,k_1}
\]

where

\[
f(i, j, k) = \frac{m^i m^j m^k}{i!(m-i)!j!(m-j)!k!(m-k)!}
\]

\[
q(i_1, j_1, k_1) = p_1^{i_1} p_2^{j_1} q_1^{k_1} q_2^{k_2}
\]

and

\[
p_1 = \frac{\sqrt{4p(U_1)\lambda T + 1}}{2\sqrt{4p(U_1)\lambda T + 1}} + 1
\]

\[
p_2 = \frac{\sqrt{4p(U_2)\lambda T + 1}}{2\sqrt{4p(U_2)\lambda T + 1}} - 1
\]

\[
q_1 = \frac{\sqrt{4p(U_1)\lambda T + 1}}{2\sqrt{4p(U_1)\lambda T + 1}} + 1
\]

\[
q_2 = \frac{\sqrt{4p(U_2)\lambda T + 1}}{2\sqrt{4p(U_2)\lambda T + 1}} - 1
\]

\[
C_{i_1,i_2,j_1,j_2,k_1,k_2} = \text{Max}\left[0, Se^{\alpha + \sigma \sqrt{\frac{T}{m}((i_1-i_2)+U_1(j_1z_1+j_2z_2)+U_2(k_1y_1+k_2y_2))}}\right]
\]

\[
\alpha = (r - \lambda T - \frac{\sigma^2}{2})T
\]
\[
\begin{align*}
X_1 &= \frac{2p(U_1)\lambda T_m}{m} + 1 - \sqrt{\frac{4p(U_1)\lambda T_m}{m} + 1} \\
X_2 &= \frac{2p(U_2)\lambda T_m}{m} + 1 + \sqrt{\frac{4p(U_2)\lambda T_m}{m} + 1} \\
Y_1 &= \frac{2p(U_2)\lambda T_m}{m} + 1 - \sqrt{\frac{4p(U_2)\lambda T_m}{m} + 1} \\
Y_2 &= \frac{2p(U_2)\lambda T_m}{m} + 1 + \sqrt{\frac{4p(U_2)\lambda T_m}{m} + 1} 
\end{align*}
\]

(2.52)

For \( n_1 = 2 \) and \( \delta = 1 \), Similar formula follows from (2.40) (with \( \mu \) replaced by \( r \),
(2.27) and (2.28).

\[
C_{2,1,2,1}(S, T, K, r) = (2r_m)^{-m} \sum_{i_1+i_2=m} \sum_{j_1+j_2 \leq m} \sum_{k_1+k_2 \leq m} f(i, j, j_2, k_1, k_2)q(i, j, j_2, k_1, k_2)C_{i_1,j_1,j_2,k_1,k_2}
\]

where

\[
f(i, j, k, g, h) = \frac{m!m!m!}{i!(m-i)!(j-k)!(m-j-k)!(g-h)!(m-g-h)!}
\]

\[
q(i, j, j_2, k_1, k_2) = p_0^j p_1^{j_1} p_2^{k_1} q_0^{k_2} q_1^{k_2}
\]

and \( j_0 = m - j_1 - j_2, k_0 = m - k_1 - k_2 \), while \( C_{i_1,j_1,j_2,k_1,k_2} \) is same as described in
(2.48) and (2.49), and

\[
p_0 = \frac{8}{(2p(U_1)\lambda T_m + 3 - \sqrt{\frac{4p(U_1)\lambda T_m}{m} + 1})(2p(U_2)\lambda T_m + 3 + \sqrt{\frac{4p(U_2)\lambda T_m}{m} + 1})},
\]

\[
p_1 = \frac{p(U_1)\lambda T_m \sqrt{\frac{4p(U_1)\lambda T_m}{m} + 1 + 1}}{(2p(U_1)\lambda T_m + 3 - \sqrt{\frac{4p(U_1)\lambda T_m}{m} + 1})\sqrt{\frac{4p(U_1)\lambda T_m}{m} + 1}},
\]

(2.54)

\[
p_2 = \frac{p(U_2)\lambda T_m \sqrt{\frac{4p(U_2)\lambda T_m}{m} + 1 - 1}}{(2p(U_1)\lambda T_m + 3 + \sqrt{\frac{4p(U_1)\lambda T_m}{m} + 1})\sqrt{\frac{4p(U_1)\lambda T_m}{m} + 1}}
\]
and

\[ q_0 = \frac{8}{(2p(U_2)^\lambda T + 3 - \sqrt{4p(U_2)^\lambda T + 1}) (2p(U_2)^\lambda T + 3 + \sqrt{4p(U_2)^\lambda T + 1})}, \]

\[ q_1 = \frac{p(U_2)^\lambda T}{(2p(U_2)^\lambda T + 3 - \sqrt{4p(U_2)^\lambda T + 1}) \sqrt{4p(U_2)^\lambda T + 1}}, \quad (2.55) \]

\[ q_2 = \frac{p(U_2)^\lambda T}{(2p(U_2)^\lambda T + 3 + \sqrt{4p(U_2)^\lambda T + 1}) \sqrt{4p(U_2)^\lambda T + 1}}, \]

and

\[ x_1 = \frac{2p(U_1)^\lambda T + 3 - \sqrt{4p(U_1)^\lambda T + 1}}{2}, \]

\[ x_2 = \frac{2p(U_1)^\lambda T + 3 + \sqrt{4p(U_1)^\lambda T + 1}}{2}, \quad (2.56) \]

\[ y_1 = \frac{2p(U_1)^\lambda T + 3 - \sqrt{4p(U_1)^\lambda T + 1}}{2}, \]

\[ y_2 = \frac{2p(U_1)^\lambda T + 3 + \sqrt{4p(U_1)^\lambda T + 1}}{2}. \]

2.8 Conclusion

The main objective of this chapter is to provide an approach of optimally discretizing a continuous random process. The main result Theorem 1 is very general. Application of the main result to special random processes are made to yield optimal approximations of the Black-Scholes's formula and Merton's formula. In the former case, explicit multinomial approximations can be given, for the later, explicit approximations can only be given in some special cases.
3.1 Introduction

Recent attempts to model the term structure of interest rates generally fall into two categories. One is general equilibrium approach which was pioneered by Cox, Ingersoll and Ross ([21], [22]) and followed by many other authors. The other is the non-arbitrage approach pioneered by Ho and Lee [42]. Ho-Lee's model, which was presented in the form of a binomial tree for discount bond prices, provides an exact fit to the current term structure of interest rates. Their model was later significantly generalized by Heath, Jarrow and Morton [40] to a general setting which allows the entire forward rate curve to depend on multiple continuous stochastic factors. Their earlier work [39] gives a discrete time approach to their continuous time framework. In particular, they found the continuous limit of the Ho-Lee model (also see Hull and White [50]). An alternative to the Ho-Lee model was proposed by Black, Derman, and Toy [7], who use a binomial tree to construct a one-factor model of the short rate
that fits the current volatilities of all discount bond yields as well as the current term structure of interest rates. The Black, Derman and Toy Model was further developed by Hull and White ([44], [50]) to treat more general cases. A trinomial one factor discrete time term structure was considered by Robert and Ehud [67] who pointed out some limitations of the Ho-Lee model and empirically supported the appropriateness of a trinomial model. Unfortunately, their trinomial model is not path independent. Hull and White [44] used trinomial trees which are path independent to construct short rates which are consistent with the initial term structure and initial volatilities. But Hull-White's trinomial tree has no resemblance to the binomial tree in the Ho-Lee model, since the trinomial tree only models the short rate. The longer rate movements are determined under the assumptions that there is no term premium involved (i.e., one period returns are the same for all maturities, which is not the case in the Ho-Lee model).

In general, a multinomial model describes reality more closely than a binomial one does. Suppose that the trading time intervals are fixed (this is usually the case since data is collected in discrete time intervals), then it would be too simplistic to assume that the state of the world could only have two outcomes (up or down) in the next period. Hence in such a situation, an \( N \)-nomial model should be preferred to a binomial model. On the other hand when a model is generalized to the mutinomial one, the path independence property is usually lost because of technical difficulties. The path independence property of a model is important for two reasons. One is that the
efficient market hypothesis implies path independence property for many securities, the other is the computational efficiency. Non-path-independent models are usually computationally inefficient, especially for multinomial ones.

This chapter is to present a $N$-nomial generalization of Ho-Lee's model (we may call it $N$-nomial Ho-Lee model) under the same framework, i.e., it is arbitrage free, path independent and consistent with the initial term structure of interest rates. It turns out that such a generalization has similar "perturbation functions" as in the original Ho-Lee model. This model is the first $N$-nomial model in literature that has path-independence property. We also obtained a non-arbitrage pricing formula for interest rate sensitive contingent claims in a general $N$-nomial environment. Applying this formula to "$N$-nomial Ho-Lee model" results in a path independent $N$-nomial pricing formula for any contingent claims, generalizing the binomial one of Ho-Lee’s. Continuous limit of the model is also considered. It turns out that the limit found by Heath-Jarrow-Merton [39] for the binomial Ho-Lee model is a special case of the limit we found.

This chapter is organized as follows. Section 2 lists some basic assumptions on the market, describes the $N$-nomial discount bond vector movements and gives an arbitrage free condition of such movements. Section 3 gives complete descriptions of the $N$-nomial generalization of Ho-Lee model. Section 4 discusses the implications of the model on the short rate and the term premium. Section 5 proves a non-arbitrage
pricing formula for interest contingent claims in a general \( N \)-nomial environment and applies it to the "\( N \)-nomial Ho-Lee model". Section 6 gives the continuous limit of the \( N \)-nomial model. Section 7 concludes the paper.

### 3.2 Non-Arbitrage \( N \)-Nomial Bond Vector Movements

This section sets up the analytical framework of \( N \)-nomial term structure movements. We first list the basic assumptions for the bond market as in Ho-Lee [42], then describe the \( N \)-nomial bond vector movements which are arbitrage free.

#### 3.2.1 The Basic Assumptions

As in Ho-Lee model, we assume the following:

A1) markets are frictionless and all securities are perfectly divisible,

A2) riskless pure discount bonds are available for all maturities,

A3) at each time there are finite number of states of the world, a term structure is determined completely in each state of the world. Each state is followed by \( N \) new possible states in the next period of time.
3.2.2 The $N$-Nomial Lattice

We now define a discount bond vector (or simply bond vector) to be an infinite dimensional vector

$$ b = (1, b_1, b_2, \ldots, b_n, \ldots) $$

where each $b_i$ stands for the price of a bond with maturity $i$ trading periods. Note that we include 1 as the first component for convenience of exposition, it is the face value of a discount bond (after normalization) at its maturity. A bond vector completely characterizes the term structure of interest rate at a given time and state.

For time $t = i$ and state $j$, suppose the discount bond vector is

$$ b_{i,j} = (1, b_{i,j}(1), b_{i,j}(2), \ldots, b_{i,j}(n), \ldots) $$

then in the next period, i.e., $t = i + 1$, there are $N$ possible states,

$$ b_{i+1,jN}, b_{i+1,jN+1}, \ldots, b_{i+1,jN+N-1} $$

where

$$ b_{i+1,Nj+k} = (1, b_{i+1,Nj+k}(1), b_{i+1,Nj+k}(2), \ldots, b_{i+1,Nj+k}(n), \ldots), \quad 0 \leq k < N $$

So the term structure movement is described by an $N$-nomial tree. Without loss of generality, we assume the tree is in non-combining form (we do allow identical subtrees to exist). Node $(i, j)$ is attached with the bond vector $b_{i,j}$ which describes the term structure at time $t = i$ and state $j$. 
3.2.3 The Non-Arbitrage Condition

We now describe the non-arbitrage condition for the $N$-nomial bond vector movements. It is a special case of more general non-arbitrage theory (see Duffie [28], or Harrison and Kreps [36], [37]). The simple treatment given here serves both for pedagogical purpose and for completeness of this paper. It also paves the way for proving a pricing formula of interest rate sensitive contingent claims in a general $N$-nomial setting. We first define portfolio and strategy in the $N$-nomial environment as follows:

A portfolio at time $t = i$ and state $j$, denoted by $\theta_{i,j}$, is defined to be a vector

$$\theta_{i,j} = (0, \theta_{i,j}(1), \theta_{i,j}(2), \ldots, \theta_{i,j}(n), \ldots)$$

where $\theta_{i,j}(n)$ are real numbers which represent the number of shares of bond $b_{i,j}(n)$ for $n > 0$ and there are only finitely many bonds in any portfolio vector, i.e., $\theta_{i,j}(n)$ is nonzero for only finitely many $n$.

A trading strategy tree of length $K$ is defined to be a finite $N$-nomial tree with time-state node $(i, j)$ associated with a holding portfolio $\theta_{i,j}$ for $0 \leq i \leq K$ and $0 \leq j < N^i$.

For any vector $b = (b_0, b_1, b_2, \ldots, b_n, \ldots)$, we define $b^* = (b_1, b_2, \ldots, b_n, \ldots)$. Then for a bond vector tree $B = \{b_{i,j}|i = 0, 1, 2, \ldots; 0 \leq j < N\}$ and a portfolio tree $\theta = \{\theta_{i,j}|0 \leq i \leq K; 0 \leq j < N^i\}$, we define the payoff tree $P = \{p_{i,j}|0 \leq i \leq K; 0 \leq j < N^i\}$ as follows: for node-(0,0), the payoff is $p_{0,0} = -\theta_{0,0} \cdot b_{0,0}$ which is simply the cost for establishing the portfolio $\theta_{0,0}$; for node-$(i, j)$ with $1 \leq i < K$, let its
predecessor be node-\((i - 1, k)\), then the payoff is defined to be \(p_{i,j} = (\theta^*_{i-1,k} - \theta_{i,j}) \cdot b_{i,j}\) which is difference between the values of the old portfolio and the new portfolio at time \(t = i\). For \(i = K\), the trader only sells his old portfolio without buying a new one, the payoff is therefore \(p_{K,j} = \theta_{K,j} \cdot b_{K,j}\) at node-\((K, j)\).

**Definition of Non-Arbitrage:** The \(N\)-nomial tree of bond vector movements is arbitrage free if and only if there exists no strategy tree of any length \(K\) such that all nodes in its payoff tree have nonnegative payoff and there is at least one node which has positive payoff.

**Lemma 12** The \(N\)-nomial bond vector tree is arbitrage free if and only there is an \(N\)-nomial tree of positive numbers \(M = \{m_{i,j}, i = 0, 1, \ldots ; j = 0, 1, \ldots, N^1\}\) such that

\[
b^*_{i,j} = \sum_{k=1}^{N-1} m_{i+1,N_j+k} b_{i+1,N_j+k}
\]

for all \((i, j)\), i.e., at each node \((i, j)\), \(b^*_{i,j}\) is a weighted average of the bond vectors in the \(N\) successor nodes.

**Proof:** For each strategy tree \(\theta\), define the payoff vector \(v(\theta)\) to be the vector which has all the payoffs of the payoff tree of \(\theta\) as its components. By varying the strategy \(\theta\), we get a vector subspace \(V\) of the Euclidean space of dimension \(D = \frac{N^{K+1} - 1}{N - 1}\). By the definition of non-arbitrage, there can be no positive payoff vector (a positive vector means that all of its components of are nonnegative, but at least one of its
them is positive), so the vector subspace \( V \) of \( R^D \) only intersect with the nonnegative cone \( R^D_+ \) (all nonnegative vectors in \( R^D \)) at the origin. By Separating Hyperplane Theorem, \( V \) and \( R^D_+ \) only intersect at the origin if and only if there is a vector \( u \) in \( R^D_+ \) with strictly positive components such that \( u \) is orthogonal to \( V \), i.e., for all \( v(\theta) \) in \( V \)

\[
v(\theta) \cdot u = 0
\]  

(3.2)

where the multiplication is inner product of two vectors. We can write the components of \( u \) as a tree similar to the strategy tree or payoff tree, i.e. \( u = \{ u_{i,j} | 0 \leq i \leq K; 0 \leq j < N^t \} \). Consider the payoff \( p_{i,j} \) in the node-(i, j) in the payoff tree, it has \( N \) successors \( p_{i+1,Nj+k}, 0 \leq k \leq N - 1 \). Let its predecessor be \( p_{i-1,s} \), then by definition

\[
p_{i,j} = (\theta_{i-1,s}^* - \theta_{i,j}) \cdot b_{i,j}
\]

\[
p_{i+1,Nj+k} = (\theta_{i,j}^* - \theta_{i,Nj+k}) \cdot b_{i+1,Nj+k}
\]

Since \( \theta_{i,j} \cdot b_{r,s} = \theta_{i,j}^* b_{r,s}^* \) for any \( r, s \) and \( i, j \), we can write \( v(\theta) \cdot u \) as linear combinations of the \( \theta_{i,j}^* \)'s. Since \( \theta_{i,j} \) is arbitrary, the coefficient of \( \theta_{i,j}^* \) must be zero, which yields the following:

\[
u_{i,j}b_{i,j}^* - \sum_{k=0}^{N-1} u_{i+1,Nj+k}b_{i+1,Nj+k} = 0
\]  

(3.3)

Dividing (3.3) by \( u_{i,j} \) and letting

\[
m_{i+1,Nj+k} = \frac{u_{i+1,Nj+k}}{u_{i,j}}
\]

then (3.1) follows. Since the tree is in non-combining form, it is easy to check that (3.1) also implies (3.3). Q.E.D.
3.3 The Path-Independent $N$-nomial Bond Vector Movements

This section sets up the path-independence condition for the $N$-nomial bond vector movements. After introducing similar "perturbation functions", as in Ho-Lee [42], the path independence condition are characterized by a set of difference equations satisfied by the "perturbation functions". These difference equations are then solved completely to give explicit formulas for the "perturbation functions".

3.3.1 The Path Independence Condition

Let $b_{i,j}(n)$ be the discount bond with maturity $n$ at time-state $(i,j)$. In the subsequent period, it becomes a discount bond of maturity $n - 1$ in each of the $N$ possible states, i.e.,

\[ b_{i+1,Nj}(n-1), b_{i+1,Nj+1}(n-1), \ldots, b_{i+1,Nj+N-1}(n-1) \]

We now define the path independence condition for the $N$-nomial bond vector movements as follows: the $N$-nomial bond vector movements are called path independent if for any integers $k_1$, $s_1$ and $k_2$, $s_2$ in between 0 and $N - 1$ such that $k_1 + s_1 = k_2 + s_2$, the following equation holds for any time-state $(i,j)$,

\[ b_{i+2,N(Nj+k_1)+s_1}(n-2) = b_{i+2,N(Nj+k_2)+s_2}(n-2) \quad (3.4) \]

If we think state $j$ as the result of $j$ jumps, then (3.4) can be explained as follows: starting from any time-state $(i,j)$, $k_1$ jumps followed by $s_1$ jumps has the same effect on the discount bond as it does $k_2$ jumps followed by $s_2$ jumps, as long as the total
number of jumps are the same, i.e., \( k_1 + s_1 = k_2 + s_2 \). (3.4) ensures that a discount bond evolves in such a way that its future value only depends on the total number of jumps it experienced in the time span between present and future, but not on the sequence in which jumps occur.

Remark: The path independence is defined for non-combining \( N \)-nomial trees, but when \( N \)-nomial non-combining tree satisfies path independence condition, it can be recombined into a combining tree in an obvious way. Conversely, any \( N \)-nomial combining tree can be written as a \( N \)-nomial non-combining tree with many identical subtrees.

### 3.3.2 The “Perturbation Functions”

We now describe the path independence condition (3.4) by a set of “perturbation functions”.

Comparing the first components on both sides of (3.1), we have

\[
\sum_{k=0}^{N-1} m_{i+1,Nj+k} = b_{i,j}(1)
\]

Let

\[
\frac{m_{i+1,Nj+k}}{b_{i,j}(1)} = a_{i,j}(k)
\]

then

\[
m_{i+1,Nj+k} = a_{i,j}(k)b_{i,j}(1) \quad 0 \leq k \leq N - 1
\]

\[
a_{i,j}(0) + a_{i,j}(1) + \ldots + a_{i,j}(N - 1) = 1
\]
(3.1) and the first equation of (3.6) imply the following:

\[
\frac{m_{i+1,N_j+k}}{b_{i,j}(1)} = a_{i,j}(k) 
\] (3.7)

We may call \( \{a_{i,j}(k), 0 \leq k \leq N - 1\} \) the risk neutral probabilities of the bond vector movements at time-state \((i,j)\). They are independent of \(n\), but in general may be dependent upon the time-state \((i,j)\) and the index \(k\). As in the Ho-Lee model, we assume that the \(a_{i,j}(k)\)'s depend on \(k\) only, i.e., they are independent of the time-state \((i,j)\), then, we may rewrite (3.5) as

\[
\frac{m_{i+1,N_j+k}}{b_{i,j}(1)} = a_k 
\] (3.8)

From (3.1) and (3.8), we have, for any \(n > 1\), that

\[
\frac{b_{i,j}(n)}{b_{i,j}(1)} = \sum_{k=0}^{N-1} a_k b_{i+1,N_j+k}(n - 1) 
\] (3.9)

We further write

\[
b_{i+1,N_j+k}(n - 1) = \frac{b_{i,j}(n)}{b_{i,j}(1)} h_k(n - 1) 
\] (3.10)

where, similar to Ho-Lee's model, the \(h_k()\)'s are assumed to be only dependent upon \(n\) and \(k\) but not on time-state \((i,j)\). These \(h_k()\)'s are called "perturbation functions" and are assumed to satisfy \(h_0(n) < h_1(n) < \ldots < h_{N-1}(n)\) for all \(n\). Then (3.9) becomes

\[
\frac{b_{i,j}(n)}{b_{i,j}(1)} = \sum_{k=0}^{N-1} a_k \frac{b_{i,j}(n)}{b_{i,j}(1)} h_k(n - 1) 
\]

which yields

\[
\sum_{k=0}^{N-1} a_k h_k(n - 1) = 1 
\] (3.11)
Ho and Lee [42] observe that in an economy without uncertainty and arbitrage opportunities, the perturbation function $h_k(\cdot)$ must equals 1. (3.10) implies that in a world without uncertainty, the next period discount function equals to the present period forward discount function. Note that the $\alpha$'s are risk neutral probabilities of the perturbation functions. We also assume that

$$h_k(0) = 1 \quad 0 \leq k \leq N - 1$$

(3.12)
since at its maturity there is no uncertainty involved in the discount functions.

With above notations and assumptions, the path independence condition (3.4) now becomes

$$\frac{b_{i,j}(n) h_{k_1}(n-1) h_{s_1}(n-2)}{b_{i,j}(1) h_{k_1}(1)} = \frac{b_{i,j}(n) h_{k_2}(n-1) h_{s_2}(n-2)}{b_{i,j}(1) h_{k_2}(1)}$$

or simply

$$\frac{h_{k_1}(n-1) h_{s_1}(n-2)}{h_{k_1}(1)} = \frac{h_{k_2}(n-1) h_{s_2}(n-2)}{h_{k_2}(1)}$$

(3.13)
for any $n > 2$ and $k_1, k_2$ and $s_1, s_2$ in between 0 and $N - 1$ such that $k_1 + s_1 = k_2 + s_2$.

### 3.3.3 Solutions of the “Perturbation Functions”

We now solve (3.13) completely for the perturbation functions $h_k(\cdot)$ for any $0 \leq k \leq N - 1$.

Choosing $k_1 = k + 1$, $k_2 = k$ and $s_1 = k$, $s_2 = k + 1$ in (3.13) and rearranging the equation gives

$$\frac{h_{k+1}(n-1)}{h_k(n-1)} = \frac{h_{k+1}(n-2)}{h_k(n-2)} \frac{h_{k+1}(1)}{h_k(1)}$$

(3.14)
for each $k$ such that $0 \leq k \leq N - 1$.

Again, choosing $k_1 = N - 1$, $k_2 = N - 2$ and $s_1 = k$, $s_2 = k + 1$ in (3.13) and rearranging the equation gives

$$\frac{h_{N-1}(n-1)}{h_{N-2}(n-1)} = \frac{h_{k+1}(n-2)h_{N-1}(1)}{h_k(n-2)h_{N-2}(1)}$$

which implies that $\frac{h_{k+1}(n-2)}{h_k(n-2)}$ is independent of $k$. In particular

$$\frac{h_{k+1}(n-2)}{h_k(n-2)} = \frac{h_k(n-2)}{h_{k-1}(n-2)}$$ (3.15)

for every $k$ such that $0 \leq k \leq N - 2$, or equivalently

$$h_k(n-2)^2 = h_{k-1}(n-2)h_{k+1}(n-2)$$ (3.16)

We now proceed to solve $h_k(\cdot)$ explicitly from (3.14),(3.15) and (3.16). From (3.14), we have

$$\frac{h_1(n-1)}{h_0(n-1)} = \frac{h_1(n-2)h_1(1)}{h_0(n-2)h_0(1)}$$

$$\frac{h_2(n-1)}{h_1(n-1)} = \frac{h_2(n-2)h_2(1)}{h_1(n-2)h_1(1)}$$

$$\ldots \ldots$$

$$\frac{h_k(n-1)}{h_{k-1}(n-1)} = \frac{h_k(n-2)}{h_{k-1}(n-2)h_{k-1}(1)}$$

Multiplying these equations side by side to get

$$\frac{h_k(n-1)}{h_0(n-1)} = \frac{h_k(n-2)}{h_0(n-2)}\frac{h_k(1)}{h_0(1)}$$

or

$$h_k(n-1) = h_0(n-1)\frac{h_k(n-2)h_k(1)}{h_1(n-2)h_0(1)}$$ (3.17)
for $1 \leq k \leq N - 1$. Substitute equations in (3.17) into (3.11) and solve for $h_0(n - 1)$ to get

$$h_0(n - 1) = \frac{h_0(n - 2)h_0(1)}{\sum_{s=0}^{N-1} a_s h_s(n - 2)h_s(1)}$$

(3.18)

Now substitute (3.18) back into (3.17) to get

$$h_k(n - 1) = \frac{h_k(n - 2)h_k(1)}{\sum_{s=0}^{N-1} a_s h_s(n - 2)h_s(1)} \quad 0 \leq k \leq N - 1$$

(3.19)

Finally, applying (3.19) recursively for $n = 3, 4, 5, \ldots$, we get the following lemma:

**Lemma 13** The $h_k()$'s defined by (3.10) can be written as follows:

$$h_k(n) = \frac{h_k(1)^n}{\sum_{s=0}^{N-1} a_s h_s(1)^n} \quad \text{for } n > 1 \text{ and } 0 \leq k \leq N - 1$$

(3.20)

where the $h_s(1)$'s are restricted by (3.16), i.e.,

$$h_s(1)^2 = h_{s-1}(1)h_{s+1}(1) \quad 0 \leq s \leq N - 2$$

(3.21)

and

$$a_0 + a_1 + \ldots + a_{N-1} = 1$$

(3.22)

Proof: We will prove the lemma by induction on $n$. First, (3.20) with $n = 2$ is equivalent to (3.19) with $n = 3$. Now suppose (3.20) is true for some $n = m \geq 2$, i.e.,

$$h_k(m) = \frac{h_k(1)^m}{\sum_{s=0}^{N-1} a_s h_s(1)^m} \quad 0 \leq k \leq N - 1$$

(3.23)

Let $n = m + 2$ in (3.19) to get

$$h_k(m + 1) = \frac{h_k(m)h_k(1)}{\sum_{s=0}^{N-1} a_s h_s(m)h_s(1)} \quad 0 \leq k \leq N - 1$$

(3.24)
Now substitute (3.23) into (3.24) and simplify the result to get

$$h_k(m + 1) = \frac{h_k(1)^{m+1}}{\sum_{s=0}^{N-1} a_s h_s(1)^{m+1}}$$

which shows (3.20) holds for $n = m + 1$, hence for all $n$ by induction. Q.E.D.

We note that when $N = 2$, (3.20) is exactly the as Ho-Lee's formula.

3.4 The $N$-nomial Model

In this section, the non-arbitrage $N$-nomial bond price movement is explicitly described using the “perturbation functions” described in the last section.

Suppose at time $t = 0$, the term structure is given by a single discount bond vector $b_{0,0} = (1, b_1, b_2, b_3, \ldots)$, where $b_n = P(0, n)$ is a discount bond with maturity $n$, and suppose during next $t$ periods, the jumps are $j_1, j_2, \ldots, j_t$ successively, let $P(t, n)$ be the discount bond evolved from $b_n = P(0, n)$ at time $t$, then it has maturity $n - t$.

By applying (3.10) repeatedly, we get the following formula for $P(t, n)$:

$$P(t, n) = \frac{P(0, n)}{P(0, t)} \prod_{s=1}^{t} \frac{h_{j_s}(n - s)}{h_{j_s}(t - s)}$$

(3.25)

where the $h_k()$'s are given by (3.20).

In order to write (3.25) in a more explicit form, let

$$\frac{h_k(1)}{h_{N-1}(1)} = e^{-c_k} \quad \text{for } 0 \leq k \leq N - 1$$

(3.26)
Since \( h_0(1) < h_1(1) < \ldots < h_{N-1}(1) \) by assumption, we must have

\[
0 = c_{N-1} < c_{N-2} < \ldots < c_1 < c_0
\] (3.27)

and condition (3.21) implies that \( c_k - c_{k+1} \) is a constant, say \( c (= c_{N-2}) \), hence

\[c_{N-1-k} = kc \quad 0 \leq k \leq N - 1\]

and

\[
\frac{h_{N-1-k}(1)}{h_{N-1}(1)} = e^{-ck}
\]

Combining the above equation with (3.20), we have

\[
\begin{align*}
\frac{h_{N-1-k}(n)}{h_{N-1}(1)} &= \frac{(h_{N-1-k}(1))_n}{\sum_{s=0}^{N-1} a_{N-1-s} (h_{N-1-s}(1))_n} \\
&= \frac{\sum_{s=0}^{N-1} e^{-ckn}}{\sum_{s=0}^{N-1} a_{N-1-s} e^{-csn}},
\end{align*}
\] (3.28)

for \( 0 \leq k \leq N - 1 \). For notational convenience, we let

\[d_s = a_{N-1-s} \quad \text{for } s = 0, 1, \ldots, N - 1,\]

then from (3.28), we have

\[
\begin{align*}
\frac{h_{j_s}(n-s)}{h_{j_s}(t-s)} &= \frac{\sum_{s=0}^{N-1} a_s e^{-c(t-s)} e^{-c(N-1-j_s)(n-s)}}{\sum_{s=0}^{N-1} a_s e^{-c(N-1-j_s)(t-s)}} \\
&= \frac{\sum_{s=0}^{N-1} a_s e^{-c(t-s)}}{\sum_{s=0}^{N-1} a_s e^{-c(N-1-j_s)(n-t)}} e^{-c(N-1-j_s)(n-t)} \\
&= \frac{\sum_{s=0}^{N-1} a_s e^{-c(t-s)}}{\sum_{s=0}^{N-1} a_s e^{-c(n-s)}} e^{-c(n-t)\eta_s} \quad (3.29)
\end{align*}
\]

where \( \eta_s = N - 1 - j_s \).
Substituting (3.29) and (3.25) gives the following (to conform with convention, we use $T$ in replacing $n$ for maturity):

$$P(t, T) = \frac{P(0, T)}{P(0, t)} \left( \prod_{s=1}^{t} \frac{\sum_{k=0}^{N-1} d_k e^{-ck(t-s)}}{\sum_{k=0}^{N-1} d_k e^{-ck(T-s)}} \right) e^{-c(T-t) \sum_{k=1}^{t} \eta_k} \quad (3.30)$$

We can view the $\eta_k$'s as i.i.d. random variables which assume values $0, 1, 2, \ldots, N-1$ with positive probabilities, then (3.30) clearly shows the path independence property of the discount functions, i.e., their terminal values depend on each individual jump only through sum of the jumps $\sum_{k=1}^{t} \eta_k$.

### 3.5 Implications of the $N$-nomial Ho-Lee Model

In this section, we analyze the implications of the $N$-nomial Ho-Lee model described in last section. First we show what the $N$-nomial Ho-Lee model implies about the short rate movement, then we discuss what this model implies about the local expectations hypothesis and the term premium.

#### 3.5.1 The Short Rate

The short rate, within the context of our generalized Ho-Lee model, is still the rate of a one-period discount bond. At time $t$, the one-period discount bond price is given by (3.30) with $T = t + 1$, i.e.,

$$P(t, t + 1) = \frac{P(0, t + 1)}{P(0, t)} \left( \prod_{s=1}^{t} \frac{\sum_{k=0}^{N-1} d_k e^{-ck(t-s)}}{\sum_{k=0}^{N-1} d_k e^{-kc(t+1-s)}} \right) e^{-c(t+1-t) \sum_{k=1}^{t} \eta_k} \quad (3.31)$$

$$= \frac{P(0, t + 1)}{P(0, t)} \frac{e^{-c \sum_{k=1}^{t} \eta_k}}{\sum_{k=0}^{N-1} d_k e^{-kc}}$$
So the one-period rate \( r(t, t + 1) \), is
\[
 r(t, t + 1) = -\ln P(t, t + 1)
\]
\[
= \ln \frac{P(0, t)}{P(0, t + 1)} + \ln \sum_{k=0}^{N-1} d_k e^{-ckt} + c \sum_{k=1}^{t} \eta_k
\]  

(3.32)

Now we assign probabilities to the \( N \)-nomial lattice. We assume that each \( \eta_k \) can attain values 0, 1, \ldots, \( N - 1 \) with probabilities \( q_0, q_1, \ldots, q_{N-1} \) respectively. Then from (3.32) it follows that \( r(t, t + 1) \) is a \( N \)-nomial distribution. To describe this distribution, we introduce the following notations (some of them are reserved for use in later sections):

\[
E_1(q) = \sum_{k=0}^{N-1} kq_k, \quad E_1(d) = \sum_{k=0}^{N-1} kd_k
\]
\[
E_2(q) = \sum_{k=0}^{N-1} k^2q_k, \quad E_2(d) = \sum_{k=0}^{N-1} k^2d_k
\]
\[
V(q) = E_2(q) - (E_1(q))^2, \quad V(d) = E_2(d) - (E_1(d))^2
\]  

(3.33)

which are the first moments, second moments and variances of the \( \eta_k \) under the true probabilities and risk neutral probabilities.

With the above notations, the mean \( \mu \) and variance \( \sigma^2 \) of the short rate \( r(t, t + 1) \) are given by

\[
\mu = \ln \frac{P(0, t)}{P(0, t + 1)} + \ln \sum_{k=0}^{N-1} d_k e^{-ckt} + ct E_1(q)
\]  

(3.34)

and

\[
\sigma^2 = e^{2t} V(q)
\]  

(3.35)

(3.34) shows that the expected forward short rate is the implied forward short rate (the first term) plus a bias (sum of the second and third terms). This feature distinguishes the Ho-Lee model from other one factor term structure models which try to determine a short rate process that can generate a non-arbitrage or equilibrium term structure.
3.5.2 Local Expectations Hypothesis and the Term Premium

Similar to the binomial case in Ho-Lee [42], we can calculate the term premium of the $N$-nomial model. The $T$-period term premium is the expected return over one period of a $T$-period bond in excess of the one-period bond return. If there is no term premium, then we say the local expectations hypothesis holds.

**Proposition 5** The $T$-period term premium is given by

$$
\tau = \frac{1}{P(0, 1)} \left( \frac{\sum_{k=0}^{N-1} q_k e^{-ck(T-1)}}{\sum_{k=0}^{N-1} d_k e^{-ck(T-1)}} - 1 \right)
$$

and the local expectations hypothesis holds if and only if $q_k = d_k$ for for $k = 0, 1, \ldots, N - 1$, i.e., the risk neutral probabilities are the same as the true probabilities.

**Proof:** Taking the expectation of (3.27) with $t = 1$ and dividing it by $P(0, T)$ gives the rate of return of a $T$-period bond over one period as follows:

$$
\frac{1}{P(0, 1)} \left( \frac{\sum_{k=0}^{N-1} q_k e^{-ck(T-1)}}{\sum_{k=0}^{N-1} d_k e^{-ck(T-1)}} \right)
$$

Since the one-period bond rate of return is $\frac{1}{P(0, 1)}$, taking the difference between the two rates of return gives the desired result. It is easy to see that the term premium is zero for all $T$ if and only if $q_k = d_k$ for all $0 \leq k \leq N - 1$. Q.E.D.

3.6 Pricing of Interest Contingent Claims

This section gives an $N$-nomial risk-neutral pricing formula for interest rate contingent claims under the assumption that the bond vector movement follows a non-arbitrage
N-nomial process. Note that we do not require that the bond vector movement is path independent. So the pricing formula given in this section applies to virtually any type of interest rate contingent claims. This is a significant improvement upon Ho-Lee's binomial formula which only applies to those class of contingent claims which have path independence properties.

Let \( C \) be an interest contingent claim with maturity \( T \). We require its price \( C(T,j) \) are known for all \( 0 < j < NT \). then we have to following:

**Proposition 6 (N-Nomial Risk-Neutral Pricing Formula):** Consider any interest rate contingent claim \( C(i,j) \) that can be bought and sold in a frictionless bond market described by \( (A1)-(A3) \). Assume that the bond vector follows a non-arbitrage \( N \)-nomial process with risk neutral probabilities \( \{a_{i,j}(k)\} \), then the \( C(i,j) \) are uniquely determined under the condition that there is no arbitrage to be realized in holding bonds, and the following equation must hold:

\[
C(i,j) = \frac{1}{b_{i,j}(1)} \sum_{k=0}^{N-1} a_{i,j}(k)(C(i+1,jN+k) + X(n+1,jN+k))
\]  

(3.37)

where \( 0 \leq i \leq T - 1, 0 \leq j \leq N^i \) and \( b_{i,j}(1) \) is the one-period discount bond price at state-time \( (i,j) \). Hence the \( C(i,j)'s \) can be solved by backward recursion from the above equation since we know each \( C(T,j) \) for every \( j \).

**Proof:** The proof uses similar argument as in section 2. Consider a non-combining \( N \)-nomial tree of which each node corresponds to a time-state pair \( (i,j) \). At each time-state \( (i,j) \), the available securities are the bond vector and the contingent claim. We
write them in a single vector called the security vector which is denoted by $s_{i,j} = (b_{i,j}, C(i,j))$ where $C(i,j)$ is the value of the contingent claim and $b_{i,j}$ is the bond vector as in section 2. A portfolio containing the contingent claim and some bonds can be written in a vector $f_{i,j} = (\theta_{i,j}, g_{i,j})$ where $g_{i,j}$ stands for the number of shares of the contingent claim and $\theta_{i,j}$ is a bond portfolio as in section 2. Strategy tree and payoff tree are also similarly defined as in section 2. The whole argument of section 2 is still valid and Lemma 1 holds with bond vectors replaced by security vectors, i.e., there is no arbitrage holding bonds and the contingent claim if and only if there exists a $N$-nomial tree of positive numbers $W = \{w_{i,j}|i = 0,1,2,\ldots; 0 \leq j \leq N^j\}$ such that

$$S_{i,j}^* = \sum_{k=0}^{N-1} w_{i+1,Nj+k} S_{i+1,Nj+k} \tag{3.38}$$

Comparing the bond part of equation (3.38) gives

$$b_{i,j}^* = \sum_{k=0}^{N-1} w_{i+1,Nj+k} b_{i+1,Nj+k} \tag{3.39}$$

Since the risk neutral probabilities $\{a_{i,j}(k)\}$ are known, the $m_{i,j}$'s in Lemma 1 are unique by (3.5), hence (3.1) and (3.39) implies that $m_{i,j} = w_{i,j}$ for all time-state $(i,j)$. Now comparing the contingent claim part of equation (3.38) gives

$$C(i,j) = \sum_{k=0}^{N-1} m_{i+1,Nj+k} b_{i+1,Nj+k} \tag{3.40}$$

Equation (3.37) now follows from (3.40) and (3.5). Q.E.D.

We note that when both the bond vector movements and the contingent claim are path independent, then after combining the identical subtrees, there are only $iN$
nodes in the bond tree at time $i > 0$, hence, after reindexing, (3.37) becomes

$$
C(i, j) = \frac{1}{b_{i,j}(1)} \sum_{k=0}^{N-1} a_{i,j}(k)(C(i + 1, j + k) + X(i + 1, j + k))
$$

(3.41)

where $0 \leq i \leq T - 1$ and $0 \leq j \leq iN - 1$.

### 3.7 Continuous Time Limit of the $N$-nomial Ho-Lee Model

In this section we find a continuous time limit of the $N$-nomial Ho-Lee model discussed in the previous section. The approach we use here is similar to that used in Heath-Jarrow-Merton [39], i.e., we find the forward discount functions first. We begin with some notational changes so that trading period could be arbitrary small (which is necessary for taking limit). In the previous discussion, the time $t$ and maturity $T$ (we have changed from $n$ to $T$) are measured by the number of trading periods. From now on, $t$ and $T$ will be measured by real time unit (whatever it is). We assume that there are $m$ trading periods in one unit time and each trading period has length $\Delta = 1/m$, then $t$ and $T$ contain $mt$ and $mT$ trading periods respectively. With this notational change, (3.30) becomes

$$
P(t, T) = \frac{P(0, T)}{P(0, t)} \prod_{s=1}^{mt} \sum_{k=0}^{N-1} d_k e^{-ck(mT-ms)} e^{-c(mT-mt) \sum_{s=1}^{mt} \eta_s} \tag{3.42}
$$

Given bond discount functions, the forward discount functions are determined. The forward rate at time $t$, for the time interval $[T, T + \Delta]$, $f_m(t, T)$ ($t \leq T$), is defined by

$$
f_m(t, T) = -\left| \log(P(t, T + \Delta)/P(t, T)) \right| / \Delta \tag{3.43}
$$
It follows from (3.42) and (3.43) that

\[ f_m(t, T) = -\frac{1}{\Delta} \left[ \log \left( \frac{P(0, T+\Delta)}{P(0, T)} \right) - \log \left( \prod_{s=1}^{m^t} \frac{\sum_{k=0}^{N-1} d_k e^{-k \lambda (mT-s)}}{\sum_{k=0}^{N-1} d_k e^{-k \lambda (mT+1-s)}} \right) - c \sum_{s=1}^{m^t} \eta_s \right] \]

\[ = -\frac{1}{\Delta} \log \frac{P(0, T+\Delta)}{P(0, T)} - \frac{1}{\Delta} \log \left( \frac{\sum_{k=0}^{N-1} d_k e^{-k \lambda (T-t)}}{\sum_{k=0}^{N-1} d_k e^{-k \lambda T}} \right) + \frac{1}{\Delta} c \sum_{s=1}^{m^t} \eta_s \] (3.44)

\[ = I_1 + I_2 + I_3 \]

By definition

\[ \lim_{\Delta \to 0} I_1 = f(0, T) \] (3.45)

Now we rewrite \( I_3 \) in the following form (note that \( \Delta = \frac{1}{m} \)):

\[ I_3 = (m^{\frac{3}{2}} c t \sqrt{V(q)}) \sum_{s=1}^{m^t} (\eta_s - E(q)) \frac{(mV(q))^{\frac{1}{2}}}{mV(q)^{\frac{1}{2}}} + m^2 c t E(q) \] (3.46)

By Central Limit Theorem, we have

\[ \lim_{m \to \infty} \frac{\sum_{s=1}^{m^t} (\eta_s - E(q))}{(mV(q))^{\frac{1}{2}}} = Z \] (3.47)

the standard normal distribution. Since only \( I_3 \) involves random variables, in order for \( f_m(t, T) \) to converge to a meaningful stochastic process, \( I_3 \) must converge in distribution, in particular the first term in (3.47) must converge in distribution to a meaningful stochastic process. But this can happen only if \( c \) is in the following form:

\[ c = \frac{\gamma_m}{n^{\frac{3}{2}}} \] (3.48)
where \( \gamma_m = \gamma + o(1) \) and \( \gamma \) is a constant independent of \( m \) and \( o(1) \) vanishes as \( m \to \infty \). With \( c \) chosen as in (3.48),

\[
I_2 = -m \log(1 + \frac{\sum_{k=0}^{N-1} d_k e^{-kmc(T-t)}(1-e^{-cmkt})}{\sum_{k=0}^{N-1} d_k e^{-cmkT}})
\]

\[
= -m \log(1 + \frac{\sum_{k=0}^{N-1} d_k e^{-\gamma m m^{-\frac{1}{2}}(T-t)}(1-e^{-\gamma m km^{-\frac{1}{2}} t})}{\sum_{k=0}^{N-1} d_k e^{-\gamma m km^{-\frac{1}{2}} T}})
\]

\[
= -m \log(1 + J)
\]

Using Taylor expansions, we get the following:

\[
J = \sum_{k=0}^{N-1} d_k (1 - \frac{2m k^2 (T-t)}{\sqrt{m}} + O(1))(\frac{2m k^2 (T-t)}{2m} + O(m^{-\frac{3}{2}}))
\]

\[
= \sum_{k=0}^{N-1} d_k (\frac{2m k^2 (T-t)}{\sqrt{m}} + O(m^{-\frac{3}{2}}))
\]

\[
= \frac{\gamma_m E_1(d)T}{\sqrt{m}} - \frac{\gamma_m^2 E_2(d)T^2}{2m} + O(m^{-\frac{3}{2}})
\]

Again using Taylor expansion, we have

\[
I_2 = -n \left(\frac{2m E_1(d)T}{\sqrt{m}} - \frac{2m^2 V(d)T^2}{m} + \frac{2m^2 E_2(d)T^2}{2m} - \frac{(2m E_1(d))^2 T^2}{2m} + O(m^{-\frac{3}{2}})\right)
\]

\[
= -n \left(\frac{2m E_1(d)T}{\sqrt{m}} - \frac{2m^2 V(d)T^2}{m} + \frac{2m^2 E_2(d)T^2}{2m} + O(m^{-\frac{3}{2}})\right)
\]

\[
= -\gamma_m E_1(d)T \sqrt{m} + \gamma_m^2 V(d)T - \frac{1}{2} \gamma_m^2 V(d)T^2 + O(m^{-\frac{3}{2}})
\]

Now write (3.46) explicitly with \( c \) as in (3.48)

\[
I_3 = \gamma_m \sqrt{V(q)} t \sum_{i=1}^{m^t} (\eta_i - E(q)) \frac{1}{\sqrt{nV(q)}} + \gamma_m E(q) t \sqrt{m}
\]
Now it is clear from (3.49) and (3.50) that $I_2 + I_3$ converges if only if

$$E_1(q) = E_1(d)$$

(3.51)

If (3.51) holds, then by (3.44), (3.45), (3.47), (3.49) and (3.50), the limit of $f_m(t, T)$ in distribution, which is denoted by $f(t, T)$ is

$$f(t, T) = f(0, T) + \gamma \sqrt{V(q)} tZ + \gamma^2 V(d) tT - \frac{1}{2} \gamma^2 V(d) t^2$$

(3.52)

We note that (3.51) implies $V(q) = V(d)$ for $N = 2$, but not for $N > 2$. We denote $\frac{V(q)}{V(d)} = \delta^2$ and denote $\sigma = \gamma \sqrt{V(d)}$, then (3.52) becomes

$$f(t, T) = f(0, T) + \delta \sigma W(t) + \sigma^2 (tT - \frac{1}{2} t^2)$$

(3.53)

where $W(t) = tZ$ is the standard Brownian random process. When $\delta = 1$, i.e., $V(q) = V(d)$, (3.53) is exactly the same as the continuous limit of the Ho-Lee model given in Heath-Jarrow-Merton [39].

From (3.43), we have

$$f(t, T) = - \frac{\partial \log P(t, T)}{\partial T}$$

hence we may integrate (3.52) to get $\log P(t, T)$ as follows:

$$\log P(t, T) = \log P(t, u)|_{u=t}^T$$

$$= - \int_t^T f(t, u) du$$

$$= - \int_t^T f(0, u) du - \delta \sigma (T - t) W(t) - \sigma^2 \int_t^T (tu - \frac{1}{2} t^2) du$$

$$= \log P(0, u)|_{u=t}^T - \delta \sigma (T - t) W(t) - \frac{\sigma}{2} t T (T - t)$$

$$= \log \frac{P(0, T)}{P(0, t)} - \delta \sigma (T - t) W(t) - \frac{\sigma}{2} t T (T - t)$$
hence

\[ P(t, T) = \frac{P(0, T)}{P(0, t)} \exp\left(-\delta(T - t)W(t) - \frac{\sigma}{2} t T(T - t)\right) \]  

(3.54)

3.8 Conclusion

We have found an \( N \)-nomial generalization of Ho-Lee model on the term structure of interest rates. Such a generalization retains all characteristics of the original Ho-Lee model. In particular, it is arbitrage free and path independent. The term structure movements are described by \( N \)-nomial bond vector movements which in turn are characterized by a set of "perturbation functions". These "perturbation functions" are determined by the path independence condition on the bond vector movements, and are natural extensions of the two-parameter "perturbation functions" described in Ho-Lee [42]. We also obtained an \( N \)-nomial pricing formula of interest rate contingent claims of any types (path dependent or path independent), which significantly generalized the corresponding formula of Ho-Lee's. The continuous limit of the \( N \)-nomial model are also found, which turns out to be more general than the limit of the binomial Ho-Lee model as found by Heath-Jarrow-Merton [39]. In binomial case, for the limit to exist, the true binomial random variable has to be the same as the risk neutral one in distribution, while in the \( N \)-nomial case, we only need to require that the true \( N \)-binomial random variable and the risk neutral one have the same expectations. This is not surprising after all since in binomial case, equal expectations imply equal distributions, but it is not the case for \( N \)-nomial random variables. This
explains why our limit formula is more general than the one found by Heath-Jarrow-Merton [39].

Empirical studies by Robert and Ehud [67] shows that a multi-nomial model, such as the one described in this chapter, provides more flexibilities and dimensions in describing the effects of real world changes on the interest rate movement, and could be used to fit empirical data more accurately. Furthermore, the computational efficiency is guaranteed by the path independence of the model.
CHAPTER IV

A NEW TERM STRUCTURE MODEL

4.1 Introduction

In this chapter, an arbitrage free term structure model is constructed. The distinct feature of this model is that it has infinitely many parameters (corresponding to infinitely many trading periods in the future) and the implied risk neutral probabilities are time dependent. Such a sequence of parameters provide flexibilities in controlling the term structure movements. For example, under some easily described conditions on the parameters, the interest rates are guaranteed to be positive for all time intervals.

Among others, The Ho-Lee model [42] and Black,Derman,Toy model [7] are most relevant to the one to be presented in this chapter. Ho-Lee's model takes the initial term structure as given and gives a closed-form solution for the discount factors in each time-state of the term structure tree, which makes computations speedy. But it is criticized because of negative interest rates it generates. Since the continuous limit of Ho-Lee model involves a random variable which is normally distributed, the
negativity of interest rates is unavoidable. The Black, Derman, Toy model, which falls into the category of lognormal model, corrects the shortcoming of negative interest rates in Ho-Lee model, but does not admit a closed form solution. Inefficient trial and error methods were recommended in various references to implement it before Jamshidian [53] introduced the forward induction method which is demonstrated to be fast and accurate.

The model presented in this chapter has merits of both Ho-Lee model and Black, Derman and Toy model. It has closed-form solutions for the discount factors in each time-state of the path independent term structure tree, and it generates positive interest rates at any time, state and maturity under appropriate choices of the parameters.

4.2 Construction of a Non-arbitrage Term Structure Vector Tree

Let the present time be $t = 0$, and term structure of interest rates is known and is characterized by a vector of unit zero coupon bonds with different maturities:

$$v_{0,0} = (1, P_1, P_2, P_3, \cdots)$$

i.e., $P_i$ is the present price of a zero coupon bond which pays $1$ when it matures at time $i$ periods from now.
We will consider how the whole term structure changes as time goes by. We assume that our expectation of the future changes of interest rates are characterized by a sequence of parameters:

$$\lambda_0(=1), \lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n, \ldots$$

where $0 < \lambda_i < 1$ for all $i > 0$.

At $t = 1$, there are two possibilities for the term structure, i.e.,

$$v_{1,0} = (1, \frac{\lambda_2 P_2}{\lambda_1 P_1}, \frac{\lambda_3 P_3}{\lambda_1 P_1}, \ldots, \frac{\lambda_n P_n}{\lambda_1 P_1}, \ldots)$$

and

$$v_{1,1} = (1, \frac{1 - \lambda_2 P_2}{1 - \lambda_1 P_1}, \frac{1 - \lambda_3 P_3}{1 - \lambda_1 P_1}, \ldots, \frac{1 - \lambda_m P_m}{1 - \lambda_1 P_1}, \ldots)$$

At time $t = 2$, $v_{1,0}$ becomes

$$v_{2,0} = (1, (\frac{\lambda_3}{\lambda_2})^2 P_3, (\frac{\lambda_4}{\lambda_2})^2 P_4, \ldots, (\frac{\lambda_m}{\lambda_2})^2 P_m, \ldots)$$

and

$$v_{2,1} = (1, \frac{\lambda_3 (1 - \lambda_3) P_3}{\lambda_2 (1 - \lambda_2) P_2}, \frac{\lambda_4 (1 - \lambda_4) P_4}{\lambda_2 (1 - \lambda_2) P_2}, \ldots, \frac{\lambda_m (1 - \lambda_m) P_m}{\lambda_2 (1 - \lambda_2) P_2}, \ldots)$$

and $v_{1,1}$ becomes $v_{2,1}$ and $v_{2,3}$ where

$$v_{2,3} = (1, (\frac{1 - \lambda_3}{1 - \lambda_2})^2 P_3, (\frac{1 - \lambda_4}{1 - \lambda_2})^2 P_4, \ldots, (\frac{1 - \lambda_m}{1 - \lambda_2})^2 P_m, \ldots)$$

For $n > 0$, $0 \leq k \leq n$ and $i \geq 0$, let

$$\alpha(n, k, i) = \lambda_{n+i}^{n-k}(1 - \lambda_{n+i})^k$$
\[ \beta(n, k, i) = \left( \frac{\lambda_{n+1}}{\lambda_n} \right)^{n-k} \left( 1 - \frac{\lambda_{n+1}}{\lambda_n} \right)^k = \frac{\alpha(n, k, i)}{\alpha(n, k, 0)} \]

It is easy to check the following property of the \( \alpha \)'s:

\[ \alpha(n, k, i) + \alpha(n, k + 1, i) = \alpha(n - 1, k, i + 1) \tag{4.1} \]

Assume that at time \( t = n \), there are \( n + 1 \) possible states of term structure vectors:

\[ v_{n,0}, v_{n,1}, \ldots, v_{n,n} \]

where

\[ v_{n,k} = (1, \beta(n, k, 1) \frac{P_{n+1}}{P_n}, \beta(n, k, 2) \frac{P_{n+2}}{P_n}, \ldots, \beta(n, k, m) \frac{P_{n+m}}{P_n}, \ldots) \tag{4.2} \]

At time \( t = n + 1 \), the time-state \((n, k)\) becomes time-states \((n + 1, k)\) and \((n + 1, k + 1)\), and the corresponding term structures are \( v_{n+1,k} \) and \( v_{n+1,k+1} \). With the above definitions, we have the following lemma:

**Lemma 14**

\[ v_{n,k}^* = \frac{\alpha(n + 1, k, 0) P_{n+1}}{\alpha(n, k, 0)} v_{n+1,k} + \frac{\alpha(n + 1, k + 1, 0) P_{n+1}}{\alpha(n, k, 0)} v_{n+1,k+1} \tag{4.3} \]

or equivalently

\[ \frac{v_{n,k}^*}{v_{n,k}(1)} = \lambda_{n+1} v_{n+1,k} + (1 - \lambda_{n+1}) v_{n+1,k+1} \tag{4.4} \]

**Proof:** The first component of the right hand side is

\[ \frac{\alpha(n + 1, k, 0) P_{n+1}}{\alpha(n, k, 0)} + \frac{\alpha(n + 1, k + 1, 0) P_{n+1}}{\alpha(n, k, 0)} \frac{P_{n+1}}{P_n} \]

\[ = \frac{\alpha(n+1,k,0) + \alpha(n+1,k+1,0)}{\alpha(n,k,0)} \frac{P_{n+1}}{P_n} \]

\[ = \frac{\alpha(n,k,1)}{\alpha(n,k,0)} P_{n+1} \frac{P_{n+1}}{P_n} \]

\[ = \beta(n, k, 1) \frac{P_{n+1}}{P_n} \]
which is exactly the first component of the right hand side of equation (4.3).

For $i > 0$, the $i + 1$-th component of the right hand side is

$$\frac{\alpha(n+1,k,0)}{\alpha(n,k,0)} \beta(n+1,k,i) \frac{P_{n+1,i}}{P_n} + \frac{\alpha(n+1,k+1,0)}{\alpha(n,k,0)} \beta(n+1,k+1,i) \frac{P_{n+1,i}}{P_n}$$

$$= \frac{\alpha(n+1,k,i)+\alpha(n+1,k+1,i)}{\alpha(n,k,0)} \frac{P_{n+1,i}}{P_n}$$

$$= \frac{\alpha(n,k,i+1)}{\alpha(n,k,0)} \frac{P_{n+1,i}}{P_n}$$

$$= \beta(n,k,i+1) \frac{P_{n+1,i}}{P_n}$$

which is exactly the $i + 1$-th component of the left hand side of the equation.

Since

$$v_{n,k}(1) = \frac{\alpha(n,k,1)}{\alpha(n,k,0)} \frac{P_{n+1}}{P_n}$$

(4.4) follows by dividing $v_{n,k}(1)$ on both side of (4.3) and using the following identities:

$$\frac{\alpha(n+1,k,0)}{\alpha(n,k,1)} = \lambda_{n+1}, \quad \frac{\alpha(n+1,k+1,0)}{\alpha(n,k,1)} = 1 - \lambda_{n+1}$$

Q.E.D.

The above lemma implies that the term structure tree constructed as above is arbitrage free (see [] or the original paper []). Also (4.4) implies that at time $t = n$, for all states and maturities, the risk neutral probabilities are the same, i.e., $\lambda_{n+1}$ and $1 - \lambda_{n+1}$. This can be explained as follows: if there is no uncertainty, at time $t = n$, then a zero coupon bond with maturity $i + 1$ period will have a determined value at time $t = n + 1$, which, in the absence of arbitrage, should be equal to value at
time $t = n$ compounded with the one period interest rate implied by $u_{n,k}(1)$, hence equal to, $v_{n+1,k}^1/v_{n,k}(1)$. Now with uncertainty, i.e., at time $t = n + 1$, the bond price could be either $v_{n+1,k}(i)$ or $v_{n+1,k+1}(i)$, (4.4) shows that the expected bond price with risk neutral probabilities $\lambda_{n+1}$ and $1 - \lambda_{n+1}$ is equal to the price when there is no uncertainty.

4.3 Constraints on the Parameters

Though the interest rate may not be an increasing function of its maturity (e.g., hump shaped yield curve), to avoid arbitrage, the price of a zero coupon bond should be a decreasing function of its maturity. For any time-state $(n, k)$, let

$$v_{n,k} = (v_{n,k}(0), v_{n,k}(1), v_{n,k}(2), v_{n,k}(3), v_{n,k}(4), \cdots)$$

where $v_{n,k}(0) = 1$, then we must have

$$1 = v_{n,k}(0) > v_{n,k}(1) > v_{n,k}(2) > v_{n,k}(3) > v_{n,k}(4) > \cdots \quad (4.5)$$

The other constraint is that for the two time-state $(n + 1, k)$ and $(n + 1, k + 1)$ which follow $(n, k)$, the corresponding term structure vector $v_{n+1,k}$ and $v_{n+1,k+1}$ should satisfy the following conditions to reflect the up and down movements:

$$v_{n+1,k}(i) < v_{n+1,k+1}(i) \quad \text{for all } n \geq 0, 0 \leq k \leq n \text{ and } i > 0. \quad (4.6)$$

So we must choose the parameter sequence $\{\lambda_i\}$ so that the inequalities (4.5) and (4.6) hold.
Since

\[ v_{n+1,k}(i) = \beta(n + 1, k, i) \frac{P_{n+i+1}}{P_n} \]

(4.6) is equivalent to

\[ \beta(n + 1, k, i) < \beta(n + 1, k + 1, i) \]

which is simplified to

\[ \frac{1 - \lambda_{n+1}}{\lambda_{n+1}} < \frac{1 - \lambda_{n+i+1}}{\lambda_{n+i+1}} \]

This inequality holds if and only if

\[ \lambda_{n+1} > \lambda_{n+i+1} \]

Since the above must hold for any \( n \geq 0 \) and \( i > 0 \), we conclude that (4.6) holds if and only if \( \{\lambda_i\} \) is a decreasing sequence, i.e.,

\[ 1 = \lambda_0 > \lambda_1 > \lambda_2 > \lambda_3 > \cdots \]  \hspace{1cm} (4.7)

Now we consider inequality (4.5). To include the case \( i = 0 \) in the following derivations, we note that \( \beta(n, k, 0) = 1 \) and

\[ v_{n,k}(0) = 1 = \beta(n, k, 0) \frac{P_n}{P_n} \]

By definition of the \( v_{n,k}(i) \), (4.5) is equivalent to

\[ \beta(n, k, i)P_{n+i} > \beta(n, k, i + 1)P_{n+i+1} \] for \( n \geq 0 \) and \( i \geq 0 \)

or

\[ \left( \frac{\lambda_{n+i}}{\lambda_{n+i+1}} \right)^{n-k} \left( \frac{1 - \lambda_{n+i}}{1 - \lambda_{n+i+1}} \right)^{k} > \frac{P_{n+i+1}}{P_{n+i}} \]  \hspace{1cm} (4.8)
We note that when \( n = 0 \), (4.8) holds automatically, so in what follows, we assume that \( n > 0 \).

Since \( \frac{\lambda_{n+i}}{\lambda_{n+i-1}} > 1 \) and \( \frac{1 - \lambda_{n+i}}{1 - \lambda_{n+i-1}} < 1 \), (4.8) holds for any \( k \) such that \( 0 \leq k \leq n \) if and only if it holds for \( k = n \), i.e.,

\[
\frac{1 - \lambda_{n+i}}{1 - \lambda_{n+i+1}} > \left( \frac{P_{n+i+1}}{P_{n+i}} \right)^{\frac{1}{n}}
\]

Denote

\[
t(n, i) = \left( \frac{P_{n+i+1}}{P_{n+i}} \right)^{\frac{1}{n}}
\]

and in particular, denote \( t_n = t(n, 0) \) for \( n > 0 \), then the above inequality is equivalent to the following:

\[
1 - (1 - \lambda_{n+i})t(n, i)^{-1} < \lambda_{n+i+1}
\]

for all \( n > 0 \) and \( i \geq 0 \).

From the nonarbitrage argument at the beginning of this section, we have \( \frac{P_{n+i+1}}{P_{n+i}} < 1 \) for all \( n \geq 0 \) and \( i \geq 0 \), hence \( t(n, i) < 1 \). Since

\[
t(n, i) = \left( \frac{P_{n+i+1}}{P_{n+i}} \right)^{\frac{1}{n}} \leq \left( \frac{P_{n+i+1}}{P_{n+i}} \right)^{\frac{1}{n+i}} = t(n + i, 0) = t_{n+i} < 1
\]

the first inequality of (4.9) holds for all \( n > 0 \) and \( i \geq 0 \) if and only if

\[
1 - (1 - \lambda_{n+i})t_{n+i}^{-1} < \lambda_{n+i+1}
\]

or equivalently

\[
1 - (1 - \lambda_i)t_i^{-1} < \lambda_{i+1} \quad \text{for all } i > 0.
\]
By combining conditions (4.7) and (4.10), we conclude that (4.5) and (4.6) hold simultaneously if and only if the following holds for all \( i > 0 \):

\[
\max\{0, 1 - (1 - \lambda_i) t_i^{-1}\} < \lambda_{i+1} < \lambda_i
\]  

(4.11)

Since \( t_i < 1 \) for \( i > 0 \), the condition (4.11) does not result in an empty set for \( \lambda_{i+1} \).

It is clear that we do have a free parameter \( \lambda_1 \) only subject to the condition \( 0 < \lambda_1 < 1 \).

Remark: Under condition (4.11), (4.5) holds, hence the prices of all zero coupon bonds at any time-state never reach or pass 1 before their maturities. This fact implies that the interest rates are positive for all time intervals under condition (4.11).

4.4 Implementation Procedure

Under the above model, the short rate of interest can be characterized easily. First we define the short rate to be the rate implied by the one-period zero coupon bond. In terms of previously defined notations, the price of one period zero coupon bond at time-state \( (n, k) \) is

\[
v_{n,k}(1) = \beta(n, k, 1) \frac{P_{n+1}}{P_n}
\]

\[
= (\frac{\lambda_{n+1}}{\lambda_n})^{n-k}(1 - \frac{\lambda_{n+1}}{1 - \lambda_n})^k \frac{P_{n+1}}{P_n}
\]

In general, the \( i \)-period interest rate is implied by the \( i \)-period zero coupon bond which has price \( v_{n,k} \) at time-state \( (n, k) \), and

\[
v_{n,k}(i) = (\frac{\lambda_{n+i}}{\lambda_n})^{n-k}(1 - \lambda_{n+i})^k \frac{P_{n+i}}{P_n}
\]
To convert zero coupon bond price to yield, let \( y_{n,k}(i) \) be the \( i \) period yield at time-state \((n,k)\), then

\[
v_{n,k}(i) = (1 + y_{n,k}(i))^{-i}
\]

Solve for \( y_{n,k}(i) \) to get

\[
y_{n,k}(i) = 1 - v_{n,k}(i)^{-1/i}
\]

or explicitly

\[
y_{n,k}(i) = 1 - (\beta(n, k, i) \frac{P_{n+i}}{P_n})^{-1}
\]

Denote \( y(j) = y_{0,0}(j) \), then

\[
P_j = (1 + y(j))^{-j}
\]

and

\[
y_{n,k}(i) = 1 - (\beta(n, k, i) \frac{(1 + y(n + i))^{-(n+i)}}{(1 + y(n))^{-n}})^{-1/i}
\]

The above equation shows that it is extremely easy to implement the model. Given the initial yield curve, we only need to choose a parameter \( \lambda_1 \) in the unit interval, then choose \( \lambda_{i+1} \) for \( i > 0 \) successively to satisfy the following conditions:

\[
\max\{0, 1 - (1 - \lambda_i) t_i^{-1}\} < \lambda_{i+1} < \lambda_i
\]

where

\[
t_i^{-1} = \left( \frac{P_{i+1}}{P_i} \right)^{1/i} = \frac{(1 + y(i + 1))^{1+1/i}}{1 + y(i)}
\]

### 4.5 Conclusion

Most term structure models including Ho-Lee model and Back-Derman-Toy model imply constant risk neutral probabilities. The implied risk neutral probabilities of the
model presented in this chapter are time dependent, but not state or maturity dependent. Numerical simulations show that this model generates rather small volatilities no matter how the parameters are chosen. But it does seem to behave well if volatilities are expected. So empirical testing of this model with small expected volatilities is a logical next step in studying this subject. This model, though it has its drawback as being only applicable to small volatility situations, may still provide us some insight on how to construct more realistic term structure models by introducing risk neutral probabilities which depends on time, states and maturities. Further research along this line seems to be quite worthwhile.
CHAPTER V

CONCLUSIONS

5.1 Summary

In this dissertation, we have considered some multinomial models of some financial instruments. The instruments we considered are options on stocks and term structure of interest rates. We build multinomial models of stocks by discretizing the corresponding continuous models. There are two problems involved in doing so, one is that we want such approximations to be optimal in some way. Second is that when time intervals are getting smaller and smaller, the resulting approximations should approach their continuous limits. Chapter 2 is devoted to these two problems. We first defined the optimality of approximations of discrete random variables (or random processes) by their moments, then showed that such optimal approximations always exist and can be explicitly calculated if we know the moments of the random variables to be approximated. We then apply this general result to some special cases such as normally distributed random variables (correspondingly, Brownian processes) and Poisson distributed random variables (correspondingly, Poisson processes). Since Black-Scholes option pricing model is based on continuous time Geometric Brownian
motion, Multinomial discretization of the normal random process yields multinomial approximations to the Black-Sholes formula. Similar, multinomial discretizations of both a normal random process and a Poisson random process yields multinomial approximations to the Merton's jump diffusion option pricing formula. The limits of such approximations to their continuous time formula are proved by using various central limit theorems.

Chapter 3 is devoted to generalize the famous Ho-Lee's binomial model of term structure of interest rates to multinomial models. The starting point of our generalization is that we want to keep all ingredients of Ho-Lee model, such as fitting initial term structure, explicit formulations of term structure movements, path-independent movements and the non-arbitrage nature of such movements, etc., yet maintaining the simplicity of the binomial model. We are quite successful in achieving our goal. The main difficulty in achieving our goal is to solve a system of $n^2$ difference equations. Fortunately, many of the difference equations are dependent. After clearing out all the redundant ones, we were able to solve the system explicitly in a nice formula which is truly a natural extension of the Ho-Lee's model. The main difference between the our multinomial model and the binomial Ho-Lee model is the number of parameters involved in the formulas. There are two parameters in Ho-Lee's models, but there are $n$ parameters in out $n$-nomial model. When $n = 2$, our model is exactly Ho-Lee's model.

Chapter 4 is devoted to develop a new non-arbitrage model of term structure of
interest rate. This model has the distinct feature that it has infinitely many parameters. It has all the ingredients of Ho-Lee's model and it also has a distinct property, i.e., it can generate interest rates which are positive for all time intervals. This property resembles that of the Black-Derman-Toy (BDT) model. But comparing with BDT model, this model has the advantage of being explicit in formulating the term structure movements. Those nice properties of this model are not without trade-off. The model only generates relatively small variances no matter how one choose the parameters (of course under the constraint that interest rates it generates are positive).

5.2 Future Research Topics

We hope that the result given in this thesis will spark some interest both theoretical researchers and practitioners in the industry. The theoretical part of chapter 2 is quite complete for the problems we set out to solve. As future research topics along this line, the following problems could be considered:

(i) Generalizations to models with time dependent variances.

(ii) Optimal multinomial approximations to multivariate models.

For Chapter 3, the model is as clean as Ho-Lee's model. Work remain to be done include:
(i) Empirical testing of the model.

(ii) Parameter choosing.

For Chapter 4, in addition to empirical testing and parameter choosing, we should look for similar models which have time-state-maturity dependent variances to overcome the limitation of the current model that variances it generates are too small.
BIBLIOGRAPHY


