Nonparametric predictive inference for option pricing based on the binomial tree model

HE, TING

Nonparametric predictive inference for option pricing based on the binomial tree model

Ting He

A Thesis presented for the degree of
Doctor of Philosophy

Durham University

Statistics and Probability
Department of Mathematical Sciences
Durham University
United Kingdom

March 2019

Nonparametric Predictive Inference for option pricing based on the binomial tree model

Ting He

Submitted for the degree of Doctor of Philosophy
2019

Abstract

Nonparametric Predictive Inference (NPI) is a frequentist statistical method based on only fewer assumptions, which has been developed for and applied to, several areas in statistics, reliability and finance. In this thesis, we introduce NPI for option pricing in discrete time models. NPI option pricing is applied to vanilla options and some types of exotic options.

We first set up the NPI method for the European option pricing based on the binomial tree model. Rather than using the risk-neutral probability, we apply NPI to get the imprecise probabilities of underlying asset price movements, reflecting more uncertainty than the classic models with the constant probability while learning from data. As we assign imprecise probabilities to the option pricing procedure, surely, we get an interval expected option price with the upper and lower expected option prices as the boundaries, and we named the boundaries the minimum selling price and the maximum buying price. The put-call parity property of the classic model is also proved to be followed by the NPI boundary option prices. To study its performance, we price the same European options utilizing both the NPI method and the Cox, Ross, and Rubinstein binomial tree model (CRR) and compare the results in two different scenarios, first where the CRR assumptions are right, and second where the CRR model assumptions deviate from the real market. It turns out that our NPI method, as expected, cannot perform better than the CRR in the first scenario with small size historical data, but as enlarging the history data size, the NPI method's performance gets better. For the second scenario, the NPI method performs better than the CRR model.

The American option pricing procedure is also presented from an imprecise statistical aspect. We propose a novel method based on the binomial tree. We prove through this method that it may be optimal for an American call option without dividends to be exercised early, and some influences of the stopping time toward option price prediction are investigated in some simulation examples. The conditions of the early exercise for both American call and put options are derived. The performance study of the NPI pricing method for American options is evaluated via simulation in the same two scenarios as the European options. Through the performance study, we conclude that the investor using the NPI method behaves more wisely in the second scenario than the investor using the CRR model, and faces to more profit and less loss than what it does in the first scenario.

The NPI method can be applied to exotic options if the option payoffs are a monotone function of the number of upward movements in the binomial tree, like the digital option and the barrier option discussed in this thesis. Otherwise, either we can manipulate the binomial tree in order to assign the upper and lower probabilities, for instance, the look-back option with the float strike price, or a new probability mass is needed to be assigned to the payoff binomial tree according to the option definition which is attractive and challenging for future study.

Declaration

The work in this thesis is based on research carried out in the Department of Mathematical Sciences at Durham University. No part of this thesis has been submitted elsewhere for any degree or qualification.

Acknowledgements

First of all, I would like to thank my supervisor Professor Frank Coolen, who offers his kind support, great advice and professional guidance. I would never live my Ph.D. life smoothly without your help. You not only taught me how to be a qualified and open-minded scholar but also to be a nice person. I am so lucky to be your supervisee. I would also thank my supervisor, Doctor Tahani Coolen-Maturi, who gives me so many good advice and comments about my research.

I am thankful for my country and Durham University that offer the scholarship to sponsor my Ph.D. study at Durham University.

I am so thankful to my parents who always take care of me and love me with all their hearts. To Yang, my boyfriend, thank you for being so supportive. Thanks to my friends, Lei and Nipada, and my colleagues, Nawapon, Themistoklis, James and Junbin for being so helpful and amazing.

Thank you to all these people for making my Ph.D. life in Durham so worthy and unforgettable.

Contents

Chapter 1

Introduction

The Binomial Tree Model (BTM) is a simple but efficient and easy to understand model which is widely implemented in option pricing [18]. In this thesis, we apply Nonparametric Predictive Inference (NPI) to the vanilla option pricing procedure based on BTM, named as the NPI method for European option pricing and the NPI method for the American option pricing. For the European option pricing method, we compare the expected payoffs and prices between the classic BTM model (CRR), presented by Cox, Ross and Rubinstein [30], and our method. Also, we calculate the expected profit and loss of the investor using our method when he trades with the only other investor in the market, predicting with the CRR model to assess our method's performance. The trade between two investors are settled in two extremes scenarios: in Scenario 1, the CRR model is making the prediction match the real market trend, while in Scenario 2, the CRR model predicts the market with the wrong assumptions. The American option is the option that can be exercised before its maturity. For the American option, because of its early exercise feature, the NPI method performance is evaluated by simulation. Same as in the NPI method for European option pricing, this performance study is also done in two extreme scenarios and displays the outcome with the profit and loss of the NPI person. Both European and American performance studies show that if there is a substantial wrong assumption in the CRR prediction, the NPI method performs better than the CRR model.

In addition to the vanilla option pricing application, the NPI method can also be applied to the exotic option pricing procedure. In this thesis, we also present the NPI pricing method for three types of exotic options, the digital option, the barrier option, and the look-back option, which are explained in Chapter 4. Among them, some can use the boundary probabilities to compute the maximum buying and the minimum selling price directly, whereas others either need to manipulate the binomial tree to apply the interval probabilities or change the imprecise probability mass assignment according to the option definition.

This chapter contains the following parts: In Section 1.1, we introduce the concepts of options; in Section 1.2, the binomial tree model is illustrated with some important definitions and properties; in Section 1.3, we discuss the estimation ways of discount rate; in Section 1.4, we introduce the fundamental concepts of the imprecise probability; in Section 1.5, the NPI method, especially NPI for Bernoulli random quantities is laid out; and in Section 1.6, we display the outline of this thesis.

1.1 Options

Options are financial products in the derivative market, which can be dated back to the 16th century in Amsterdam [36] and then became popular in London in 17th century [63]. The option was formally introduced with standard trading rules when the Chicago Board of Options Exchange (CBOE) was formed [34]. It gives the option holder the right but not the obligation to buy or sell the underlying asset at a predetermined price, the strike price K, and the option seller the obligation to sell or buy the underlying asset at K when this option is exercised. If the option buyer has the right to buy the underlying asset at the exercise time at a price of K, this option contract is called the call option. Or if the option buyer has the right to sell the underlying asset when he exercises it, this option contract is the put option. An option is described as "*in the money*" when its exercise gives the option holder profit, while an option is "*out of the money*" if its exercise produces a loss to the option holder. When the strike price is equal to the underlying asset price, the option is described as "*at the money*". The character described above, which is the option strike price and underlying asset price comparison is called "*the moneyness*" [13].

In the option category, there are two most straightforward types of options, the European option and the American option, which are named as vanilla options. The European option only allows the option holder to exercise the option at maturity. Then the expected option value at t ($t \leq T$) of the European option is equal to the discounted expectation of the payoff at maturity. $B(t,T)$ denotes the discount factor from t to maturity. The expected value of V of the European call option can be described below.

$$V(S,t) = E[B(t,T)(S_T - K_c)^+] \tag{1.1}$$

where S is the underlying asset price at time t, S_T is the underlying asset price at maturity, and K_c is the strike price of the European call option. Correspondingly, the expected value V of the European put option expected value at time t is,

$$V(S,t) = E[B(t,T)(K_p - S_T)^+] \tag{1.2}$$

where K_p is the strike price of the European put option. Therefore, the expected value of the put option also equals the discounted expectation of the payoff at maturity.

The American option allows its holder to exercise anytime during the option life period. The American option is generally more valuable than the European option because of its leeway [13]. Due to the early exercise feature of the American option, there is no closed form option pricing formula for American options. Then the value of the American option (without any dividend) $V(S,t)$ at time t, with stock price

$S_t = S$, is different from that of the European options. In terms of the American option, there is the stopping time τ, when the option is exercised, which for each possible path, the American option is exercised at the optimized stopping time τ giving us the optimization of this American option payoff. Here is the definition of the American option. Let $V(S,t)$ denotes the expected value of the American option, and it equals to the expected payoff discounted from the exercise time τ with the stock exercise price S_τ. For an American call option, the expected value is given by,

$$V(S,t) = \max_{\tau} E\left[B(t,\tau)(S_\tau - K_c)^+ \middle| S_t = S\right] \tag{1.3}$$

where $B(t,\tau)$ is the discount factor from t to τ. This formula defines the value of this call option at the time t, as being equal to the discounted instant payoff of this call option at the stopping time τ. For an American put option, the expected value is described as,

$$V(S,t) = \max_{\tau} E\left[B(t,\tau)(K_p - S_\tau)^+ \middle| S_t = S\right] \tag{1.4}$$

Therefore, the value of this put option at time t is equal to the discounted maximum payoff at τ. Other than vanilla options, there also exist exotic options with more complicated settings, for instance, the digital option, the barrier option, the look-back option and so on.

In the finance literature, the most prevalent two methods are the Binomial Trees Option Pricing Model (CRR), presented by Cox, Ross and Rubinstein [30] and the Black-Scholes Model [12]. Both methods assume that investors know the underlying asset from every perspective, for instance, for both the CRR model and the Black-Scholes model are set up in a risk-neutral world. Besides, for these models, the market is complete without any arbitrage opportunity. The Black-Scholes Model is not practical in a valuation of early exercised options, like American option, for it tends to exhibit systematic empirical biases related to the exercise price, the time to maturity and the variance when it is used in pricing the American option [38]. As a continuous time model, the Black-Scholes model is not considered in this thesis, because the option pricing method in this thesis is set up in the discrete time environment. All these assumptions are unlikely to be satisfied in the real world. Many papers are challenging those unrealistic assumptions and presenting new option pricing models. For example, Jackwerth and Rubinstein used a nonparametric method to deduce the option probabilities from option prices by using the quadratic minimization criterion [46]. GMPOP, short for generalized multi-period option pricing model, is a binomial tree model with subjective probability in the real world to price options, but this subjective probability is still constant [4]. While the NPI method provides an interval probability for each step, updating with the observed information. From this perspective, undoubtedly, implementing the NPI method in the option pricing procedure is reasonable, for in reality, the situations change all the time. In this thesis, we propose a novel approach to price the discrete time option using the NPI method. The NPI method is a frequentist statistical method inference based on historical data, which is under fewer assumptions of the market completeness and underlying asset and contains more uncertainty from the interval probability updating in every time step.

Some of the existing option pricing models are performed based on the Bayesian paradigm. Boyle and Ananthanarayanan [16] proposed an estimation method of the variance in the option pricing model with the Bayesian approach. Bauwens and Lubrano [9] conduct the Bayesian inference in the GARCH option pricing model to bring in the risk-neutral measurement. Jacquier and Jarrow [47] introduced the Bayesian inference to the Black-Scholes model to reduce the model error. Polson and Stroud [64] use the Bayesian simulation method to set up the stochastic volatility models. Martin et al. [55] defined an option price using the Bayesian approach that allows the time-varying volatility and non-normality in the conditional distribution. However, the Bayesian method has its drawbacks that it generates biased

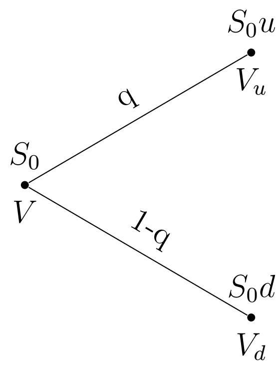


Figure 1.1: Stock and option prices in a general one-step tree

estimators of market option prices, while the NPI method gets all information from the historical data without bias.

1.2 Binomial Tree Model (BTM)

Since our option pricing method is based on the CRR binomial tree option pricing model together with the NPI method, in this section, the CRR binomial tree is introduced. The CRR mode is a discrete model which has been proved it converges to the Black-Scholes formula when time increments approach to zero [30]. Due to the flexibility and the ease of computation in the CRR model, it can be used to price the European option as well as the American option.

The binomial tree model was designed as a time-discrete pricing model dividing the life of option into a large number of small time intervals of length Δt [45]. It assumes there are only two possible prices for the underlying asset paying no dividend on the next time step. From the initial price S_0, the price either goes up to S_0u or goes down to S_0d, where u and d are the up and down movement factors respectively. This approach is depicted in Figure 1.1. Generally, $u > 1$ and $d < 1$, and the probability in the real world of the underlying asset up movement is denoted by p, while the real probability of the underlying asset down movement is indicated by $1 - p$. Then the expected stock price at the next time step can be calculated by $S_0(1 + r)^{-\Delta t} = pS_0u + (1 - p)S_0d$, where r is the expected return of the stock.

However, in the CRR binomial model, the valuation is under the assumption that the world is risk-neutral and complete, which means:

- Options are calculated based upon an absence of arbitrage profit.
- All investors in the market are risk-neutral, i.e., investors are regardless of risk preferences assigning the same value to the same options.
- Markets are frictionless, i.e., there is no transaction cost or other fees in the markets.

Therefore, the real probability of the underlying asset p barely plays a role in this model, whereas option prices are valued by the risk-neutral probability measure q for an upward movement, and $1-q$ for a downward movement. Moreover, the binomial tree model is irrelevant of the underlying asset's expected return, the reason is that options in a risk-neutral world have the same prices as in our real world, the risk-averse world, but with the risk-free rate as the expected return. It is reasonable that all risk-neutral measures change the real probability p to the risk-neutral probability q, for in the risk neutral world, the expected return of all financial product is risk-free rate, no matter it is a stock or a derivative upon an underlying asset, like the option. So it can effectively avoid difficulties from the expected return estimation of derivatives, as the expected return of a derivative is higher than the underlying asset and hard to be estimated [45]. Since the derivative is priced based on the value of the underlying asset, it has a higher leverage leading to a riskier position than its underlying assets. However, its risk premium is hard to be assessed because of the illiquidity of the derivatives market, causing a problem of discount rate estimation discussed in Section 1.3.

As the binomial tree is settled according to the risk-neutral valuation, option prices can be valued by the following procedure:

1. Compute the risk-neutral probability q, based upon the risk-free rate r_f which assumed to be constant during the option life. And the risk-neutral probability

calculation formula is gain by the following underlying expected equations:

$$S_0(1+r_f)^{-\Delta t} = qS_0u + (1-q)S_0d \tag{1.5}$$

or

$$(1+r_f)^{-\Delta t} = qu + (1-q)d \tag{1.6}$$

Then

$$q = \frac{(1+r_f)^{-\Delta t} - d}{u-d} \tag{1.7}$$

2. Value expected option payoffs and get option prices by discounting the payoffs at the risk-free rate. The detailed calculation procedure is displayed as below:

$$V^{CRR} = (1+r_f)^{-\Delta t}[qV_u^{CRR} + (1-q)V_d^{CRR}] \tag{1.8}$$

where V^{CRR} is the option price at time t, and V_u^{CRR} and V_d^{CRR} are option values of up movement and down movement at time $t+\Delta t$, respectively, regardless of option types, i.e. for call option, $V_u^{CRR} = \max\{S_0u - K, 0\}$, while for put option, $V_u^{CRR} = \max\{K - S_0u, 0\}$, where K is the strike price.

What we have discussed above is only a one-time step valuation procedure, and as we add more steps to the binomial tree, the risk-neutral valuation principle continues to be held. Since option prices are always identical to risk-neutral payoffs discounted at the risk-free rate, after all of the iterative processes through the whole option life, we can get the ultimate result which is the option price we wish to derive. There are closed formulae of European options, for a call option with maturity $T = m$, the strike price K_c based on an underlying asset with an initial price S_0 and movement factors u and d [30].

$$V_c^{CRR}[S_m - K_c]^+ = (1+r_f)^{-T} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m}{k} q^k (1-q)^{(m-k)} \tag{1.9}$$

where k_c^* is such that $u^{k_c^*}d^{m-k_c^*}S_0 - K_c = 0$, and $\lceil k_c^* \rceil$ denotes the smallest integer greater than or equal to k_c^*. T is the maturity time. For a put option, the expected price from the CRR model is given by [30],

$$V_p^{CRR}[K_p - S_m]^+ = (1+r_f)^{-T} \sum_{k=0}^{\lfloor k_p^* \rfloor} [K_p - u^k d^{m-k} S_0] \binom{m}{k} q^k (1-q)^{(m-k)} \quad (1.10)$$

k_p^* is such that $K_p - u^{k_p^*}d^{m-k_p^*}S_0 = 0$ and where $\lfloor k_p^* \rfloor$ denotes the largest integer less than or equal to k_p^*.

Compared to the celebrated Black-Scholes Model, the binomial tree is simpler and easier to use. It is also more versatile, for it can value varies types of options that can be exercised before maturity, like American option. In the limit, when the time intervals are squeezed and approach to zero, the binomial tree leads to the lognormal assumption for stock prices and underlies the Black-Scholes model [30]. Since the CRR model converges weakly to the geometric Brownian motion, Jarrow and Rudd [49] presented a binomial tree model matching the first two moments of the tree. Tian [71] introduced a new model matching the first three moments of the tree and the underlying geometric Brownian motion. Kim et al. [53] generalized all three model and set up a tree model fitting all moments to the approximated geometric Brownian motion. Tian [72] also developed a flexible binomial tree model with a "tilt" parameter to enhance the accuracy of the binomial tree prediction. Ji and Brorsen [50] developed a relaxed binomial tree model accounting for the skewness of the underlying distribution to relax the assumption of lognormality. The binomial tree model is widely used and developed as a result of its understandability and intuition. Jump diffusion has been added to the binomial tree that helps price the American option [65], the Asian option [52] and the look-back option[51]. Gerbessiotis [37] presented a latency- tolerant parallel algorithm to price the vanilla option for the multiplicative binomial tree model, which achieves theoretical speedup.

1.3 Discount rate

Originally, the discount rate is from the debt process. When a debtor is permitted to delay the payment to a creditor, there will be a charge for the debtor, which is the difference between the original amount of owed money and the amount of money needed to be paid in the future for the same unit, with the discount rate is the yield of it. For instance, a debtor A loan money l_t at time t, and admits to repay the money along with the extra compensation totally as $l_T = l_t + \varepsilon$ at time T. l_T is equal to the l_t times its yield rate function according to two interest calculation, single interest or compound interest. Then the yield of this transaction, as well as the discount rate, is $r = \left(\frac{l_T}{l_t}\right)^{\frac{1}{T-t}} - 1$ for the single interest situation or $r = \frac{\ln\left(\frac{l_T}{l_t}\right)}{(T-t)}$ for the compound interest situation. Concerning financial investment, the discount rate is the yield rate that an investor expects to gain during the holding period, namely that it is the expected return of the investment piece. Prices of assets are highly upon their risk levels. For example, there are two financial products, W and H, that they have the same value at maturity, $W_T = H_T$ but H is riskier than W. So an investor who is willing to buy these two products would ask a higher return of H because it is more precarious, $r_W < r_H$. The asset's value at the initial time is equal to the discounted expected value of the asset with the expected return as the discount rate, $W_0 = (1 + r_W)^{-T} W_T$ and $H_0 = (1 + r_H)^{-T} H_T$. After the discount, the price of H is lower than the price of W. This is what "high risk, high reward" means.

As acknowledged, in the CRR binomial tree this discount rate is the risk-free rate, because the CRR binomial tree is settled based on the risk-neutral valuation. The reason why the risk-neutral valuation is utilized is that in a risk-neutral world all products are riskless, and all individuals are indifferent to risk, where their expected returns for all products are the risk-free rate. Whereas, in a risky world, different investors have different risk levels they can tolerate, so every investment needs to be adjusted according to investors' risk aversion, which is time-consuming and difficult to estimate precisely. Thus, the risk-neutral valuation is commonly used because of its simplicity and efficiency.

Suppose an investor would like to invest an asset in the risk-neutral world because in this world there is no risk at all, then this investor would expect to gain a profit at the risk-free rate. If the same investor invests the same asset in a risk world, because of the uncertainty, this investor would like to ask more profit during this holding period, so the expected return would be higher than the risk-free rate. As mentioned earlier, the discount rate is the yield rate during the holding period, so in the risk-neutral world, the discount rate is risk-free rate while in the risk world it is equal to the asset return, which for a risk asset its discount rate is higher than the risk-free rate.

The example of two financial products, W and H are the underlying assets (not derivatives), when it comes to derivatives things become complex in the risk world. As in the risk world, derivatives involving options are much riskier than underlying assets, resulting in investors have a higher expectation of profit. In the risk-neutral world option prices are derived by Equation(1.8). But in the risk world, suppose the stock up movement probability is a constant value $p > q$, then the option prices are calculated by the equation $V_c[S_T - K_c]^+ = (1+r)^{-T} \sum_{k=\lceil k_c^* \rceil}^{T} [u^k d^{T-k} S_0 - K_c] \binom{T}{k} p^k (1-p)^{(T-k)}$, with $V_c[S_T - K_c]^+ = V_c^{CRR}[S_T - K_c]^+$, meaning $r > r_f$. Besides, the same option is going to have the same price using these two formulae [45], which means the discounted procedure eliminated the effects of risk on the same product. And the discount rate in this world is

$$r = r_f + r_{pr} \tag{1.11}$$

where r_{pr} is the yield of the risk premium referring to the financial product risk. When it relates to the underlying assets' discount rate, it equals the risk-free rate plus the risk premium of this asset in this market. However, it is hard to estimate the time discount rate for options, because the risk premium parameter of options is higher than that of stock, and it is not easy to judge the fair risk premium according to the information available from the market [45].

Although it is complicated to get the discount rate in the real world, there is no reason for us to overlook it or weaken it. As the discount rate is typically defined as 'the equilibrium expected rate of return on securities equivalent in risk to the project being value' [59], we could use the expected rate of return as the discount rate. According to option pricing analysis, there are two solutions to estimate the discount rate in the empirical market. The first one is using the expected return of the underlying asset as the discount rate. However, as its corresponding derivative has a higher risk, when there exist better solutions, this estimation method is less appropriate. Another way is that under the assumption of the completed market, seeking for a portfolio of securities that can perfectly replicate the payoff of the derivatives is always achievable, then the expected return of this portfolio is identical to the expected return of the derivative. Since the NPI method for option pricing is in the theoretical study, finding a replicated portfolio of securities is not feasible. Thus, we use the first strategy to estimate the discount rate for American options, and the discount rate is equal to the non-negative expected return of the underlying asset.

1.4 Imprecise probability

As the NPI method is a frequentist method based on imprecise probability theory, in this section, the idea and property of the imprecise probability theory are presented.

For a non-empty space $(\Omega, \mathcal{A})$, where Ω is a sample space and $\mathcal{A}$ is a set of events. For an event $A \in \mathcal{A}$, in precise probability theory, there is a specific probability, $p(A) \in [0, 1]$, with this probability p satisfying Kolmogorov's axioms [6]. Due to the lack of information gained from the real world for the event A, the precise probability misses estimating uncertainties in the real world so that it can not give a good result of the prediction of event A. Imprecise probability offers an alternative way to investigate event A, and the imprecise probability is an umbrella term comprised of all related quantitative uncertainty measurement providing multiple-valued probabilities, e.g., interval probabilities, as the outcomes [22]. Imprecise probabilities are becoming more and more popular nowadays, making it widely applied in different backgrounds, like artificial intelligence, engineering, chemistry, and biology [1].

The idea of using the imprecise probabilities is dated back at least to the middle of the nineteenth century [14]. The main idea of imprecise probabilities is that instead of using a certain probability to describe the uncertainty of the event A that is the path with the non-negative payoff in the binomial tree, we assign an interval probability, $[\underline{P}(A), \overline{P}(A)]$, to event A, where $0 \leq \underline{P}(A) \leq \overline{P}(A) \leq 1$. Imprecise probability extends the classic precise theory of probability, and here the classic probability becomes a special case, if $\underline{P}(A) = \overline{P}(A)$ is true, and it is identical to say that we have all information about event A in order to let us value the explicit probability of this event. Apart from this case, there exists another special case, $\overline{P}(A) - \underline{P}(A) = 1$, where $\overline{P}(A) = 1$ and $\underline{P}(A) = 0$, which means we have no information about event A. Weichselberger [74] defined a structure $\mathcal{M}$:

$$\mathcal{M} = \{p : \underline{P}(A) \leq p(A) \leq \overline{P}(A), \forall A \in \mathcal{A}\}$$

where p is a probability function on $\mathcal{A}$ in classical probability theory. According to the expression of imprecise probabilities stated by Augustin and Coolen [5], the imprecise probability can be expressed as optimal bounds for a set of probabilities of event $A \in \mathcal{A}$,

$$\underline{P}(A) = \inf_{p \in \mathcal{P}} p(A)$$

$$\overline{P}(A) = \sup_{p \in \mathcal{P}} p(A)$$

where $0 \leq \underline{P}(A) \leq \overline{P}(A) \leq 1$. Furthermore, the lower and upper probabilities hold a conjugacy property which links these two probabilities, $\overline{P}(A) = 1 - \underline{P}(A^c)$, where A^c is the complementing event of A. The lower and upper boundary probabilities have several interpretations, in general, $\underline{P}(A)$ reflects the information in support of A where $\overline{P}(A)$ reflects the information not supporting A^c [5].

Imprecise probability generalizes probability theory, for the circumstances that information is too limited to conclude a precise probability for an event of interest. So imprecise probability reflects more uncertainty about the event. Imprecise probabilities have been introduced to describe financial markets and to solve financial problems. For example, Berleant et al. [11] provide criteria and a measure for portfolio selection problems by utilizing the concept of imprecise probabilities. Imprecise probabilities also help with decision making in case of imprecise risk [48]. Muzzioli and Reynaerts [58] proposed a model to price American options with imprecise probabilities, where they introduced the fuzzy theory to the option pricing. The fussy set theory provides multiple outcomes of the underlying asset movements given an imprecise expected option value. Based on the successful application of imprecise probability theory, it is evident to believe that the implementation of the NPI method is also applicable.

1.5 Nonparametric Predictive Inference (NPI)

Coolen [21] introduced a statistical methodology called 'Nonparametric Predictive Inference' (NPI), to calculate the lower and upper probabilities for Bernoulli random quantities. It is on the basis of imprecise probability with frequentist statistical framework and strong consistency properties [5]. NPI is an inferential framework based on the assumption $A_{(n)}$ presented by Hill [41], which directly provides probabilities for events involving future observations by using few model assumptions and observed values of relevant random quantities without prior knowledge. Suppose that there exists a sequence of real-valued and exchangeable random quantities, $X_1, ..., X_n, X_{n+1}$. Assume that $X_1, ..., X_n$ be ordered and their realizations denoted as $x_{(1)} < ... < x_{(n)}$ and let $x_{(0)} = -\infty$ and $x_{(n+1)} = \infty$ for ease of notation. We assume there is no tie between any of them, if not, the results can be generalized to allow ties [56]. These ordered observed data partition the real line into $n+1$ open intervals $I_j = (x_{(j-1)}, x_{(j)})$, where $j = 1, 2, ..., n+1$. For the first predictive observation X_{n+1} on the basis of n observed values, $A_{(n)}$ [41] is

$$P(X_{n+1} \in (x_{(j-1)}, x_{(j)})) = \frac{1}{n+1} \text{ for } j = 1, 2, ..., n+1$$

So the probability for the event that the next observation falls in the interval $I_j = (x_{(j-1)}, x_{(j)})$ is $\frac{1}{n+1}$, for each interval I_j. As what Hill discussed [42], $A_{(n)}$ does not assume any knowledge of the distribution of random quantities of interest. By introducing imprecise probability theory, $A_{(n)}$ provides optimal bounds for the probability of any event of interest involving X_{n+1}, namely lower and upper probabilities in imprecise probability theory [73] and interval probability theory [75], following from De Finetti's fundamental theorem of probability [31].

The NPI method, a data-based imprecise probability method has been developed for a range of problems in operational research, including queueing [27], replacement problems [28], and many applications in reliability [24] and statistics [21, 25]. NPI has been applied to finance prediction, providing a relatively straightforward way to study future stock return when little further information is available providing an interval probability of the stock return greater than the target return and also a way of the pairwise comparison between stock returns [8]. The NPI method can also be implemented in credit rate for banking based on ROC analysis, which is under few assumptions and uses the imprecise probabilities to qualify the uncertainty [26]. Owing to the attractive properties of the NPI method [23], with fewer assumption in the method but embrace more uncertainty by using the imprecise probabilities, its implementation in option pricing is appealing. Unlike the CRR model, where the probability of stock movement is constant and precise, the probabilities from the NPI method are in the form of an interval with lower and upper bounds, gained through studying the observed data within a frequentist statistics framework, which makes it an appealing forecasting method [23]. Another exceptional property of NPI is that it keeps learning from data. When predicting non-independent multiple

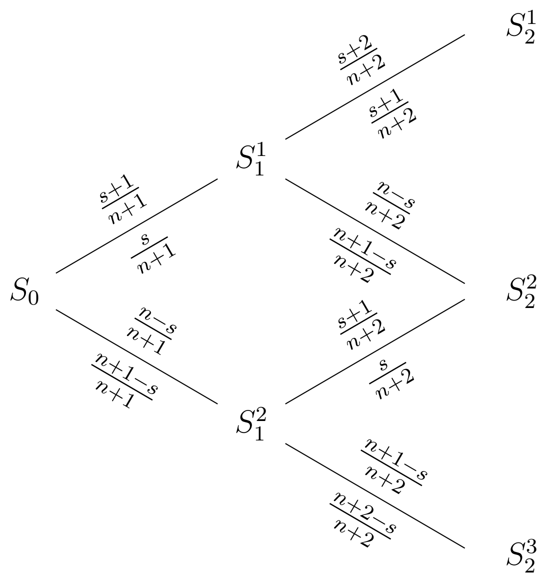


Figure 1.2: The binomial tree based on the NPI method

future observations, NPI considers all the predicted observations as observed data and uses the new imprecise probabilities determined from the predicted data and historical data to forecast the next future observation [21]. Thanks to utilizing imprecise probabilities from NPI method, outcomes of the predictions exhibit more uncertainty of the market than those of the CRR model.

As in the Binomial tree model, the underlying asset price is assumed to be a Bernoulli random quantities. The NPI method has been developed for Bernoulli data [21], which is used in this thesis for option pricing. Regarding the option pricing, there are two possible outcomes, the underlying asset price going up or down. Let (n, s) represent s increasing underlying asset prices in n historical underlying asset prices. Let $Y(m)$ represent the number of increasing underlying asset prices in m future underlying asset prices. Here n and m underlying asset prices are exchangeable. Denote $Z_t = \{z_1, .., z_t\}$, where $1 \leq t \leq m$ and $0 \leq z_1 < z_2 < ... < z_t \leq m$, and for the ease of notation, specify $\binom{s+z_0}{s} = 0$. Thus, the upper probability of the event $Y(m) \in Z_t$ based on the information (n, s), for $s \in \{0, ..., n\}$ [21], is

$$\overline{P}\left(Y(m) \in Z_t | (n,s)\right) = \binom{n+m}{n}^{-1} \sum_{j=1}^{t} \left[\binom{s+z_j}{s} - \binom{s+z_{j-1}}{s} \right] \binom{n-s+m-z_j}{n-s} \tag{1.12}$$

We can also deduce the comparative NPI lower probability by the conjugacy property $\overline{P}(A) = 1 - \underline{P}(A^c)$, where A^c is the complementary event to A,

$$\underline{P}(Y(m) \in Z_t | (n,s)) = 1 - \overline{P}(Y(m) \in Z_t^c | (n,s)) \tag{1.13}$$

where $Z_t^c = \{0, 1, ..., m\} \setminus Z_t$. On the basis of the NPI method, we can structure a binomial tree for a underlying asset with the price S_t^i, $t \in \{0, \ldots, m\}$ and $i \in \{1, \ldots, t+1\}$, as drawn in Figure 1.2. Based on Equation (1.12), for a one step binomial tree, the upper probability for the event $Y(m) = 1$, given data (n, s), for $s \in \{0, ..., n\}$, is [21]

$$\overline{P}\left(Y(1) = 1 | (n,s)\right) = \frac{s+1}{n+1} \tag{1.14}$$

The lower probability can be deduced by the conjugacy property,

$$\underline{P}\left(Y(1) = 1 | (n,s)\right) = \frac{s}{n+1} \tag{1.15}$$

The general formulae for any one time step in the binomial tree are also written below.

$$\underline{P}(Y(t+1) = 1 | (n+t, s+t-i+1)) = \frac{s+t-i+1}{n+t+1} \tag{1.16}$$

$$\overline{P}(Y(t+1) = 1 | (n+t, s+t-i+1)) = \frac{s+t-i+2}{n+t+1} \tag{1.17}$$

According to Coolen's [21], the derivation of NPI lower and upper probabilities for a m step option can be calculated by counting arguments directly. In this method, we predict m future observations given n observed values under the assumption of $A_{(n)}$, and a latent variable representation constituted of Bernoulli quantities is represented by all observations on the real line with a threshold between successes

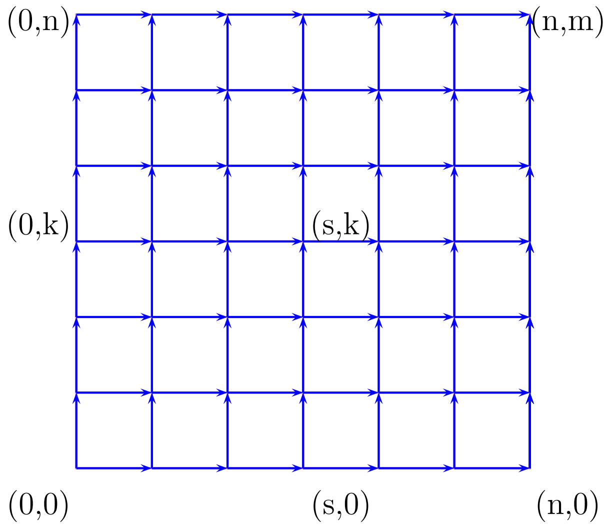


Figure 1.3: Paths counting of Bernoulli Quantities

and failures [1]. Among two types of observations, n observed data with s successes and m future data containing k successes, there exist $\binom{n+m}{m}$ different orderings of observations, which are equally likely. Alternatively, we can think in a way using the lattice from $(0,0)$ to (n,m), like what is shown in Figure 1.3, with movements either go right or upward, and the lower and upper probabilities are given by counting the paths. Specifically, the lower NPI probability $\underline{P}(Y(m) \in Z_t)$ is obtained by counting the number of all paths go only through the point (s,k) with $k \in Z_t$. And the upper NPI probability $\overline{P}(Y(m) \in Z_t)$ is given by counting the number of all paths go through the point (s,k) with $k \in Z_t$, including paths which might also go through the point (s,l) with $l \in Z_t^c$

The NPI lower and upper probabilities for events that are of interest in this thesis and the formulae are as follows [1]. The upper probability of the event $\{Y(m) = k|(n,s)\}$, given when all paths go through the point (s,k) are counted. There are $\binom{s+k}{k}\binom{n-s+m-k}{m-k}$ paths go through the point (s,k), so the formula is described as

$$\overline{P}(Y(m) = k|(n,s)) = \binom{n+m}{m}^{-1}\binom{s+k}{k}\binom{n-s+m-k}{m-k} \tag{1.18}$$

While the lower probability of the event $\{Y(m) = k|(n, s)\}$ is

$$\underline{P}(Y(m) = k|(n, s)) = \binom{n+m}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k-1}{m-k} \tag{1.19}$$

which is based on the number of all the paths go through both the point $(s-1, k)$ and the point $(s+1, k)$ that is $\binom{s+k-1}{k}\binom{n-s+m-k-1}{m-k}$. Formulas for special cases, $k = 0$ and $k = m$ are discussed below:

$$\underline{P}(Y(m) = m|(n, s)) = \binom{s+m-1}{m} \binom{n+m}{m}^{-1} \tag{1.20}$$

$$\overline{P}(Y(m) = m|(n, s)) = \binom{s+m}{m} \binom{n+m}{m}^{-1} \tag{1.21}$$

$$\underline{P}(Y(m) = 0|(n, s)) = \binom{n-s+m-1}{m} \binom{n+m}{m}^{-1} \tag{1.22}$$

$$\overline{P}(Y(m) = 0|(n, s)) = \binom{n-s+m}{m} \binom{n+m}{m}^{-1} \tag{1.23}$$

By using path counting method, we discover the probability relationships between the events $\{Y(m) \geq k^*|(n, s)\}$ and $\{Y(m) \geq k^*+1|(n, s)\}$ for $k^* \in \{0, 1, ..., m-1\}$, as well as the relationship between the event $\{Y(m) \leq k^{**}|(n, s)\}$ and $\{Y(m) \leq k^{**}+1|(n, s)\}$ where $k^{**} \in \{1, ..., m\}$, according to their lower and upper probabilities, respectively [21].

$$\begin{aligned} &\underline{P}(Y(m) \geq k^*|(n, s)) - \underline{P}(Y(m) \geq k^*+1|(n, s)) \\ &\quad = \binom{n+m}{m}^{-1} \binom{s+k^*-1}{k^*} \binom{n-s+m-k^*}{m-k^*} \end{aligned} \tag{1.24}$$

$$\begin{aligned} &\overline{P}(Y(m) \geq k^*|(n, s)) - \overline{P}(Y(m) \geq k^*+1|(n, s)) \\ &\quad = \binom{n+m}{m}^{-1} \binom{s+k^*}{k^*} \binom{n-s+m-k^*-1}{m-k^*} \end{aligned} \tag{1.25}$$

$$\begin{aligned} &\underline{P}(Y(m) \leq k^{**}|(n, s)) - \underline{P}(Y(m) \leq k^{**}+1|(n, s)) \\ &\quad = \binom{n+m}{m}^{-1} \binom{s+k^{**}}{k^{**}} \binom{n-s+m-k^{**}-1}{m-k^{**}} \end{aligned} \tag{1.26}$$

$$\overline{P}(Y(m) \leq k^{**}|(n,s)) - \overline{P}(Y(m) \leq k^{**}+1|(n,s))$$
$$= \binom{n+m}{m}^{-1} \binom{s+k^{**}-1}{k^{**}} \binom{n-s+m-k^{**}}{m-k^{**}} \quad (1.27)$$

These equations are derived by counting the movements for the right-upward paths from $(0,0)$ to (n,m). Equation (1.24), for instance, can be attained by counting all the paths go through the point (s,k^*) and points above other than (s,y) for $y < k^*$, identical to all the paths via both two specific points, $(s-1,k^*)$ and (s,k^*). We can get the lower and upper probabilities of events $\{Y(m) \leq k^*|(n,s)\}$ and $\{Y(m) \geq k^{**}|(n,s)\}$ for $k^* \in \{0,1,...,m-1\}$ and $k^{**} \in \{1,...,m\}$. Detailed derivation is shown as following:

$$\underline{P}(Y(m) \geq k^*|(n,s)) = \binom{n+m}{m}^{-1} \sum_{k=k^*}^{m} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \quad (1.28)$$

$$\overline{P}(Y(m) \geq k^*|(n,s)) = \binom{n+m}{m}^{-1} \sum_{k=k^*}^{m} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \quad (1.29)$$

$$\underline{P}(Y(m) \leq k^{**}|(n,s)) = \binom{n+m}{m}^{-1} \sum_{k=0}^{k^{**}} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \quad (1.30)$$

$$\overline{P}(Y(m) \leq k^{**}|(n,s)) = \binom{n+m}{m}^{-1} \sum_{k=0}^{k^{**}} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \quad (1.31)$$

As we discussed, the NPI method for Bernoulli data [21] provides a set $\mathcal{P}$ of classical, precise, probability distributions for which the presented lower and upper probabilities are optimal bounds. In imprecise probability theory, this set $\mathcal{P}$ is called a structure. So for any event $Y(m) \in Z_t$

$$\underline{P}(Y(m) \in Z_t) = \inf_{p \in \mathcal{P}} p(Y(m) \in Z_t) \quad (1.32)$$

$$\overline{P}(Y(m) \in Z_t) = \sup_{p \in \mathcal{P}} p(Y(m) \in Z_t) \quad (1.33)$$

Similarly, lower and upper expected values for a real-valued function g of $Y(m)$ can be derived. In the option pricing method, this real-valued function equation is equal to the positive payoffs according to the options' definitions.

$$\underline{E}(g(Y(m))) = \inf_{p \in \mathcal{P}} E^p(g(Y(m))) \tag{1.34}$$

$$\overline{E}(g(Y(m))) = \sup_{p \in \mathcal{P}} E^p(g(Y(m))) \tag{1.35}$$

where $E^p(g(Y(m)))$ is the expected value or payoff based on the probability function p. On behalf of option trading positions, $\underline{E}(g(Y(m)))$ denotes the maximum payoffs an investor would be willing to pay and $\overline{E}(g(Y(m)))$ denotes the minimum payoffs an investor would be willing to sell.

We combine the NPI method with the binomial tree option pricing model, getting a novel option pricing method learning from the historical data and concerning more uncertainty by introducing an interval probability instead of a constant probability to the binomial tree.

1.6 Outline

In this thesis, we present the NPI method for various types of options, both vanilla and exotic options based on the BTM. The thesis is organized as follows. In Chapter 2, Nonparametric Predictive Inference for European option pricing based on the Binomial Tree Model is proposed. The whole method is set up and studied examples comparing with the CRR model in order to assess its performance. In this chapter, we also discuss the put-call parity from the NPI method perspective. In Chapter 3, we introduce the NPI option pricing method based on the binomial tree model for American options. Some examples of early exercise call options based on the NPI method are investigated to manifest that the rational trading theory: 'Never early exercise an American call option without dividends' is not valid in our method. We also do the simulation of the trade between two investors, one using the NPI method and the other using the CRR model, to study the performance of our method. In Chapter 4, the NPI pricing methods for exotic options are introduced.

The NPI option pricing methods for the digital option and the barrier option are set up based on the underlying asset price binomial tree according to their option definition. The NPI option pricing method for the look-back option with a float strike price is based on a new binomial tree manipulated to be monotonic according to its definition. The NPI option pricing method raises interesting questions for future research, some brief comments and general conclusions are included in Chapter 5.

A paper based on the content in Chapter 2 has been published by the Journal of the Operational Research Society [40]. The content of this chapter also has been presented at several conferences, Research Students' Conference in Probability and Statistics (Durham, UK, April 2017) and the 7th International Conference of the Financial Engineering and Banking Society (Glasgow, UK, June 2017). A paper related to Chapter 3 is ready to submit. The results of Chapter 3 has been presented at some conferences and the seminars, Stats4grads weekly seminar in statistics (Durham, UK, November 2017), the 17th Winter school on Mathematical Finance (Lunteren, Netherlands, January 2018) and at the training school of Uncertainty Treatment and Optimization (Durham, UK, July 2018). A paper in the light of Chapter 4 is in preparation to be submitted to an academic journal.

Chapter 2

NPI for European Option Pricing

In this chapter, we present the NPI method for European option pricing based on the binomial tree model. The European option is one of the vanilla options, which is the most basic but popular option type in the market [45]. To start the investigation of our novel option pricing method in Section 2.1, the NPI method is applied to the European option pricing both without and with the discount factor. We also discuss the put-call parity for the NPI method in Section 2.2. In Section 2.3, we study its performance compared to the CRR model in two extreme scenarios and also inquire into the performance of the NPI method with the discount procedure. We conclude the content of the whole chapter and discuss further interesting topics in Section 2.4.

2.1 NPI option pricing method

In this section, we use the NPI method for Bernoulli random quantities [21] discussed in Section 1.5 to evaluate European option payoffs. The assumptions of our NPI European option pricing method are as follows. The initial underlying asset S_0 has two possible outcomes at the next time step, either going up to uS_0 or going down to dS_0, with $u > 1$, $d < 1$ and S_0 is the initial stock price without paying any dividends during the period considered. So far, what described above is the same

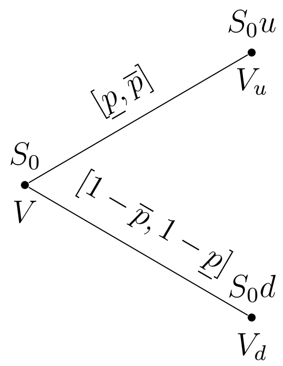


Figure 2.1: Stock and option prices in one-step tree

as those in the CRR binomial trees in Section 1.2.

The NPI method is applied to the binomial trees. As reviewed in Section 1.5, these m future stock prices are seen as Bernoulli random quantities. Unlike the CRR binomial trees which use an explicit probability from risk-neutral valuation, we assume that the n historical data are sufficient to analyze option prices, and among n observed data the stock price went up s times and down $n-s$ times. To simplify our model, there is no effect of the discount factor at first by assuming the time of trading is close to maturity, so the influence of any discount factor is neglectable. Excluding the discount factor eliminates the error caused by the inappropriate estimation of discount rate to expose the real results of the NPI method in the later performance study. Same settings as in Section 1.5, suppose the random number of up movements during future m time steps is $Y(m)$. Then the stock price at time m is expressed as:

$$S_m = u^{Y(m)} d^{m-Y(m)} S_0 \tag{2.1}$$

Based on Equations (1.34) and (1.35), we can compute the upper and lower expectations of the European option payoff, where the real-value function g is equal to $[S_m - K]^+$ with the strike price K for the European call option and $[K - S_m]^+$ for the European put option referring to the European option defined in Section 1.1. The binomial tree for the European option can be structured based on the NPI method in Figure 2.1.

Referring to Figure 2.1, V represents the option value at each node, then V_d and V_u represent the option value for nodes with stock prices $S_0 d$ and $S_0 u$, respectively. The interval probability from the NPI method is written as $[\underline{p}, \overline{p}]$ for the upward movement and $[1-\overline{p}, 1-\underline{p}]$ for the downward movement. For each type of option, only paths with positive payoffs are taken into account because an option is a right for the buyer, and the buyer would like to exercise the option if the payoff is positive. For call options, only paths which have payoff $S_m - K_c$ greater than zero, are taken into account, where S_m is the stock price at maturity and K_c is the strike price, then,

$$S_m - K_c = u^{Y(m)} d^{m-Y(m)} S_0 - K_c > 0 \tag{2.2}$$

$$Y(m) > \frac{\ln K_c - \ln S_0 - m \ln d}{\ln u - \ln d} =: k_c^* \tag{2.3}$$

The NPI lower and upper probabilities for the paths having positive payoffs in the European call option binomial tree, all stock prices at the m step higher than the strike price of this call option, are calculated according to the NPI method for Bernoulli data.

$$\underline{P}(Y(m) \geq \lceil k_c^* \rceil) = \binom{n+m}{m}^{-1} \sum_{k=\lceil k_c^* \rceil}^{m} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \tag{2.4}$$

$$\overline{P}(Y(m) \geq \lceil k_c^* \rceil) = \binom{n+m}{m}^{-1} \sum_{k=\lceil k_c^* \rceil}^{m} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \tag{2.5}$$

For the put option, the paths with payoffs $K_p - S_m > 0$ are considered, where K_p is the strike price. By this definition, the payoff of a put option is $[K_p - S_m]^+$, then,

$$K_p - S_m = K_p - u^{Y(m)} d^{m-Y(m)} S_0 > 0 \tag{2.6}$$

$$Y(m) < \frac{\ln K_p - \ln S_0 - m \ln d}{\ln u - \ln d} =: k_p^* \tag{2.7}$$

Following the same steps, as we did for call options, we find paths valued for the put option, and the interested event is $Y(m) \leq \lfloor k_p^* \rfloor$.

$$\underline{P}(Y(m) \leq \lfloor k_p^* \rfloor) = \binom{n+m}{m}^{-1} \sum_{k=0}^{\lfloor k_p^* \rfloor} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \tag{2.8}$$

$$\overline{P}(Y(m) \leq \lfloor k_p^* \rfloor) = \binom{n+m}{m}^{-1} \sum_{k=0}^{\lfloor k_p^* \rfloor} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \tag{2.9}$$

In addition to the lower and upper probabilities above we are also interested in the lower and upper expected values, given by Equations (1.34) and (1.35), where the real-valued function $g(Y(m))$ is equal to $[S_m - K_c]^+$ for a call option and $[K_p - S_m]^+$ for a put option, because S_m is a random variable depending on m. According to the trading actions, expected boundary payoffs are renamed, $\underline{E}$ denotes the maximum payoff an investor would buy at, and $\overline{E}$ denotes the minimum payoff an investor would sell for. As we have already computed the lower and upper probabilities for the call options (Equations (2.4) and (2.5)), as well as for the put options (Equations (2.8) and (2.9)), then formulae for European option expected payoffs can be generated [40] as follows:

The minimum selling payoff of the call option

$$\begin{aligned}
&\overline{E_c}[S_m - K_c]^+ \\
&= \binom{n+m}{m}^{-1} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c](\overline{P}(Y(m) \geq k) - \overline{P}(Y(m) \geq k+1)) \\
&= \binom{n+m}{m}^{-1} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{s+k}{k} \binom{n-s+m-k-1}{m-k}
\end{aligned} \tag{2.10}$$

Here for each term with k from $\lceil k_c^* \rceil$ to m, we assign a probability $\overline{P}(Y(m) \geq k) - \overline{P}(Y(m) \geq k+1)$ to ensure that we give the maximum possible probability to the largest possible value for k, then the maximum possible remaining probability to the second largest value for k, and so on. The corresponding minimum selling payoff of the put option can be written as follows.

The minimum selling payoff of the put option

$$\overline{E_p}[K_p - S_m]^+ = \binom{n+m}{m}^{-1} \sum_{k=0}^{\lfloor k_p^* \rfloor} [K_p - u^k d^{m-k} S_0](\overline{P}(Y(m) \leq k) - \overline{P}(Y(m) \leq k-1))$$
$$= \binom{n+m}{m}^{-1} \sum_{k=0}^{\lfloor k_p^* \rfloor} [K_p - u^k d^{m-k} S_0] \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \tag{2.11}$$

For downward paths with k from 0 to $\lfloor k_p^* \rfloor$, each path is assigned a probability $\overline{P}(Y(m) \leq k) - \overline{P}(Y(m) \leq k-1)$, which ensures that we give the maximum possible probability to the lowest possible value for k, then the maximum possible remaining probability to the second lowest value for k, and so on.

Using similar derivations, we can formulate the lower expected payoff for a call option and a put option as follows.

The maximum buying payoff of the call option

$$\underline{E_c}[S_m - K_c]^+ = \binom{n+m}{m}^{-1} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \tag{2.12}$$

The maximum buying payoff of the put option

$$\underline{E_p}[K_p - S_m]^+ = \binom{n+m}{m}^{-1} \sum_{k=0}^{\lfloor k_p^* \rfloor} [K_p - u^k d^{m-k} S_0] \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \tag{2.13}$$

Therefore, for each type of option, there is an interval of expected payoffs with bounds as the maximum buying payoff and the minimum selling payoff. As we calculated, for the call and the put options we get an interval of the expected values, which means with limited information any value in this interval is reasonable to the NPI investor, and any value outside this interval is appealing to the NPI investor.

When the NPI investor is offered a payoff higher than the minimum selling payoff, it is overvalued according to NPI outcomes. Similarly, the NPI investor would see any value less than the maximum buying payoff as undervalued, while the value in between expected value bounds does not trigger any trading action.

In the thesis, the formulae with the discount factor are also provided to complete the NPI method from the time value perspective. As discussed in Section 1.3, after introducing the discount rate to the evaluation formula, we rewrite the European option prices with maturity m. For the call option, the maximum buying price and the minimum selling price are listed below.

The maximum buying price of the call option

$$\underline{V_c}[S_m{-}K_c]^+ = (1{+}r)^{-m}\binom{n+m}{m}^{-1}\sum_{k=\lceil k_c^*\rceil}^{m}[u^k d^{m-k}S_0{-}K_c]\binom{s+k-1}{k}\binom{n-s+m-k}{m-k} \tag{2.14}$$

The minimum selling price of the call option

$$\overline{V_c}[S_m{-}K_c]^+ = (1{+}r)^{-m}\binom{n+m}{m}^{-1}\sum_{k=\lceil k_c^*\rceil}^{m}[u^k d^{m-k}S_0{-}K_c]\binom{s+k}{k}\binom{n-s+m-k-1}{m-k} \tag{2.15}$$

These formulae are deduced by adding the discount factor $(1+r)^{-m}$, where r is the discount rate equal to the non-negative expected return of the underlying asset in this thesis. The function of the discount factor is to discount the time value of the payoff. After the discount procedure, we get the expected call option price at the initial time. By adding the same discount factor in the put option formulae, we get the formulae of the expected put option price at time 0.

The maximum buying price of the put option

$$\underline{V_p}[K_p - S_m]^+ = (1+r)^{-m}\binom{n+m}{m}^{-1}\sum_{k=0}^{\lfloor k_p^* \rfloor}[K_p - u^k d^{m-k} S_0]\binom{s+k}{k}\binom{n-s+m-k-1}{m-k} \tag{2.16}$$

The minimum selling price of the put option

$$\overline{V_p}[K_p - S_m]^+ = (1+r)^{-m}\binom{n+m}{m}^{-1}\sum_{k=0}^{\lfloor k_p^* \rfloor}[K_p - u^k d^{m-k} S_0]\binom{s+k-1}{k}\binom{n-s+m-k}{m-k} \tag{2.17}$$

By far, we finish setting up the NPI method for the European option pricing and present the formulae of the boundary results, the maximum buying price and the minimum selling price, either with or without the discount procedure.

2.2 The put-call parity

After constructing the NPI option pricing method, a popular property in the classic option pricing model is of interest. In financial mathematics, there is a relationship, called the put-call parity, between the price of a European call option and a European put option, with the same exercise price and the exercise date, namely that a portfolio containing a long call option and a short put option has the same value as a single forward contract at the identical strike price and maturity, and this future contract can also be constituted by an underlying asset and an opposite position of the riskless asset. The put-call parity traces back to the 17th century. Both de la Vega[32] and de Pinto[33] wrote statements indicating that the put-call parity was described and understood based on its application in the late 17th century and 18th century Amsterdam option markets. The put-call parity is valid because if the underlying asset has a higher price than the strike price at maturity, the call option is exercised, while if the price is below the strike price, the put option is exercised. Thus, in either case, one unit of the asset is purchased at the strike price, exactly as in a forward contract. Besides, no matter which option is exercised, the payoff of this portfolio is $S_T - K$ at time T.

The assumptions of the put-call parity need to be clarified before we study this property in the setting of the NPI method. First of all, neither a put nor a call option can be exercised before maturity, meaning that the type of options here is "European". And stock dividends are protected, which the value of dividends will be subtracted from the stock exercise price at maturity, or the stock does not pay any dividend during the option holding period. Also, no transaction costs or other fees exist in the market. Last but not least, the market is risk-neutral that investors can borrow or lend money at a risk-free rate.

The put-call parity can be expressed as

$$c_e(t) - p_e(t) = S_t - K \cdot B(t,T), \qquad \text{with } B(t,T) = (1+r_f)^{-(T-t)} \tag{2.18}$$

where $c_e(t)$ and $p_e(t)$ are fair prices of a European call option and a European put option, respectively, having the same strike price K and the underlying asset at price S_t, and $B(t,T)$ is the time discount factor consisting of riskless rate r_f. It denotes that the cost of a European call option can be inferred from the value of a European put option with the same strike price and exercise date and vice versa [45]. As it is mentioned earlier, a put option, a call option, and the underlying security constitute an interrelated securities complex, leading to a profit or loss can be yielded when the put or call price in the deviates market substantially from the parity price [54].

To make this relationship between put and call option clearer, we illustrate this in Table 2.1. There are four trading strategies in Table 2.1, Strategy A, B, C and D. Strategy A tells us to buy a call option for $-c_e$ at time 0, then at maturity we either get the payoff if $S_T \geq K$ or nothing if $S_T \leq K$. Strategy B is a portfolio containing a buying position of the stock at a price S_0, a buying position of a put option with the price $-p_e$ and a loan. So at the initial time the total cash flow of Strategy B is $-p_e - S_0 + K \cdot B(t,T)$, and the maturity cash flow is $S_T - K$ when $S_T \geq K$ or 0 when $S_T \leq K$. Strategy C is to buy a put option at a price $-p_e$, and

Strategy	Cashflow at Time 0	$S_T \leq K$ (maturity)	$S_T \geq K$ (maturity)
A Buy call	$-c_e$	0	$S_T - K$
B Buy stock	$-S_0$	S_T	S_T
Buy put	$-p_e$	$K - S_T$	0
Borrow	$K \cdot B(t,T)$	$-K$	$-K$
Total	$-p_e - S_0 + K \cdot B(t,T)$	0	$S_T - K$
C Buy put	$-p_e$	$K - S_T$	0
D Sell stock	S_0	$-S_T$	$-S_T$
Buy call	$-c_e$	0	$S_T - K$
Lend	$-K \cdot B(t,T)$	K	K
Total	$S_0 - c_e - K \cdot B(t,T)$	$K - S_T$	0

Table 2.1: The relationship of European options and stock without dividend

get payoff $K - S_T$ if $S_T \leq K$ or nothing if $S_T \geq K$. Strategy D is also a portfolio with selling a stock, buying a call option and a deposit. The initial cash flow is $S_0 - c_e - K \cdot B(t,T)$, and the maturity cash flow is $K - S_T$ if $S_T \leq K$ or zero if $S_T \geq K$. From the table, it is evident that strategies A and B yield same cash flow at maturity, so as strategies C and D. Therefore, it is sure that these two combined strategies should have the same cash flow at the initial time because the market is under the risk-neutral assumption.

If call prices are overvalued, arbitrage occurs by selling a call option and longing a call position, taking an action of strategy C and strategy D, for a call option can be converted into complex securities within a put option. The sure profit of M from this action is :

$$c_e - p_e - S_0 + K \cdot B(t,T) = M \tag{2.19}$$

The procedure is the same when it comes to an overpriced put option. By taking a trade strategy, selling a call option and longing a put position, identical to a combination of strategy A and strategy B, a profit N can be achieved, which can be represented as:

$$p_e - c_e + S_0 - K \cdot B(t,T) = N \tag{2.20}$$

And when these two situations appear in the market, the put-call parity is invalid, and there exist arbitrage opportunities in the market. According to our method, this phenomenon will frequently happen when option prices fall outside the interval prices, for investors would not like to take any action when prices are in the interval, and this problem is detailedly discussed below.

In terms of the assumptions, there is one main difference between our method and the classic one. On the contrary to the original environment in which the classic put-call parity was proved, the market in which NPI option pricing model is settled is incomplete, which means that there exist arbitrage opportunities, and the bid-ask spread is non-zero. To avoid the influence caused by the discount procedure on the NPI method, we do not consider the discounted factor in our payoff valuation in NPI pricing model, so we only demonstrate the put-call parity in non-discounted version:

$$E_c - E_p = S_T - K \tag{2.21}$$

As it is described in the put-call parity both call and put options of the portfolio structured in the put-call parity have the same exercise value $K_c = K_p = K$, then for ease of notation Equation (2.13) is rewritten to

$$\underline{E_p}[K - S_m]^+ = \binom{n+m}{m}^{-1} \sum_{k=0}^{\lceil k^* \rceil - 1} [K - u^k d^{m-k} S_0] \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \tag{2.22}$$

where $\lceil k^* \rceil$ is the same value as it is in the call option. Now we intend to prove that the boundary prices from NPI follow the put-call parity by replacing E_c and E_p with the minimum selling payoff of a call option and the maximum buying payoff of a put option, respectively, in order to see if it can get the same result as the original one does. The proof is given as follows.

$$
\begin{aligned}
&\overline{E_c}[S_m - K]^+ - \underline{E_p}[K - S_m]^+ \\
&\quad = \binom{n+m}{m}^{-1} \sum_{k=\lceil k^* \rceil}^{m} [u^k d^{m-k} S_0 - K] \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \\
&\quad - \binom{n+m}{m}^{-1} \sum_{k=0}^{\lceil k^* \rceil - 1} [K - u^k d^{m-k} S_0] \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \\
&\quad = \binom{n+m}{m}^{-1} \sum_{k=0}^{m} [S_k - K] \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \\
&\quad = \overline{S_m} - K
\end{aligned} \tag{2.23}
$$

where $\overline{S_m}$ is the upper stock price at maturity:

$$
\overline{S_m} = \binom{n+m}{m}^{-1} \sum_{k=0}^{m} S_k \binom{s+k}{k} \binom{n-s+m-k-1}{m-k}
$$

$\overline{S_m}$ is the minimum stock price a person is willing to sell for. Equation (2.23) is identical to the non-discounted version of put-call parity, proving boundary option prices based on NPI follows the put-call parity.

For completeness, it is necessary to investigate into the other case, which is holding a portfolio encompassing two actions of selling a put option and buying a call option. Then the put-call parity can be described as the version down below,

$$
E_p - E_c = K - S_T \tag{2.24}
$$

Doing the same steps as in the first case, E_c is taken place by the maximum buying payoff of a call option Equation(2.12). And E_p is replaced by the formula deduced based on Equation(2.11) valuing the minimum selling payoff of the put option as following:

$$
\overline{E_p}[K - S_m]^+ = \binom{n+m}{m}^{-1} \sum_{k=0}^{\lceil k^* \rceil - 1} [K - u^k d^{m-k} S_0] \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \tag{2.25}
$$

Then, the alternative put-call parity can be investigated:

$$
\begin{aligned}
&\overline{E_p}[K-S_m]^+ - \underline{E_c}[S_m-K]^+ \\
&\qquad = \binom{n+m}{m}^{-1} \sum_{k=0}^{\lceil k^*\rceil -1} [K-u^k d^{m-k} S_0] \binom{s+k-1}{k}\binom{n-s+m-k}{m-k} \\
&\qquad - \binom{n+m}{m}^{-1} \sum_{k=\lceil k^*\rceil}^{m} [u^k d^{m-k} S_0 - K] \binom{s+k-1}{k}\binom{n-s+m-k}{m-k} \\
&\qquad = \binom{n+m}{m}^{-1} \sum_{k=0}^{m} [K-S_m] \binom{s+k-1}{k}\binom{n-s+m-k}{m-k} \\
&\qquad = K - \underline{S}(m)
\end{aligned} \tag{2.26}
$$

where $\underline{S}(m) = \binom{n+m}{m}^{-1} \sum_{k=0}^{m} S_m \binom{s+k-1}{k}\binom{n-s+m-k}{m-k}$ and shows the maximum buying stock price that an investor is willing to buy at.

Note that all the cases we discussed above, only boundary prices of the interval are involved, in other words, only bounds of the price interval hold the equilibrium like the put-call parity in classic theory. In our method, there exists bid-ask spread, the interval between the maximum buying price and the minimum selling price, meaning that unlike the market in which the classical model is made, in our market there are arbitrage opportunities. Since the NPI investor is only willing to buy an option at a lower price than the lower boundary price and sell one for a price over the higher bound, so apart from boundary prices, all actions that are going to be taken by investors will create an arbitrage opportunity. The arbitrage opportunities are clearer to be demonstrated by Figure 2.2, in which strategy, buying a call option and selling a put option. From the figure, we can see that other than bounds, the value of a portfolio, buying a call option and selling a put option is greater than the portfolio $K \cdot B(t,T) - \underline{S}(0)$. When $c_e < \underline{c_e}$, $p_e > \overline{p_e}$, then $p_e - c_e \geq K \cdot B(t,T)\underline{S}(0)$, where c_e and p_e in this case is the prices of the action options. An investor can gain a certain profit N, by taking action as discussed in Table 2.1, buying the call option at c_e, and selling the call position, involving selling the underlying stock at $\underline{S}(0)$, and the corresponding put option at p_e and lending $K \cdot B(t,T)$ money

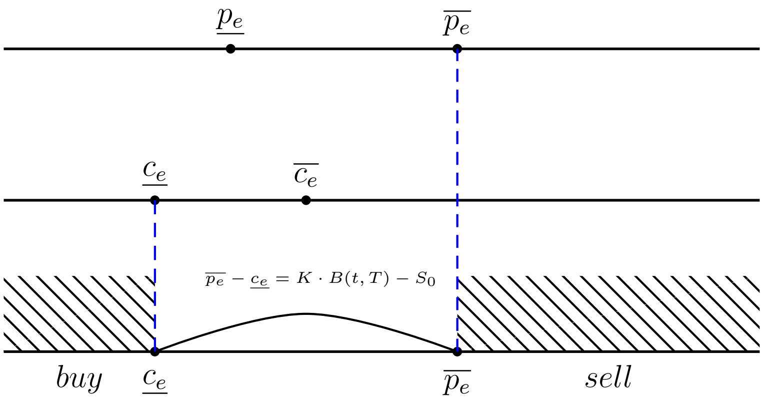


Figure 2.2: The changes of put-call parity for interval prices

from the bank. And when it comes to the other situation, $c_e > \overline{c_e}$, $p_e < \underline{p_e}$, then $c_e - p_e \geq \overline{S}(0) - K \cdot B(t,T)$, one can get money M by selling the call option and buying a call position, including buying the corresponding put option at p_e, the underlying stock at $\overline{S}(0)$, and borrowing money $K \cdot B(t,T)$ from bank.

It is evident that as if the NPI investor can find the chance and take action in the market, the investor will gain a certain profit by manipulating the products in this market. This property also supports the view that imprecise probability generalizes classical probability theory from the perspective of easing representation of uncertainty [6]. Moreover, imprecise prices offer an opportunity to consider investors' risk attitudes during the pricing process. From aforementioned statements, we know that an investor will buy an option when the price is lower than or equal to the lower boundary, and sell the option when the price is greater than or equal to the upper boundary. When it comes to values between the boundaries, an investor may choose to hold the security and watch the markets. The investor considered here is the risk-averse investor. In reality, whether an investor takes action or not depends on one's risk attitude, which can be estimated by risk aversion function, and more discussion will be studied in the future.

2.3 Performance study

In this section, we study the performance of the NPI method for European options in comparison to the CRR model. If there are only two investors in the option market, the CRR investor and the NPI investor, we would like to see the expected profit and loss (P&L) of the NPI investor in the trades with the CRR investor. We consider two scenarios, first the CRR model correctly captures the future market trend, meaning the real market stock price at maturity equals the expected stock price from the CRR model. In the second scenario, the CRR model is wrong about the future market trend, indicating the real market probability is different from the risk-neutral probability. Whereas in both scenarios the NPI method predicts based on the historical data.

In these two scenarios, we would like to compare payoffs from the NPI method and CRR model based on the same option with the same underlying stock at first to eliminate the error of using the expected stock return as the discount rate. By ignoring time discounting at this stage, we can regard the expected payoffs as the expected price and denote the NPI upper and lower expected prices as $\overline{V}$ and $\underline{V}$, as well as the expected CRR price V^{CRR}, equal to the expected option payoffs $\overline{E}$, $\underline{E}$ and E^{CRR}, respectively. When m is fixed, the key factor in the comparison is s. Other factors of the NPI method and the CRR model are also fixed, including the number of historical data n. Different values of s lead to different NPI payoffs, so compared to the CRR payoff this will result in different trading actions. The CRR payoff formulated by Equation (1.9) for a call option or Equation (1.10) for a put option with $r_f = 0$ is a constant value, while NPI payoffs vary with s. There are three trading cases according to s. When the CRR payoff is in between the maximum buying and the minimum selling NPI payoffs, there is no trade between the two investors. Otherwise, the NPI investor will either sell an option or buy it depending on whether the CRR payoff is lower than the maximum buying NPI payoff or higher than the minimum selling NPI payoff. As the worst result that the NPI investor will have eventually is the focus of evaluation, the NPI investor is assumed to quote the maximum buying price or the minimum selling price, and any trade occurs at these prices.

2.3.1 Scenario 1: The CRR investor is correct

In this scenario, we assume that the CRR investor is correct, meaning the upward movement probability in the real market p is equal to the risk-neutral probability q used by the CRR investor, and the option payoff has the same value as the expected payoff from the CRR model. Equations (2.10) and (2.12) are applied to compute the NPI call option bound payoffs, and Equations (2.11) and (2.13) for the NPI put option bound payoffs. Each of the expected minimum selling payoff and the expected maximum buying payoff have an intersection point with the expected CRR payoff, and we note each two intersection points as s_1 and s_2 for call option ($s_1 \leq s_2$) and s_3 and s_4 for put option ($s_3 \leq s_4$). In this case the value s_q, the number of success historical data under the constraint $\frac{s_q}{n} = q$, is in the interval of two intersection points, $s_1 \leq s_q \leq s_2$ for the call option and $s_3 \leq s_q \leq s_4$ for the put option, for if $s = s_q$ the CRR expected value is between the lower and upper expected values. Therefore, for the call option there exist inequalities $\frac{s_1}{n} \leq q \leq \frac{s_2}{n}$ and for the put option there exist inequalities $\frac{s_3}{n} \leq q \leq \frac{s_4}{n}$. After the analytic study for NPI payoff patterns, we learn for the call option the maximum buying payoff and the minimum selling payoff increase as s increases whereas for the put option they decrease as s increases. Then different trading actions of the NPI investor according to different s are presented below.

For the call option:

Case 1.1: $s \geq s_2$

In this case, because of $\frac{s}{n} \geq \frac{s_2}{n} \geq q$, the NPI investor would be more optimistic than the CRR investor about the underlying stock future price, and the expected maximum buying price $\underline{V_c}$ is higher than the fair price V_c^{CRR} from the CRR model, so the NPI investor would like to buy a call option. As in this scenario the CRR investor is right, the loss of the NPI investor in this case depends on this option exercise.

Under the situation that at maturity this call option is exercised the loss of the NPI investor can be formulated as below. If the NPI maximum buying price $\underline{V_c}$ is quoted, so the NPI investor needs to pay the buying price as the payment for this call option and gain the profit from the payoffs $S_T - K_c$. Then the loss L for the NPI investor in this case under this situation is:

$$\begin{aligned}
&L(n, m, s : s \geq s_2 | V = \underline{V_c}) = \underline{V_c} - S_T + K_c \\
&= \underline{E_c}[S_m - K_c]^+ - E_c^{CRR}[S_m - K_c]^+ \\
&= \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \\
&\times \left[\binom{m+n}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} - \binom{m}{k} q^k (1-q)^{m-k} \right] \\
&= \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m}{k} \left[\binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{s}{s+k} - q^k (1-q)^{m-k} \right] \qquad (2.27)
\end{aligned}$$

Here $\underline{V_c} - V_c^{CRR} = \underline{E_c}[S_m - K_c]^+ - E_c^{CRR}[S_m - K_c]^+$. Since the prediction of the CRR investor is totally right, the payoff at maturity $S_T - K_c$ is equal to the CRR expected payoff $E_c^{CRR}[S_m - K_c]^+$.

The other circumstance is that the NPI investor does not exercise this call option causing a loss of this call option price $\underline{V_c}$ as calculated in Equation (2.28).

$$\begin{aligned}
&L(n, m, s : s \geq s_2 | V = \underline{V_c}) = \underline{V_c} \\
&= \underline{E_c}[S_m - K_c]^+ \\
&= \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m+n}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \qquad (2.28)
\end{aligned}$$

Case 1.2: $s_2 > s > s_1$

In this case, there is no action between the two investors, as the market price is higher than the NPI investor's expected maximum buying price and lower than the expected minimum selling price, $\underline{V_c} < V_c^{CRR} < \overline{V_c}$;

$$L(n, m, s : s_2 > s > s_1) = 0 \tag{2.29}$$

Case 1.3: $s \leq s_1$

In this case, the expected CRR price V_c^{CRR} is higher than the minimum selling price of the NPI investor $\overline{V_c}$, so the NPI investor would like to sell a call option. If we want to value the loss of the NPI investor L, two situations happen according to this option exercise. First, when the call option is exercised eventually, longing a call option is the wise action. As we assume the CRR investor is correct, which means the opposite action taken by the NPI investor is wrong, so the NPI investor will face an amount to lose for selling a call option. In this case, two parts are constituting this profit and loss; one part is the payoffs spread $S_T - K$, where S_T is the actual stock price at maturity. The other part is the profit gained by selling this call option $\overline{V_c}$, and under our assumptions we use the expected payoff instead of the price $\overline{V_c} = \overline{E_c}[S_m - K_c]^+$. Due to the assumption about the perfection of the CRR model, that the stock price at maturity equals to the expected stock price, the payoff spread at maturity is equal to the expected CRR payoff, $S_T - K = E_c^{CRR}[S_m - K_c]^+$. The formula below calculates the loss of the NPI investor when the minimum selling price is quoted:

$$
\begin{aligned}
&L(n,m,s:s_1 \geq s|V=\overline{V_c}) = S_T - K_c - \overline{V_c} \\
&= E_c^{CRR}[S_m - K_c]^+ - \overline{E_c}[S_m - K_c]^+ \\
&= \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \\
&\times \left[\binom{m}{k} q^k (1-q)^{m-k} - \binom{m+n}{m}^{-1} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \right] \\
&= \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m}{k} \left[q^k (1-q)^{m-k} - \binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{n-s}{n-s+m-k} \right]
\end{aligned}
\tag{2.30}
$$

For the other situation, when this call option is not exercised, selling a call option is a good choice leading to a profit of the NPI investor, the call option price. The loss of the NPI investor in this situation is harmful as follows:

$$
\begin{aligned}
&L(n,m,s:s_1 \geq s|V=\overline{V_c}) = -\overline{V_c} \\
&= -\overline{E_c}[S_m - K_c]^+ \\
&= -\sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m+n}{m}^{-1} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k}
\end{aligned}
\tag{2.31}
$$

Expected loss of the NPI investor for a call option

After the calculation of the loss of the NPI investor comparing to the correct CRR investor according to each different s cases, we would like to evaluate the expected loss. Given what we have discussed above, s follows the binomial distribution $s \sim Bin(n,q)$ in this example, so the expected loss $L(q)$ can be formulated as:

If the call option is exercised, the expected loss of the NPI investor is

$$
\begin{aligned}
& E_c[L(s)] \\
&= \sum_{s=0}^{s_1} L(n,m,s : s < s_1 | V = \overline{V_c}) \binom{n}{s} q^s (1-q)^{n-s} \\
&+ \sum_{s=s_2}^{n} L(n,m,s : s > s_2 | V = \underline{V_c}) \binom{n}{s} q^s (1-q)^{n-s} \\
&= \sum_{s=0}^{s_1} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m}{k} \binom{n}{s} \\
&\times \left[q^{k+s} (1-q)^{n+m-s-k} - \binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{n-s}{n-s+m-k} q^s (1-q)^{n-s} \right] \\
&+ \sum_{s=s_2}^{n} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m}{k} \binom{n}{s} \\
&\times \left[\binom{n+m}{s+k}^{-1} \binom{n}{s} q^s (1-q)^{n-s} \frac{s}{s+k} - q^{k+s} (1-q)^{n+m-s-k} \right]
\end{aligned}
\tag{2.32}
$$

If not, the expected loss of the NPI investor is

$$
\begin{aligned}
& E_c[L(s)] \\
&= \sum_{s=0}^{s_1} L(n,m,s : s < s_1 | V = \overline{V_c}) \binom{n}{s} q^s (1-q)^{n-s} \\
&+ \sum_{s=s_2}^{n} L(n,m,s : s > s_2 | V = \underline{V_c}) \binom{n}{s} q^s (1-q)^{n-s} \\
&= -\sum_{s=0}^{s_1} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \\
&\times \binom{m+n}{m}^{-1} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \binom{n}{s} q^s (1-q)^{n-s} \\
&+ \sum_{s=s_2}^{n} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \\
&\times \binom{m+n}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \binom{n}{s} q^s (1-q)^{n-s}
\end{aligned}
\tag{2.33}
$$

For the put option:

Case 1.4: $s \geq s_4$

The CRR price is higher than the minimum NPI selling price, so the NPI investor would like to sell this put option and gain the put option price $\overline{V_p}$. If this put option has a negative payoff at maturity, then, in this case, the loss L can be calculated as:

$$\begin{aligned}
&L(n, m, s : s \geq s_4 | V = \overline{V_p}) = -\overline{V_p} \\
&= -\overline{E_p}[K_p - S_m]^+ \\
&= -\sum_{k=0}^{\lfloor k_p^* \rfloor}[K_p - u^k d^{m-k} S_0] \binom{n+m}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k}
\end{aligned} \tag{2.34}$$

Here V is the price that the NPI investor quoted in the market equal to the minimum NPI selling price. Like what happened in the call option, rather than actual prices $\overline{V_p}$ we use the minimum selling payoff $\overline{E_p}[K_p - S_m]^+$, because we try to avoid influences by the discount factor at the start of our study.

However, if this put option has a positive payoff, selling a put option is not smart, as it will cause some loss from this put option exercise by the CRR investor, then the NPI investor needs to pay the payoff $K_p - S_T$. The loss of the NPI investor is,

$$\begin{aligned}
&L(n, m, s : s \geq s_4 | V = \overline{V_p}) = K_p - S_T - \overline{V_p} \\
&= E_p^{CRR}[K_p - S_m]^+ - \overline{E_p}[K_p - S_m]^+ \\
&= \sum_{k=0}^{\lfloor k_p^* \rfloor}[K_p - u^k d^{m-k} S_0] \\
&\times \left[\binom{m}{k} q^k (1-q)^{m-k} - \binom{n+m}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \right] \\
&= \sum_{k=0}^{\lfloor k_p^* \rfloor}[K_p - u^k d^{m-k} S_0] \binom{m}{k} \left[q^k (1-q)^{m-k} - \binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{s}{s+k} \right]
\end{aligned} \tag{2.35}$$

The payoff of this put option at maturity is identical to the expected value of this put option from the CRR model, $K_p - S_T = E_p^{CRR}[K_p - S_m]^+$, under the assumption of the CRR model perfection.

Case 1.5: $s_4 > s > s_3$

In this case, the CRR price is in the interval of NPI prices, so there is no transaction when this case is encountered. Therefore, in this case, there is no loss.

$$L(n, m, s : s_4 > s > s_3) = 0 \tag{2.36}$$

Case 1.6: $s \leq s_3$

The CRR price is lower than the maximum NPI buying price. The NPI investor will buy a put option from the CRR investor paying this put option price $\underline{V_p}$. If the NPI investor buys the right to exercise from the market but not exercise at maturity, the put option price quoted at the maximum NPI buying price is the loss.

$$\begin{aligned}
&L(n, m, s : s \leq s_3 | V = \underline{V_p}) = \underline{V_p} \\
&= \underline{E_p}[K_p - S_m]^+ \\
&= \sum_{k=0}^{\lfloor k_p^* \rfloor} [K_p - u^k d^{m-k} S_0] \binom{m+n}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k}
\end{aligned} \tag{2.37}$$

Under the same assumptions as that in the first scenario of this put option, the price is taken place by the payoff $\underline{V_p} = \underline{E_p}[K_p - S_m]^+$. But if this put option is exercised, the NPI investor can get the payoff of this put option $K_p - S_T$. The loss of the NPI investor in this situation is:

$$
\begin{aligned}
&L(n,m,s:s\leq s_3|V=\underline{V_p}) = \underline{V_p} - K_p + S_T \\
&= \underline{E_p}[K_p - S_m]^+ - E_p^{CRR}[K_p - S_m]^+ \\
&= \sum_{k=0}^{\lfloor k_p^* \rfloor}[K_p - u^k d^{m-k} S_0] \\
&\times \left[\binom{m+n}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} - \binom{m}{k} q^k (1-q)^{m-k} \right] \\
&= \sum_{k=0}^{\lfloor k_p^* \rfloor}[K_p - u^k d^{m-k} S_0] \binom{m}{k} \left[\binom{m+n}{s+k}^{-1} \binom{n}{s} \frac{n-s}{n-s+m-k} - q^k (1-q)^{m-k} \right]
\end{aligned} \tag{2.38}
$$

Expected loss of the NPI investor for a put option

After calculating the loss in each case, we would like to explore the value of the expected loss for this put option with the same underlying stock according to $s \sim Bin(n,q)$, and the formulae are listed below:

The put option payoff is negative, then the expected loss for this put option is

$$
\begin{aligned}
&E_p[L(s)] \\
&= \sum_{s=0}^{s_3} L(n,m,s:s\leq s_3|V=\underline{V_p}) \binom{n}{s} q^s (1-q)^{n-s} \\
&+ \sum_{s=s_4}^{n} L(n,m,s:s\geq s_4|V=\overline{V_p}) \binom{n}{s} q^s (1-q)^{n-s} \\
&= \sum_{s=0}^{s_3} \sum_{k=0}^{\lfloor k_p^* \rfloor}[K_p - u^k d^{m-k} S_0] \binom{m+n}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \binom{n}{s} q^s (1-q)^{n-s} \\
&- \sum_{s=s_4}^{n} \sum_{k=0}^{\lfloor k_p^* \rfloor}[K_p - u^k d^{m-k} S_0] \binom{n+m}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \binom{n}{s} q^s (1-q)^{n-s}
\end{aligned} \tag{2.39}
$$

Otherwise, the expected loss for this put option is

$$
\begin{aligned}
&E_p[L(s)] \\
&= \sum_{s=0}^{s_3} L(n,m,s: s \le s_3 | V = \underline{V_p}) \binom{n}{s} q^s (1-q)^{n-s} \\
&+ \sum_{s=s_4}^{n} L(n,m,s: s \ge s_4 | V = \overline{V_p}) \binom{n}{s} q^s (1-q)^{n-s} \\
&= \sum_{s=0}^{s_3} \sum_{k=0}^{\lfloor k_p^* \rfloor} [K_p - u^k d^{m-k} S_0] \binom{m}{k} \binom{n}{s} \\
&\times \left[\binom{m+n}{s+k}^{-1} \binom{n}{s} \frac{n-s}{n-s+m-k} q^s (1-q)^{n-s} - q^{k+s} (1-q)^{n-s+m-k} \right] \\
&+ \sum_{s=s_4}^{n} \sum_{k=0}^{\lfloor k_p^* \rfloor} [K_p - u^k d^{m-k} S_0] \binom{m}{k} \binom{n}{s} \\
&\times \left[q^{k+s} (1-q)^{n-s+m-k} - \binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{s}{s+k} q^s (1-q)^{n-s} \right] \qquad (2.40)
\end{aligned}
$$

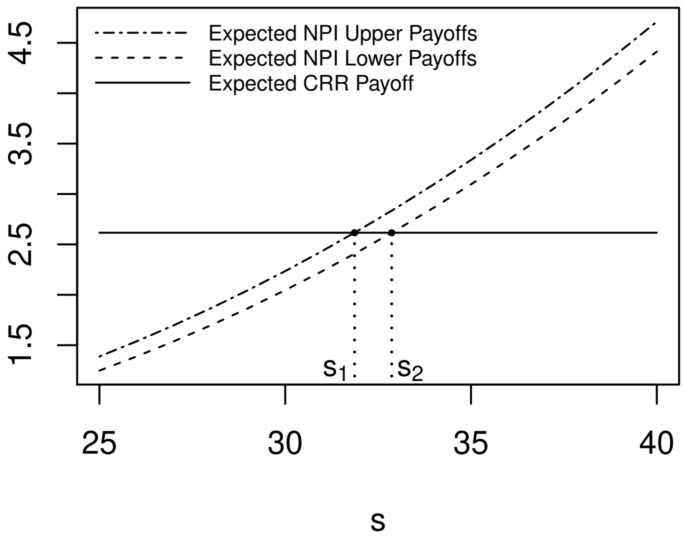
An interesting characteristic is disclosed in the NPI expected loss formulae for both the call and the put option in this scenario. We started with imprecise NPI prices, but ended up getting a precise expected loss, for each s, the trading action for the NPI investor is determined compared to the CRR price, so only one NPI price is taken into account for each case, and the loss becomes an explicit value as action price is settled. This explicit value of expected NPI loss is convenient for us to compare the two pricing methods.

Example 2.1

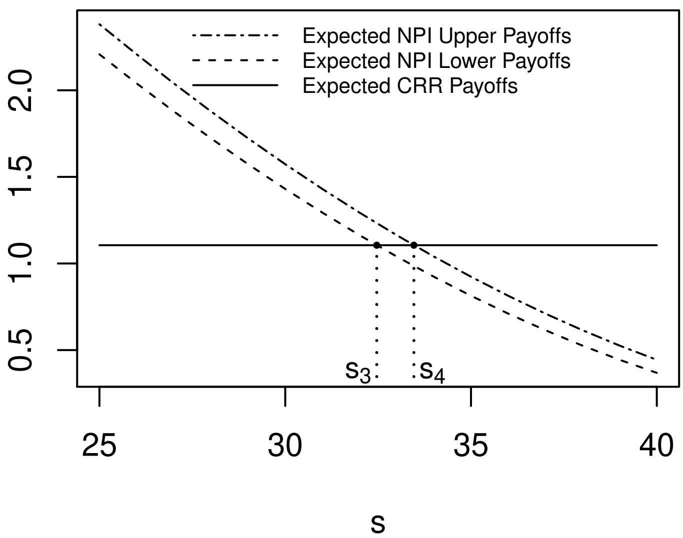
After discussing three trading cases for each type of option according to s, we would like to compare the payoffs in an example. At first, we need to define some input values. For the binomial tree, the initial stock price $S_0 = 20$, and at every next step this stock price will either go up with the upward factor $u = 1.1$ or go down with the downward factor $d = 0.9$. We set the same strike price $K_c = K_p = 21$ for both the call option and the put option. We set a risk-neutral probability q equal to

0.65, which is identical to the real market probability of movements, $q = p = 0.65$. Since we assume the CRR model is right in this scenario, then the proportion of upward movements $\frac{s}{n}$ of historical data should follow the CRR prediction. To do this analytical study, understanding patterns of payoffs according to s and calculating the expected profit or loss of the NPI investor, if total historical data n is equal to 50, s will follow the binomial distribution $s \sim Bin(50, 0.65)$. In this example, we will first plot the patterns of all payoffs with a fixed maturity, then we would like to know the expected loss of the NPI investor with varying m. Finally, since the NPI method is based on historical data, we want to check if the expected loss of the NPI investor will get better when we enlarge the historical data n.

We want to compare the two methods with $m = 4$. Payoff patterns for call and put options are plotted in Figure 2.3, and it displays that three cases for each type option we mentioned above. In this example, the values of intersection points between NPI payoffs and the CRR payoff are gained using Newton-Raphson root-finding method [60, 66]. For a call option, the intersection between s_1 value of the NPI upper payoff and the CRR payoff equals to 31.86541, and the intersection s_2 of the NPI lower payoff and the CRR payoff equals to 32.8654. For this put option, s_3 equal to 32.46275, and s_4 equals to 33.46276, which are the points of the NPI lower payoff intersecting with the CRR payoff and the NPI upper payoff intersecting with the CRR payoff. According to s values of the intersection, we can tell when $\frac{s}{n}$ is equal to values near to q, the CRR payoff in the interval of NPI payoffs, no trading action exists in this circumstance. When s falls outside the intersection interval $[s_1, s_2]$ or $[s_3, s_4]$, the NPI investor and the CRR investor will trade with each other, then NPI investor will either gain profit from the CRR investor or lost the money. Since we assume the CRR investor is always right, so we expect NPI investor to face an amount of loss during their trades. As the loss of the NPI investor for different cases can be estimated, and we know the distribution s follows, the expected loss in this scenario for the NPI investor is available.



(a) Expected payoffs of the call option

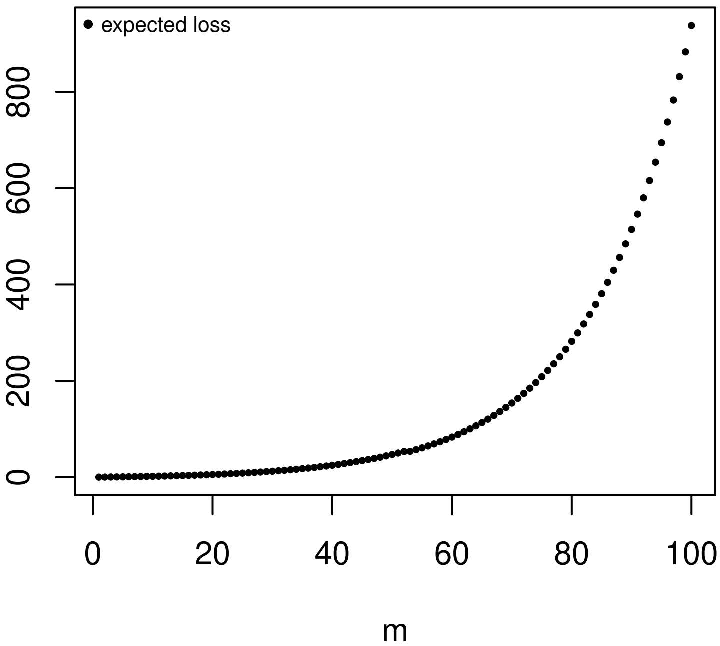


(b) Expected payoffs of the put option

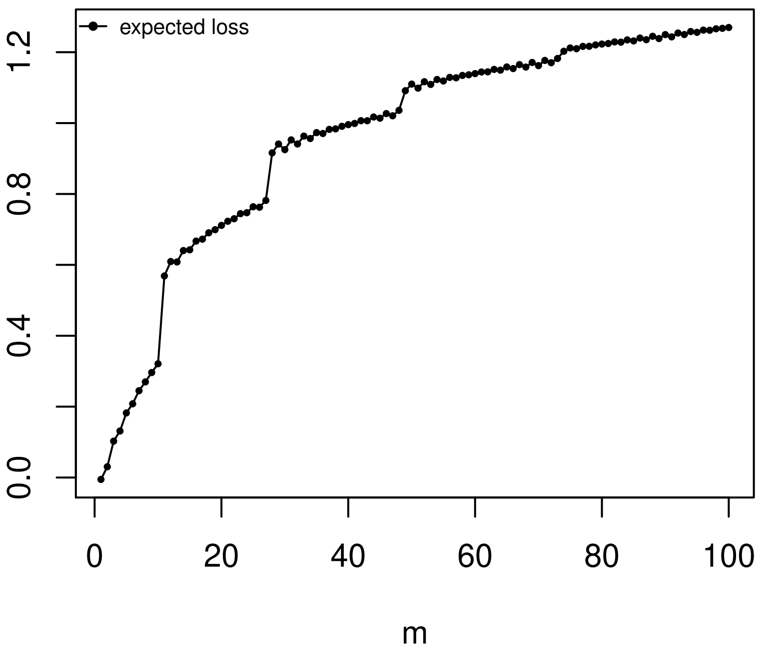
Figure 2.3: Expected payoff of European options from both the NPI method and the CRR model in Scenario 1

As we know, if the NPI investor has decided to invest in an option based on a specific underlying stock, all input values are fixed except the number m of future steps. Then the influence of varying m toward expected NPI losses is of interest. In this example, the call option is exercised but the put option not, so Equations (2.32) and (2.39) are used to calculate the expected loss of the NPI investor, and we reveal expected NPI losses with various m in Figure 2.4.

From Figure 2.4, there is no doubt no matter call or put option the NPI investor decides to invest in and how long the maturity is, the NPI investor is always expected to face an amount of loss. What is more, the expected loss manifests that it is wise to take part in short-term investment rather than long-term one for the NPI investor, because the NPI expected loss increases more than linearly as m increases. The reason that the expected loss pattern shows the convexity as m increases is the pattern of the NPI payoff for call option gets more and more convex along with increasing m, and the part $s \geq s_2$ of the NPI expected payoff for call option takes a big part of the expected loss. However, there is a gap in the graph of expected loss for the call option, and for the put option, the expected NPI loss is in a stairs type raise. The reason for these is payoffs' intersections movements when m increases, shown in Figure 2.5. In this figure, we plotted s integer value of varying intersection points, s_1 and s_2 for the call option and s_3 and s_4 for the put option, along with increasing m. As we illustrated in the expected loss formulas, the expected loss consists of differently weighted losses in different cases according to s, and intersection movements affect probabilities of each part losses. Since s_1, s_2, s_3, and s_4 in the NPI pricing formulas are supposed to be integers, and s_1 and s_3 are transferred to the first integer less than their values, while s_2 and s_4 become the first integer greater than their values. It is clear from Figure 2.5 that when we change intersection values into their corresponding integers, for the call option there is a step down of s_1 and s_2 which explains why there is a sudden decrease gap in the expected loss pattern for the call option. For put option s_3 and s_4 increase



(a) Expected loss of the call option



(b) Expected loss of the put option

Figure 2.4: Expected loss for the NPI investor in Scenario 1

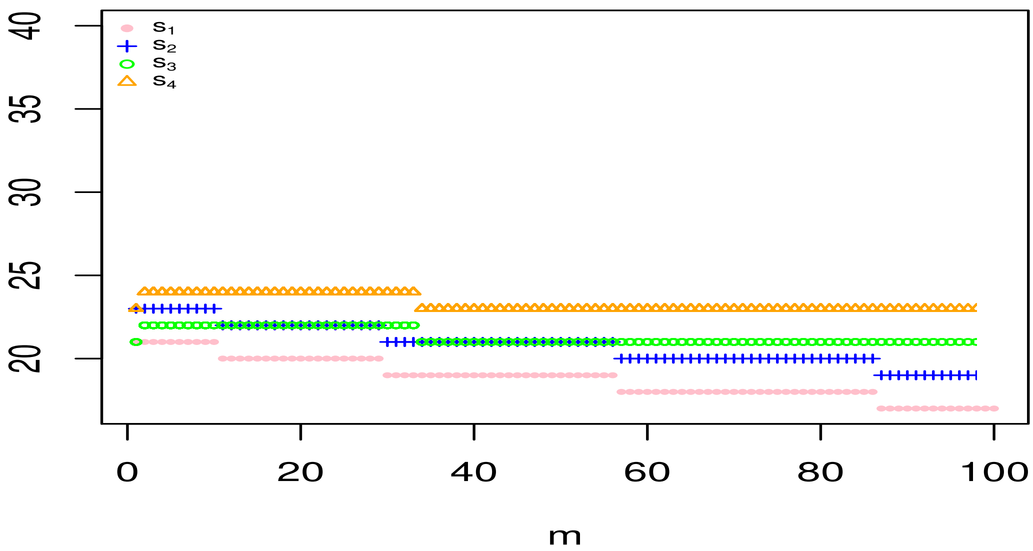


Figure 2.5: Intersection s move with varying m

values in steps with m, resulting in a stairs type expected loss increase. All these characteristics of s intersection values can explain the pattern of the expected loss.

Although under assumptions of this scenario the NPI investor will always pay for the wrong prediction, this situation can be improved if more historical information can be reached. A 3D plot for a call option shown in Figure 2.6, the expected loss of the NPI investor with increasing historical data n and varying maturity m, supports this statement. It denotes that for each maturity m, as we raise n, the expected loss decreases, except when n and m are both very small. When n is low, the interval between the maximum buying price and minimum selling price is vast, and when m is very small, the patterns of the NPI prices resemble a straight line. Therefore, a small amount of loss from the $s \geq s_2$ and a small amount of profit $s \leq s_1$ cancel each other out. However, as m is not too small, we can minimize the expected loss by increasing n. When n is small each increment of probability in each time step, for instance $\frac{s}{n+1}$ moving to $\frac{s+1}{n+2}$, changes greatly. Whereas for larger n, lower and upper probabilities at every step are more stable and approaching to q for calculating the expected loss of the NPI method in this scenario $s \sim Bin(n, q)$. Illustration from the financial aspect also makes sense, which when an investor has more trustable

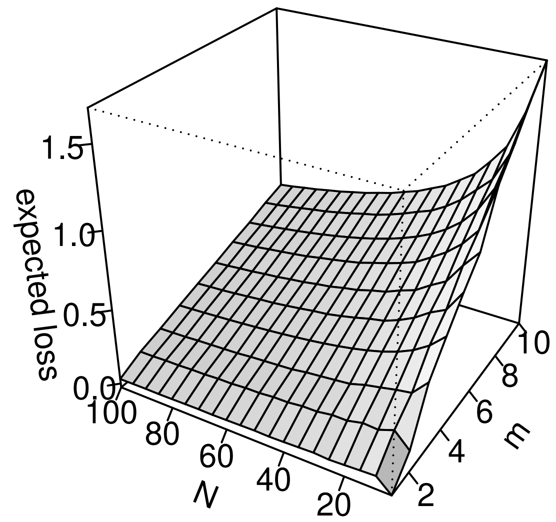


Figure 2.6: Influence on the expected loss with increasing historical data (call option): $n = N \times m$

historical information, his prediction is more accurate compared to the market, and there is less chance he will lose money. Overall, under the assumption that the CRR investor knows every information to forecast a right price, the NPI investor would not be expected to perform better than the CRR investor, and the longer the NPI investor in this game the more expected losses he will give up on. However, since our method keeps learning from historical data, if there are more historical data available the loss will decrease.

2.3.2 Scenario 2: The CRR investor is wrong

In this scenario, there are also three possible trading actions the NPI investor may take according to the value of s. As in Scenario 1, for the call and put options there are two intersections between the NPI expected prices and the CRR expected price, and these intersections' positions depend on q. We note the intersections for call option as s_5 and s_6 ($s_5 \leq s_6$) and for put option as s_7 and s_8 ($s_7 \leq s_8$).

The relationship happening in Scenario 1 still valid, $\frac{s_5}{n} \leq q \leq \frac{s_6}{n}$ for the call option $\frac{s_7}{n} \leq q \leq \frac{s_8}{n}$ for the put option. However, in this scenario the real market probability p is different from the risk-neutral probability q, and we assume that historical data can reflect the market at some level, $s \sim Bin(n,p)$. Thus, even though there still exist three cases of trading according to the value of s, the case that has the highest chance to occur is $\frac{s}{n}$ around p but q. We expect the NPI investor will get some profit since the CRR investor is wrong. The profit of the NPI investor in three cases for each type option is listed below:

For the call option:

Case 2.1: $s \leq s_5$

When $s \leq s_5$ the NPI investor would like to sell a call option to the CRR investor and gain the call option price $\overline{V_c}$ as the profit. If this call option will have a positive payoff, then selling a call option will cause some loss from this call option exercise by the CRR investor, $S_T - K_c$. Then the profit $Pro(\cdot)$ of the NPI investor is:

$$
\begin{aligned}
& Pro(n, m, s : s \leq s_5 | V = \overline{V_c}) = \overline{V_c} - S_T + K_c \\
& = \overline{E_c}[S_m - K_c]^+ - E_c^{CRR}(p) \\
& = \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \\
& \times \left[\binom{m+n}{m}^{-1} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} - \binom{m}{k} p^k (1-p)^{m-k} \right] \\
& = \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m}{k} \left[\binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{n-s}{n-s+m-k} - p^k (1-p)^{m-k} \right]
\end{aligned} \tag{2.41}
$$

Here we neglect the difference between option payoffs and option premiums $\overline{V_c} = \overline{E_c}[S_m - K_c]^+$ because of the assumption that the contract settlement date is close to the expiration date. As the call option payoff at maturity is hard to estimate, we used the expected value from the CRR model with the probability p.

Another situation is that this call option has a negative payoff, then the NPI investor can earn this call option price without worrying about the CRR investor will exercise it at maturity. And the profit for this case can be formulated as follows:

$$\begin{aligned}
&Pro(n, m, s : s \leq s_5 | V = \overline{V_c}) = \overline{V_c} \\
&= \overline{E_c}[S_m - K_c]^+ \\
&= \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m+n}{m}^{-1} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k}
\end{aligned} \tag{2.42}$$

Case 2.2: $s_5 < s < s_6$

There is no transaction in this case, for the CRR price falls in the interval between the minimum selling price and the maximum buying price.

$$Pro(n, m, s : s_6 > s > s_5) = 0 \tag{2.43}$$

Case 2.3: $s \geq s_6$

When $s \geq s_6$ occurs, the CRR expected price is lower than the NPI maximum buying price, so that the NPI investor will buy this call option. So if at maturity this call option is exercised, then this trading is active. The NPI investor needs to pay this call option price but wins the payoff at maturity.

$$
\begin{aligned}
&Pro(n, m, s : s \geq s_6 | V = \underline{V_c}) = S_T - K_c - \underline{V_c} \\
&= E_c^{CRR}(p) - \underline{E_c}[S_m - K_c]^+ \\
&= \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \\
&\times \left[\binom{m}{k} p^k (1-p)^{m-k} - \binom{m+n}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \right] \\
&= \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m}{k} \left[p^k (1-p)^{m-k} - \binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{s}{s+k} \right]
\end{aligned} \tag{2.44}
$$

If this call option is not exercised at maturity, although the CRR investor made the wrong prediction for the stock upward movement probability, the historical data providing the information is worse than the CRR prediction. The result misleads the NPI investor to a wrong decision, buying a call option, and this will cause an amount of loss. The loss is the maximum buying price from our NPI method. Therefore, the loss of the NPI investor is,

$$
\begin{aligned}
&L(n, m, s : s \geq s_6 | V = \underline{V_c}) = \underline{V_c} \\
&= \underline{E_c}[S_m - K_c]^+ \\
&= \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m+n}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k}
\end{aligned} \tag{2.45}
$$

Expected profit of the NPI investor for a call option

After listing the profit and loss formulas for three cases, it is time to calculate the expected profit of the NPI investor. In this scenario, $s \sim Bin(n, p)$ and the intersections depend on the expected value from the CRR model, so both p and q influence the expected profit of the NPI investor.

When the call option is exercised, then the expected profit of the NPI investor is

$$
\begin{aligned}
E_c[Pro(p,s)] &= \sum_{s=0}^{s_5} Pro(n,m,s : s \le s_5 | V = \overline{V_c}) \binom{n}{s} p^s (1-p)^{n-s} \\
&+ \sum_{s=s_6}^{n} Pro(n,m,s : s \ge s_6 | V = \underline{V_c}) \binom{n}{s} p^s (1-p)^{n-s} \\
&= \sum_{s=0}^{s_5} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m}{k} \binom{n}{s} \\
&\times \left[p^s (1-p)^{n-s} \binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{n-s}{n-s+m-k} - p^{s+k} (1-p)^{n-s+m-k} \right] \\
&+ \sum_{s=s_6}^{n} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m}{k} \binom{n}{s} \\
&\times \left[p^{s+k} (1-p)^{n-s+m-k} - p^s (1-p)^{n-s} \binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{s}{s+k} \right] \qquad (2.46)
\end{aligned}
$$

Note that this formula depends on q, because values of intersections s_5 and s_6 are calculated according to q. When this call option is not exercised at maturity, the expected profit of the NPI investor is,

$$
\begin{aligned}
E_c[Pro(p,s)] &= \sum_{s=0}^{s_5} Pro(n,m,s : s \le s_5 | V = \overline{V_c}) \binom{n}{s} p^s (1-p)^{n-s} \\
&- \sum_{s=s_6}^{n} L(n,m,s : s \ge s_6 | V = \underline{V_c}) \binom{n}{s} p^s (1-p)^{n-s} \\
&= \sum_{s=0}^{s_5} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \\
&\times \binom{m+n}{m}^{-1} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \binom{n}{s} p^s (1-p)^{n-s} \\
&- \sum_{s=s_6}^{n} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \\
&\times \binom{m+n}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \binom{n}{s} p^s (1-p)^{n-s} \qquad (2.47)
\end{aligned}
$$

For the put option:

Case 2.4: $s \leq s_7$

The CRR price is lower than the maximum NPI buying put option price, which will result in the NPI investor buy this put option from the CRR investor. If this put option is not exercised, then, in this case, the NPI investor will lose the put option fee. The formula of the NPI loss is listed below.

$$\begin{aligned}
&L(n, m, s : s \leq s_7 | V = \underline{V_p}) = \underline{V_p} \\
&= \underline{E_p}[K_p - S_m]^+ \\
&= \sum_{k=0}^{\lfloor k_p^* \rfloor} [K_p - u^k d^{m-k} S_0] \binom{m+n}{m}^{-1} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k}
\end{aligned} \tag{2.48}$$

Here $\underline{V_p} = \underline{E_p}[K_p - S_m]^+$ means we assume the discount procedure can be neglect in this step. When the put option is exercised at maturity, the NPI investor pays the put option price and gain the payoff of this put option $K_p - S_T$. The profit earned by the NPI investor is,

$$\begin{aligned}
&Pro(n, m, s : s \leq s_7 | V = \underline{V_p}) = K_p - S_T - \underline{V_p} \\
&= E_p^{CRR}(p) - \underline{V_p} \\
&= E_p^{CRR}(p) - \underline{E_p}[K_p - S_m]^+ \\
&= \sum_{k=0}^{\lfloor k_p^* \rfloor} [K_p - u^k d^{m-k} S_0] \\
&\times \left[\binom{m}{k} p^k (1-p)^{(m-k)} - \binom{m+n}{m}^{-1} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \right] \\
&= \sum_{k=0}^{\lfloor k_p^* \rfloor} [K_p - u^k d^{m-k} S_0] \binom{m}{k} \left[p^k (1-p)^{(m-k)} - \binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{n-s}{n-s+m-k} \right]
\end{aligned} \tag{2.49}$$

As we can not know the real put option payoff in this scenario, we assume the real payoff is approximately equal to the value calculated by the CRR model with risk-neutral probability p.

Case 2.5: $s_7 < s < s_8$

In this case, s falls in between of s_7 and s_8, where no trading action occurs, for the CRR price is higher than the maximum buying price and lower than the minimum selling price.

$$Pro(n, m, s : s_7 < s < s_8) = 0 \tag{2.50}$$

Case 2.6: $s \geq s_8$

In this case, the CRR expected price is higher than the minimum selling price, and then the NPI investor sells this put option. If the CRR investor is not able to exercise this put option at maturity, the NPI investor will gain the price without any payment, then the profit in this case is,

$$\begin{aligned}
&Pro(n, m, s : s \geq s_8 | V = \overline{V_p}) = \overline{V_p} \\
&= \overline{E_p}[K_p - S_m]^+ \\
&= \sum_{k=0}^{\lfloor k_p^* \rfloor} [K_p - u^k d^{m-k} S_0] \binom{m+n}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k}
\end{aligned} \tag{2.51}$$

But when the stock price is not optimistic, and the CRR investor will exercise the put option, the NPI investor will take a wrong action, selling the put option, which violates to the real market. The NPI investor will face a payment as put option payoffs, which is larger than the profit earned by selling put option price.

$$
\begin{aligned}
&L(n,m,s:s\geq s_8|V=\overline{V_p})=K_p-S_T-\overline{V_p}\\
&=E_p^{CRR}(p)-\overline{E_p}[K_p-S_m]^+\\
&=\sum_{k=0}^{\lfloor k_p^*\rfloor}[K_p-u^kd^{m-k}S_0]\\
&\times\left[\binom{m}{k}p^k(1-p)^{(m-k)}-\binom{m+n}{m}^{-1}\binom{s+k-1}{k}\binom{n-s+m-k}{m-k}\right]\\
&=\sum_{k=0}^{\lfloor k_p^*\rfloor}[K_p-u^kd^{m-k}S_0]\binom{m}{k}\left[p^k(1-p)^{(m-k)}-\binom{n+m}{s+k}^{-1}\binom{n}{s}\frac{s}{s+k}\right]
\end{aligned}
\tag{2.52}
$$

Expected profit of the NPI investor for a put option

Eventually, we can get the expected profit of put option computing as below:

When the investor is optimistic about the underlying asset and this put option is not going to be exercised, the expected profit of the NPI investor is

$$
\begin{aligned}
&E_p=[Pro(p,s)]=\sum_{s=0}^{s_7}L(n,m,s:s\leq s_7|V=\underline{V_p})\binom{n}{s}p^s(1-p)^{n-s}\\
&+\sum_{s=s_8}^{n}Pro(n,m,s:s\geq s_8|V=\overline{V_p})\binom{n}{s}p^s(1-p)^{n-s}\\
&=-\sum_{s=0}^{s_7}\sum_{k=0}^{\lfloor k_p^*\rfloor}[K_p-u^kd^{m-k}S_0]\\
&\times\binom{m+n}{m}^{-1}\binom{s+k}{k}\binom{n-s+m-k-1}{m-k}\binom{n}{s}p^s(1-p)^{n-s}\\
&+\sum_{s=s_8}^{n}\sum_{k=0}^{\lfloor k_p^*\rfloor}[K_p-u^kd^{m-k}S_0]\\
&\times\binom{m+n}{m}^{-1}\binom{s+k-1}{k}\binom{n-s+m-k}{m-k}\binom{n}{s}p^s(1-p)^{n-s}
\end{aligned}
\tag{2.53}
$$

When this put option is exercised, the expected profit of the NPI investor is

$$
\begin{aligned}
E_p[Pro(p,s)] &= \sum_{s=0}^{s_7} Pro(n,m,s: s \leq s_7 | V = \underline{V_p}) \binom{n}{s} p^s (1-p)^{n-s} \\
&- \sum_{s=s_8}^{n} L(n,m,s: s \geq s_8 | V = \overline{V_p}) \binom{n}{s} p^s (1-p)^{n-s} \\
&= \sum_{s=0}^{s_7} \sum_{k=0}^{\lfloor k_p^* \rfloor} [K_p - u^k d^{m-k} S_0] \binom{m}{k} \binom{n}{s} p^s (1-p)^{n-s} \\
&\times \left[p^k (1-p)^{(m-k)} - \binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{n-s}{n-s+m-k} \right] \\
&- \sum_{s=s_8}^{n} \sum_{k=0}^{\lfloor k_p^* \rfloor} [K_p - u^k d^{m-k} S_0] \binom{m}{k} \binom{n}{s} p^s (1-p)^{n-s} \\
&\times \left[p^k (1-p)^{(m-k)} - \binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{s}{s+k} \right] \qquad (2.54)
\end{aligned}
$$

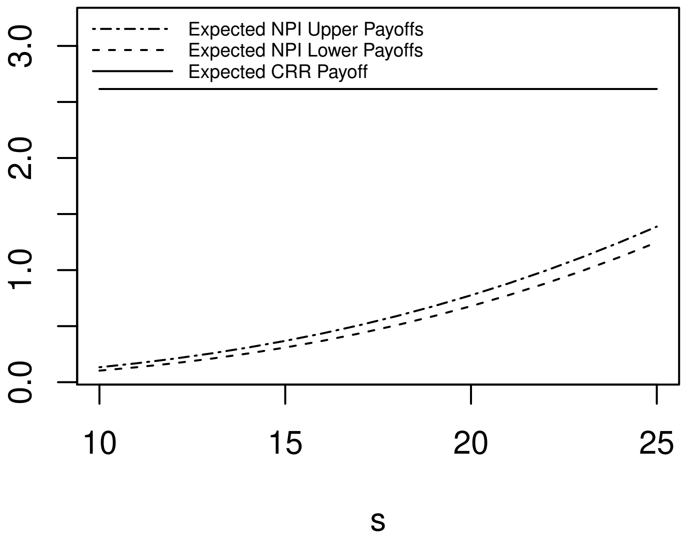
Example 2.2

We have learned about how much profit or loss the NPI investor will face in every case as well as the expected profit. Then we would like to illustrate the performance comparison in an example. As we discuss the NPI method for European options versus the CRR model under the assumption that the CRR model deviates from the real market value in the future, we input the risk-neutral probability $q = 0.65$, while the real market probability of upward movement is $p = 0.25$. This means based on information from the market, the stock's future is not bright, and its price will drop. However, the CRR investor overvalues this stock believing its price will rise, whereas the NPI investor has a high chance to predict it right based on the historical data. All other inputs in this example stays the same as in Example 2.1, $S_0 = 20, K_c = K_p = 21, u = 1.1, d = 0.9, n = 50$. Since q does not change, payoffs calculated from Equations (1.9) and (1.10) stay the same. But NPI results computed with Equations (2.10) and (2.12) for call option and Equations (2.11) and (2.13) for put option are different, for the analytical study of payoff patterns according to s

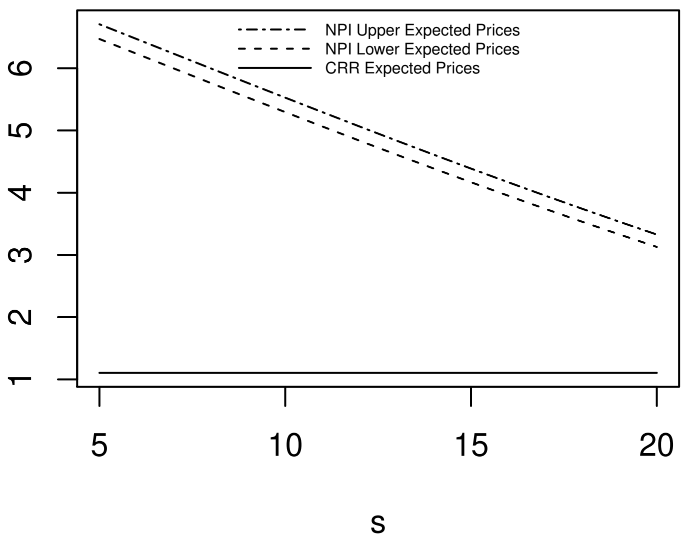
and expected profit or loss calculation in this scenario $s \sim Bin(50, 0.25)$ meaning $\frac{s}{n}$ is around p rather than q. As Example 2.1, we would like to study the performance in three ways. First, we would like to learn the pattern of each expected value from both pricing method with fixed m. Then knowing the expected profit of the NPI investor with varying m is what we intend to do. Finally, we want to check the influence of n on the expected profit of the NPI investor.

As an example, we predict option payoffs in four future steps, $m = 4$, and plot them in Figure 2.7. The whole patterns of NPI payoffs and the CRR payoff with s from 0 to 50 are the same as in Scenario 1, and intersections between our NPI payoffs and the CRR payoff are the same. But to distinguish from intersections in Scenario 1, we use different notations, then $s_5 = s_1 = 31.86541$, $s_6 = s_2 = 32.8654$, $s_7 = s_3 = 32.46275$ and $s_8 = s_4 = 33.46276$. The only different part is the area that s has a high chance to fall in, which is the part of payoffs we mainly focus on, shown in two graphs standing for each option type in Figure 2.7. From the figure, it is clear that for call option the NPI investor has higher chance to sell the call option as well as for put option the NPI investor is more willing to buy the put option from the CRR investor, and both two actions gain profit. There also exist possibilities that the historical data offers worse information than the CRR model's assumptions, issuing in the NPI investor will buy a call option or sell a put option, although the likelihood of that happening is quite low. Overall, we look forward to seeing that in Scenario 2, the NPI investor will gain some profit, and this guess is confirmed by plotting expected profit for both call and put options in Figure 2.8.

As in Scenario 1, after investigating trading actions in all kinds of s cases, we know the exact quote price for each case, finally leading us to a precise expected profit or loss. Based on the expected profit formulas Equations (2.47) and (2.54), it is easier for an investor to choose the maturity with the NPI method. After all, once an option to a specific underlying asset has been settled the only factor which will influence the price is the maturity. We plot the expected profit with varying m

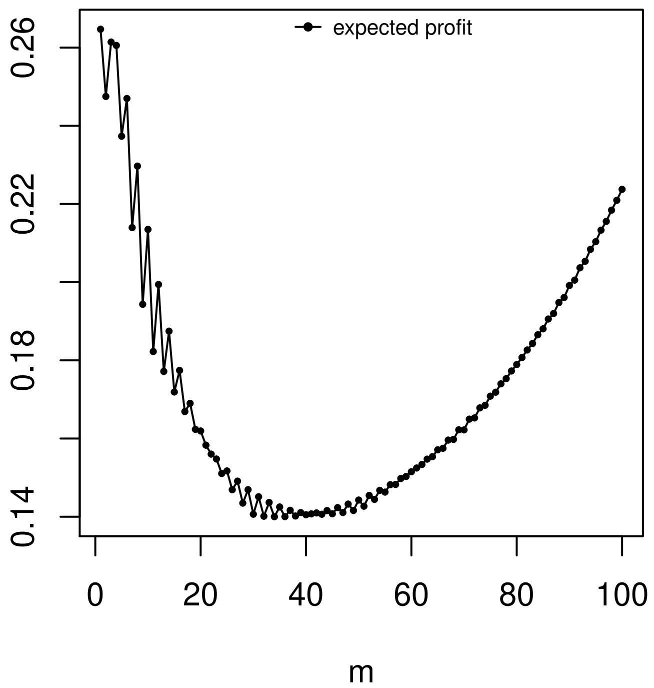


(a) Expected payoffs of the call option

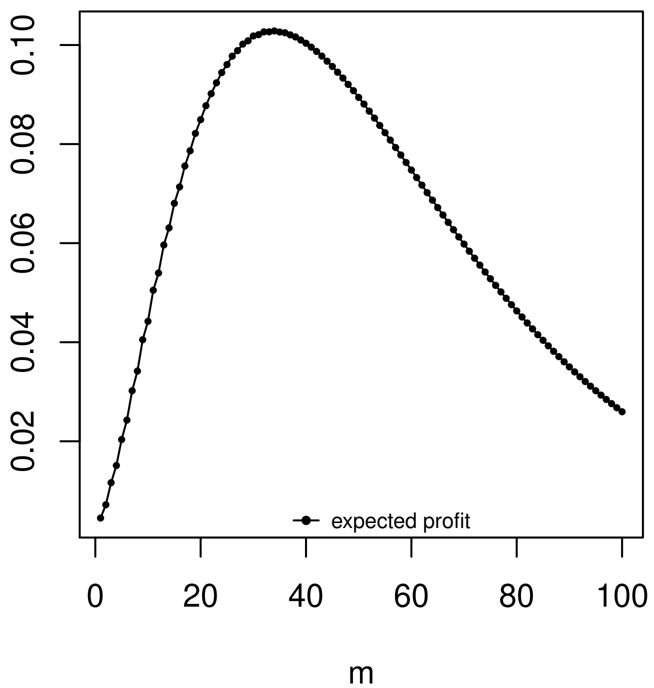


(b) Expected payoffs of the put option

Figure 2.7: Payoffs from the NPI method for European options and the CRR model in Scenario 2



(a) Expected profit of the call option



(b) Expected profit of the put option

Figure 2.8: Expected profit from the NPI method for European options and CRR model in Scenario 2

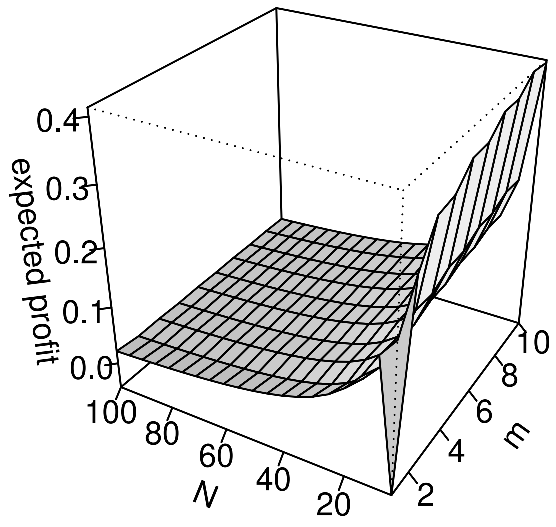


Figure 2.9: Influence on the expected profit with increasing historical data (put option): $n = N \times m$

for the call option and the put option in Figure 2.8.

It turns out, under the assumption that the CRR investor makes a wrong prediction with opposite trading direction, the NPI investor is expected to gain some positive profit no matter which type of option. The price of this call option causes the fluctuation and the convex shape of the NPI profit for the call option. As m increases, the pattern of NPI expected prices for this call option according to s becomes more convex. The concave shape of expected profits for the put option is caused by the competition between the payoff $K_p - S_T$ and the option price, for they have opposite moving directions when m increases. In our example, the stock price will end up going down, so playing with a put option is a safer choice with less profit because the more risk exists the higher return an investor can get. In general, according to the predicted direction of stock price movement, buying a relevant option is better and safer than selling an opposite direction option, which is already commonly agreed in the real market. When we increase n in this scenario, the profit of the NPI investor reduces, except when n and m are both small. The reason for no gain with small n and m is the same as that in Example 2.1, and we plotted the NPI expected a profit for the put option to confirm our point in Figure 2.9. Like what we have discussed in Scenario 1, enlarging historical data makes the forecasting from the NPI method close to market behavior with more stable movement probabilities in the binomial tree, then the difference between the NPI prediction and the CRR real market prediction narrows down. From the perspective of the financial market, the more effective prediction closing to the real market an investor gets, the less opportunity exists to beat the market, so an investor would never take action if the forecast is the same as the real market.

Example 2.3

In this example, we want to see if the CRR model predicts a wrong probability but with the same direction as the real market, $p > 0.5$ and $q > 0.5$ with $q \neq p$ or $p < 0.5$ and $q < 0.5$ with $q \neq p$. All inputs in this example are the same as in Example 2.2, except $q = 0.45$ leading to a different value of the intersections with fixed m. However, the most interesting problem is the expected profit of the NPI investor in this example.

Even though in this example the CRR prediction and the real market have the same direction, the stock price will go downwards, so we should still use Equations (2.47) and (2.54) to calculate the expected profit of the NPI investor. The results of the put option prediction are plotted in Figure 2.10, showing that when n and m are both small the NPI investor will face some loss, because the historical information is not enough for the NPI investor to act effectively and small m makes the NPI put option payoff pattern according to s steeper for the area holding greatest probability. However, if we increase n, both profit and loss will approach zero as we have already discussed in Examples 2.1 and 2.2.

We have performed a more detailed study of the expected profit of the NPI investor for varying values of q given p. In Table 2.2, we set the real market prob-

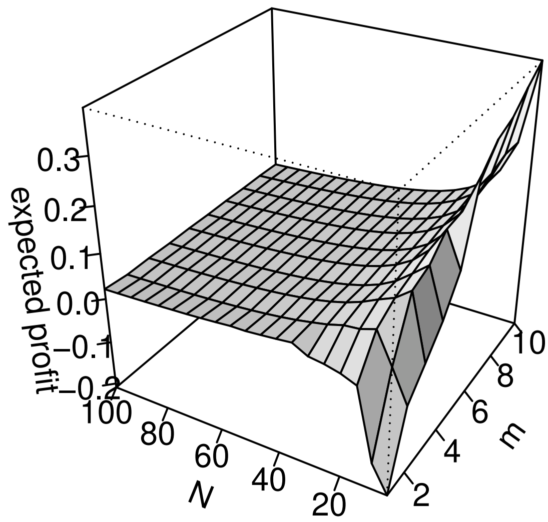


Figure 2.10: Influence on the expected profit with increasing historical data (put option): $n = N \times m$

ability $p = 0.25$, and $S_0 = 20, K_c = K_p = 21, u = 1.1, d = 0.9$. We calculated the expected profit according to different q. In this study, by 40 historical data we discover that for a difference between q and $p = 0.25$ is greater than 0.1, the NPI investor is expected to gain some profit by investing in either a European call or put option with maturity identical to 2. As n increases, the absolute value of the expected profit narrows down like what displays in Figures 2.9 and 2.10. The reason is when n increases, meaning the NPI investor has more information from the market, the interval between the maximum buying price and the minimum selling price gets smaller, approaching to the fair market price, this gives fewer opportunities for the NPI investor to beat the market. Therefore, along with increasing n, the NPI investor will gain less profit from a two time-step European option with a large difference between p and q, while losing less money with a small difference between p and q. If we increase the maturity to $m = 4$ time steps, with fixed n, the trend of results referring to different levels of differences between p and q is identical to the one when $m = 2$. But for the option with $m = 4$, based on the same fixed corresponding historical data, the NPI investor will face more loss when the p and q are close to each other comparing to that when $m = 2$. Thus if there exists more data information given a fixed option maturity, the interval of the difference between q and p leading to a negative profit for the NPI investor are smaller than those based on less data information. We have investigated further cases, including other values of p, for the problem of the expected profit according to differences between p and q is similar as discussed above.

$p = 0.25$	Expected Profit											
	$m = 2$						$m = 4$					
	$n = 40$		$n = 100$		$n = 200$		$n = 40$		$n = 100$		$n = 200$	
q	call	put	call	put	call	put	call	put	call	put	call	put
0.15	-0.22	0.00	-0.21	-0.10	-0.20	0.01	-0.23	0.00	-0.19	0.03	-0.18	0.02
0.25	-0.17	-0.24	-0.11	-0.17	-0.08	-0.13	-0.19	-0.46	-0.12	-0.34	-0.09	-0.25
0.35	0.13	-0.10	0.21	0.01	0.21	0.02	0.05	-0.18	0.18	0.02	0.19	0.03
0.45	0.25	0.07	0.22	0.04	0.21	0.02	0.25	0.14	0.21	0.07	0.19	0.04
0.55	0.26	0.09	0.22	0.04	0.21	0.02	0.28	0.18	0.22	0.07	0.19	0.04
0.65	0.26	0.09	0.22	0.04	0.21	0.02	0.28	0.18	0.22	0.07	0.19	0.04
0.75	0.26	0.09	0.22	0.04	0.21	0.02	0.28	0.18	0.22	0.07	0.19	0.04

Table 2.2: Expected profit and loss changing with p and q difference ($p = 0.25$)

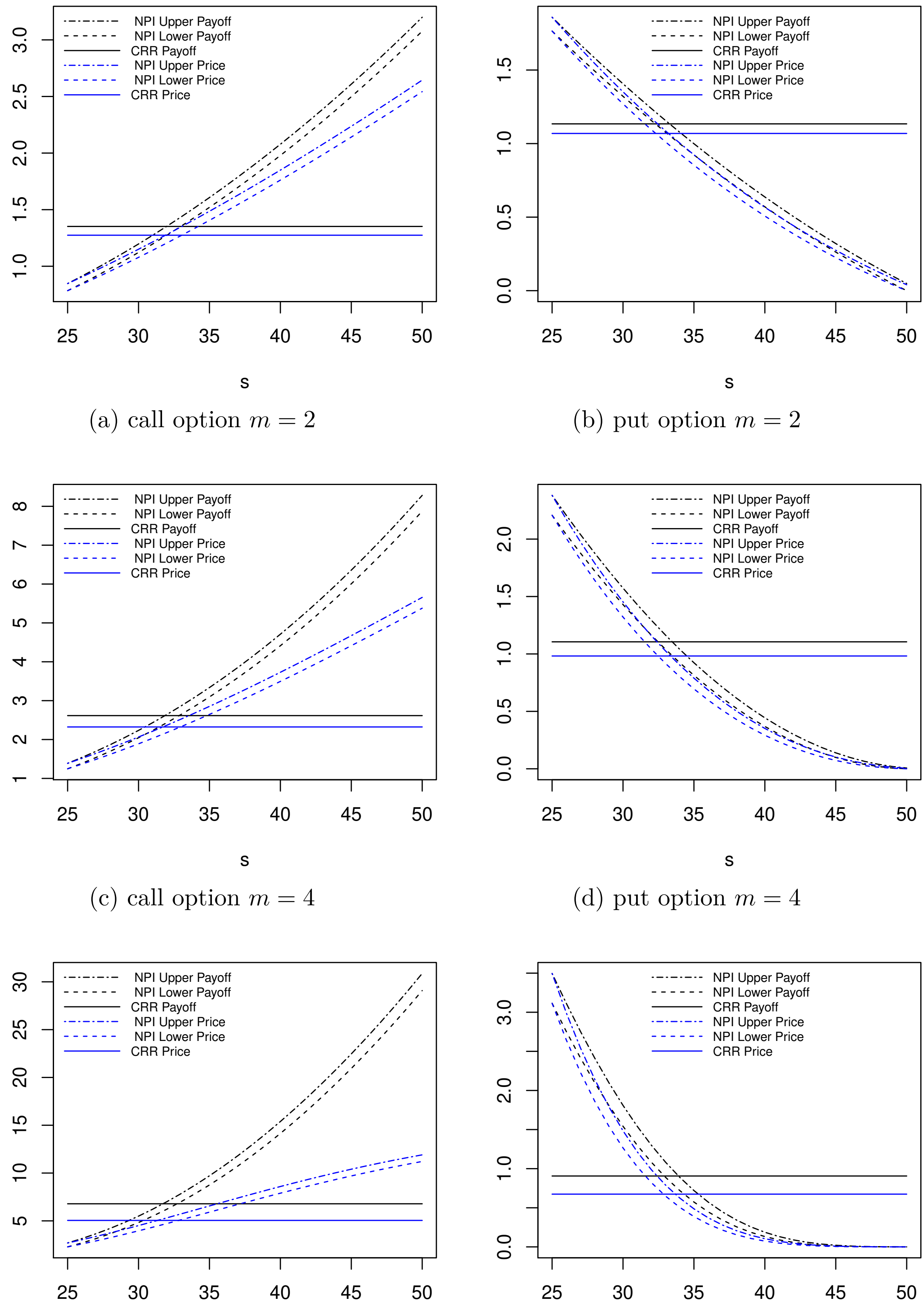
From the examples above, it is evident that the NPI method performs better than the CRR model when the CRR model is under the wrong assumptions. When the CRR model is right about the market, the prediction of the NPI method is different from that of the CRR model, but this difference can be reduced by enlarging the historical data size. A massive difference between the real market probability and the risk-neutral probability also improves the NPI performance.

2.3.3 Performance study including discount rate

To be more comprehensive, we now included the discount rate in the performance study. As we mentioned, in our method we assume the discount rate is equal to the non-negative expected return of the underlying asset, as the expected return of the underlying asset is changing along with the time step, we assume the discount rate as a constant value is the expected return of the underlying asset from the initial time to time 1. The discount rate cannot be negative because if the discount rate is negative, then the future payment is less than the initial amount. Nobody would like to get involved in a sure loss trade. Although the expected return of the underlying asset is lower than the actual discount rate, an option suppose to have, for the analytical study it is the best choice of approximation of the discount rate.

All other assumptions in this study are the same as in the one without the discount rate. We can still use the same formulae to compute the profit and loss. But now the NPI investor will quote at the price calculated from Equations (2.14) and (2.15) for call options and Equations (2.16) and (2.17) for put options, and get or pay the maturity payoff calculated from Equations (2.10) and (2.12) for call options and Equations (2.11) and (2.13) for put options. As we would like to value the profit and loss at maturity, we need to calibrate the profit earned by the option seller. The option seller gets option price V as the payment from the option buyer at the initial time, and after holding it to the maturity the value amount of money is equal to $V(1+r_f)^T$. There are also scenarios in this study, both of them based on a stock with the initial price $S_0 = 20$, the upward movement factor $u = 1.1$ and the downward movement factor $d = 0.9$. The option based on this underlying asset has a maturity $m = 4$. Two investors predict the options based on two methods, in which the CRR investor uses the CRR model to make the prediction with a constant probability q, while the NPI investor uses the NPI method to predict the options based on $n = 50$ historical stock price among them s increased. Then we can calculated the discount rate for the NPI method, $r = \frac{s}{n}u + (1 - \frac{s}{n})d - 1$ depending on s, while the CRR discount rate is $r_f = qu + (1-q)d - 1 = 0.03$. Since we assume that the discount rate is a non-negative value, which leads to no trading action of $s < 25$. We study the prices from the NPI method as well as the CRR model.

The expected price pattern of the option is similar to the expected payoff pattern when the option life period is shorter, e.g. options with the maturity $m = 2$ in Figures 2.11 (a) and (b). However, when it comes to long period options, e.g. options



(a) call option $m = 2$

(b) put option $m = 2$

(c) call option $m = 4$

(d) put option $m = 4$

(e) call option $m = 10$

(f) put option $m = 10$

Figure 2.11: Expected prices and payoffs from the NPI method for European options and the CRR model in Scenario 1

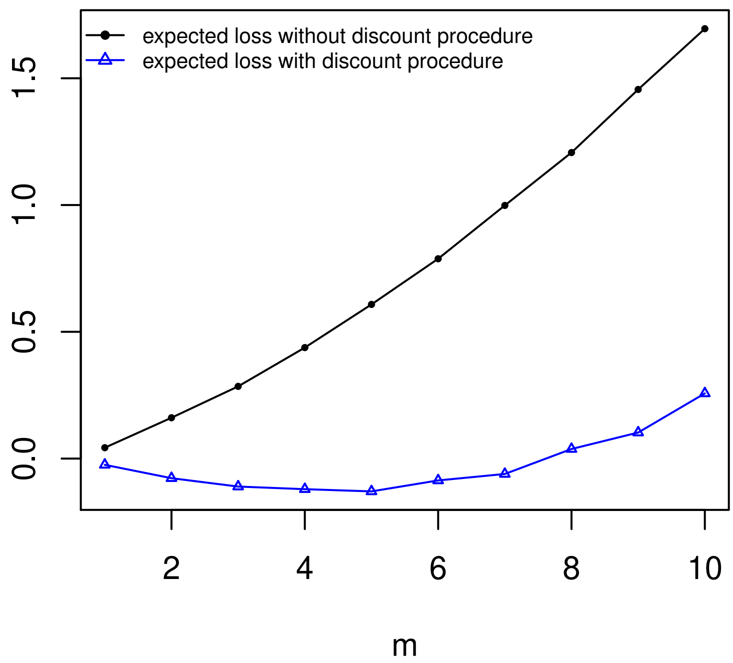
with the maturity $m = 10$ in Figures 2.11 (e) and (f), the difference between expected prices and expected payoffs is obvious, especially for the one with a great payoff. Ultimately, the expected price is from the expected payoff divided by the discount rate. Now we know the influence of discount procedure towards the expected price, then we would like to see how the discount procedure effects the NPI profit and loss (P&L) in two different scenarios. In the first scenario $p = q = 0.65$, then $s \sim Bin(50, 0.65)$. As the stock is expected to rise in the future, then the call option is exercised at maturity while the put option not. Then based on the P&L formulas with the discount procedure: Equations (2.27) and (2.30) for the call option, and Equations (2.34) and (2.37) for the put option. We get the expected NPI loss for the call option, which is,

$$
\begin{aligned}
&E_c[L(s)] \\
&= \sum_{s=0}^{s_1} L(n, m, s : s < s_1 | V = \overline{V_c}) \binom{n}{s} q^s (1-q)^{n-s} \\
&+ \sum_{s=s_2}^{n} L(n, m, s : s > s_2 | V = \underline{V_c}) \binom{n}{s} q^s (1-q)^{n-s} \\
&= \sum_{s=0}^{s_1} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m}{k} \binom{n}{s} \\
&\times \left[q^{k+s} (1-q)^{n+m-s-k} - (1+r)^{-m} \binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{n-s}{n-s+m-k} q^s (1-q)^{n-s} \right] \\
&+ \sum_{s=s_2}^{n} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m}{k} \binom{n}{s} \\
&\times \left[(1+r)^{-m} \binom{n+m}{s+k}^{-1} \binom{n}{s} q^s (1-q)^{n-s} \frac{s}{s+k} - q^{k+s} (1-q)^{n+m-s-k} \right]
\end{aligned}
\tag{2.55}
$$

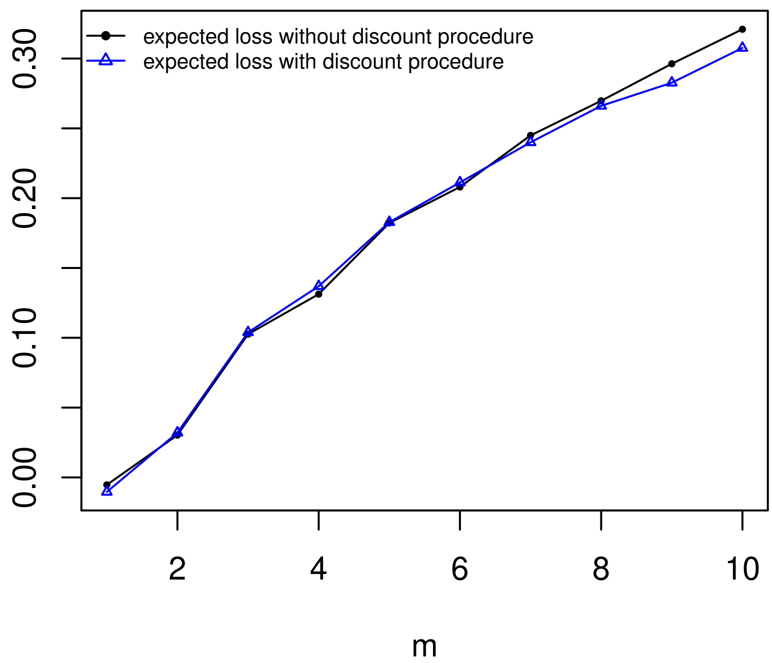
For the put option, the expected NPI loss is,

$$
\begin{aligned}
&E_p[L(s)] \\
&= \sum_{s=0}^{s_3} L(n,m,s : s \leq s_3 | V = \underline{V_p}) \binom{n}{s} q^s (1-q)^{n-s} \\
&+ \sum_{s=s_4}^{n} L(n,m,s : s \geq s_4 | V = \overline{V_p}) \binom{n}{s} q^s (1-q)^{n-s} \\
&= \sum_{s=0}^{s_3} \sum_{k=0}^{\lfloor k_p^* \rfloor} (1+r)^{-m} [K_p - u^k d^{m-k} S_0] \binom{m+n}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \\
&\times \binom{n}{s} q^s (1-q)^{n-s} \\
&- \sum_{s=s_4}^{n} \sum_{k=0}^{\lfloor k_p^* \rfloor} (1+r)^{-m} [K_p - u^k d^{m-k} S_0] \binom{n+m}{m}^{-1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \\
&\times \binom{n}{s} q^s (1-q)^{n-s}
\end{aligned} \tag{2.56}
$$

We calculated the expected NPI loss with various maturities, m from 1 to 10, to explain the influence of discount procedure as well as the benefit for further P&L comparison in the second scenario. In Figure 2.12, we listed graphs of expected NPI loss with and without the discount procedure for both call and put options. For the call option, we can see the influence of involving discount procedure is significant. This is because the call option in this example is exercised, then the expected NPI loss consists of in both the payoff at maturity and the option price. Including the discount procedure to some extent weakens the option price part. Since $s \sim Bin(50, 0.65)$ and when $s \geq 33$, the NPI investor is more likely to buy the call option, paying the option price and earning the payoff. So apart from the option price, the NPI investor is expected to have a positive profit at maturity. Therefore, the disadvantage of NPI prediction in this scenario is insignificant. Whereas the put option is not exercised at maturity, then all expected NPI loss contains is the put option price. Therefore, the difference with or without discount procedure is not apparent.



(a) Expected loss for the call option



(b) Expected loss for the put option

Figure 2.12: Expected NPI loss in Scenario 1

In the second scenario, the real market probability $p = 0.85$ is different from the CRR probability assumption $q = 0.65$, while the NPI method does the prediction based on the historical data $s \sim Bin(50, 0.85)$. The underlying asset is still expected to have a higher value in the future, making for the call option exercise but not for the put option. To calculate the expected NPI profit, we need to use the discount version of Equations (2.41) and (2.44) for the call option and Equations (2.48) and (2.51) for the put option. The specific formula of the expected NPI profit for the call option is displayed below.

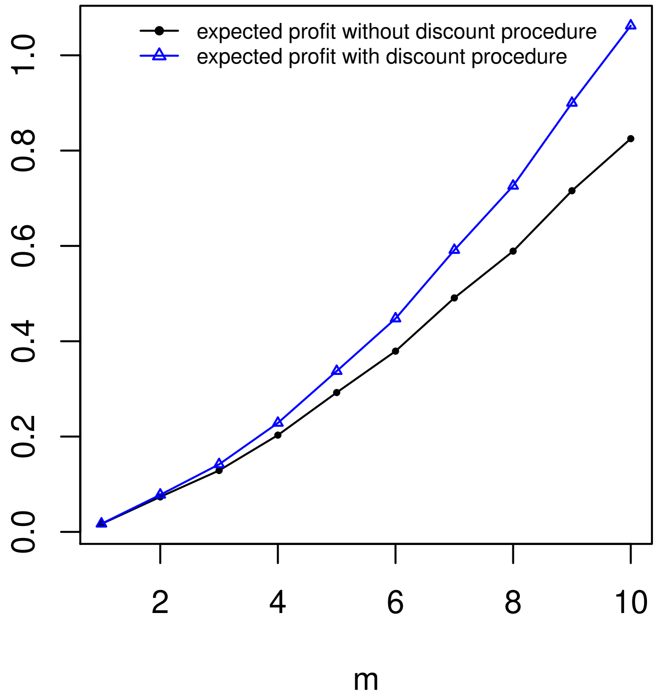
$$
\begin{aligned}
&E_c[Pro(p, s)] \\
&= \sum_{s=0}^{s_5} Pro(n, m, s : s \leq s_5 | V = \overline{V_c}) \binom{n}{s} p^s (1-p)^{n-s} \\
&\quad + \sum_{s=s_6}^{n} Pro(n, m, s : s \geq s_6 | V = \underline{V_c}) \binom{n}{s} p^s (1-p)^{n-s} \\
&= \sum_{s=0}^{s_5} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m}{k} \binom{n}{s} \\
&\times \left[(1+r)^{-m} p^s (1-p)^{n-s} \binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{n-s}{n-s+m-k} - p^{s+k} (1-p)^{n-s+m-k} \right] \\
&\quad + \sum_{s=s_6}^{n} \sum_{k=\lceil k_c^* \rceil}^{m} [u^k d^{m-k} S_0 - K_c] \binom{m}{k} \binom{n}{s} \\
&\times \left[p^{s+k} (1-p)^{n-s+m-k} - (1+r)^{-m} p^s (1-p)^{n-s} \binom{n+m}{s+k}^{-1} \binom{n}{s} \frac{s}{s+k} \right] \qquad (2.57)
\end{aligned}
$$

For the put option, the formula of the expected NPI profit is,

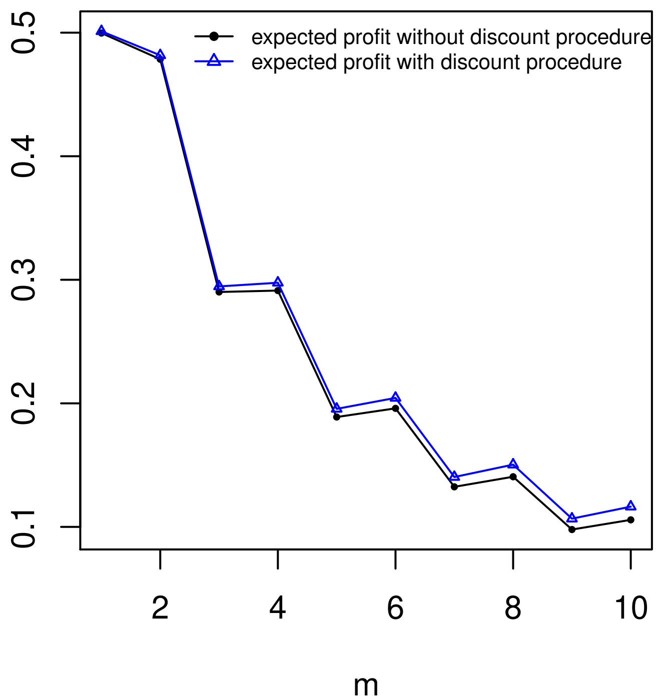
$$
\begin{aligned}
&E_p[Pro(p,s)] \\
&= \sum_{s=0}^{s_7} L(n,m,s:s\leq s_7|V=\underline{V_p})\binom{n}{s}p^s(1-p)^{n-s} \\
&\quad + \sum_{s=s_8}^{n} Pro(n,m,s:s\geq s_8|V=\overline{V_p})\binom{n}{s}p^s(1-p)^{n-s} \\
&= -\sum_{s=0}^{s_7}\sum_{k=0}^{\lfloor k_p^* \rfloor}[K_p-u^kd^{m-k}S_0](1+r)^{-m}\binom{m+n}{m}^{-1}\binom{s+k}{k}\binom{n-s+m-k-1}{m-k} \\
&\times \binom{n}{s}p^s(1-p)^{n-s} \\
&+ \sum_{s=s_8}^{n}\sum_{k=0}^{\lfloor k_p^* \rfloor}[K_p-u^kd^{m-k}S_0](1+r)^{-m}\binom{m+n}{m}^{-1}\binom{s+k-1}{k}\binom{n-s+m-k}{m-k} \\
&\times \binom{n}{s}p^s(1-p)^{n-s}
\end{aligned}
\tag{2.58}
$$

Then we plotted Figure 2.13, the expected NPI profit with and without the discount procedure. From the plot, we can see that the discount procedure does not make a lot of difference to the expected NPI profit. For both the call and put options, the expected NPI profit with the discount procedure has the similar values and shape to the one without the discount procedure, indicating the NPI investor can earn some money from this option trade with the CRR investor. The reason that when m is large, e.g. $m = 10$, the expected NPI profit with the discount procedure is higher than that without the discount procedure is that in the discount procedure we use the risk-free rate to do the calibration of NPI earning as an option seller. In this case, the risk-free rate is calculated from p, which is the real market probability. Although the NPI method is based on the historical information $s \sim Bin(50, 0.85)$, it cannot entirely follow the market pattern of the market. But after the calibration, the result is more robust than without the discount procedure leading to a higher profit.

From the comparison, it is clear that the discount procedure does influence the



(a) Expected profit for the call option



(b) Expected profit for the put option

Figure 2.13: Expected NPI profit in Scenario 2

price prediction. A proper discount procedure can weaken the disadvantage of the inaccurate prediction from the NPI method when the CRR is right. If we compare the expected NPI P&L with the discount procedure in Scenario 2 to the results in Scenario 1, we can see that our conclusion revealed in the performance without the discount procedure still valid: the NPI method performs better if there are some wrong assumptions in the CRR model.

2.4 Concluding remarks

The NPI performances comparing to the CRR model performances		
Influencing factor	Scenario 1:The CRR model is right	Scenario 2: The CRR model is wrong
	The NPI investor faces amount of loss trading with the CRR investor.	The NPI investor benefits from the trade with the CRR investor.
Increase the number of historical data	The loss of the NPI investor decreases.	The profit of the NPI investor drops.
Enlarge the difference between the risk-neutral probability and the real market probability		The profit of the NPI investor increases.
Include an appropriate discount rate	The loss of the NPI investor decreases.	The profit of the NPI investor increases.

Table 2.3: A summary of the NPI performance results

The NPI method for European option pricing, a way keeping learning from historical data, relaxes some traditional assumptions, one of the most important ones is that we do not assume the probability of upward movement to remain constant. In Section 2.2, we prove that the NPI boundary prices also follow the put-call parity. As the classic put-call parity only valid in an arbitrage-free market, but our method generates an interval price indicating the existence of arbitrage opportunity, the boundary prices put-call parity is reasonable. After setting up the NPI method for European options, we compared our model with the CRR model. In this analytical study, two extreme scenarios were investigated. Scenario 1 is the CRR investor predicts with the same knowledge as that in the market, meaning the CRR price is identical to the market price. In this scenario, the NPI investor is not expected to beat the CRR investor, but with more historical information the NPI investor performs better. Scenario 2 is the opposite of Scenario 1, in which the CRR investor made a mistake during prediction. In this scenario, investing in a corresponding option in the same direction as your prediction is a good move for the NPI investor. As the forecast from the CRR model gets closer to the truth, the advantages of the NPI method dwindles. After settling the expected underlying asset return as the discount rate, we study the NPI method's performance, that the results show that an appropriate discount procedure can release the wrong prediction influence when the CRR model does the perfect job, and do not change the conclusion when the CRR model uses the false assumption. A tabular summary of the NPI performance study is listed in Table 2.3.

Some further topics are still interesting to be investigated. The sufficient size of historical data to support an accurate prediction is not expounded, the necessity to include all the historical data available in the market. Also, at the start point, we only involve two traders for the method performance study. The work based on multiple traders with complex scenarios is also challenging and appealing. Populating the empirical historical data to the NPI method and studying its performance need more hard work. Last but not least, the application of the NPI method to a trinomial model is also attractive, which gives more variability to the underlying asset price movement. And the application of NPI method to a trinomial model is different by the order of the possible outcomes from one node.

Chapter 3

NPI for American Option Pricing

After applying the NPI method to the European option pricing, we want to find out if the NPI method can perform the same as for another type of vanilla options, the American option, as it does for the European option.

American options giving the right of early exercise are an essential type of options in the market. Due to the path dependence feature of the American option, it is difficult to find a closed formula for pricing. The CRR model can be used for American option pricing, and other scholars extent this model to fit more complicated situations [30]. Boyle [15] set up a binomial tree model for an American option based on two underlying state variables making the model handle the early exercise feature of the American option. Amin [2] improved the original CRR model by adding jump diffusion to fit in the path-dependent options' evaluation like the American option, under the assumption of market completeness and the risk-neutral world. To contain more uncertainty, the CRR model has been converted to some new versions. Hu and Cao [43] propose a binomial tree model with randomized stock price movement. Zdenek [77] implemented the fuzzy set theory in the American real option pricing procedure to embrace more uncertainty and study its completeness and the non-arbitrage property. However, these models are still under the original assumptions, like the risk-neutral world and the constant probability of stock price upward movements, and overlook the information from historical data.

In Chapter 2, we presented a novel European option pricing method based on Nonparametric Predictive Inference [40]. Instead of using a precise probability for each step in the binomial tree, we applied NPI and concluded its advantages when the investor has less specific information about the underlying asset. In this chapter, we investigate this imprecise statistical method for American options pricing. We apply the NPI method to the discrete binomial tree model to price American options.

There is no closed form formula for the American option pricing, so a backward optimization method can be used for the American option pricing based on the CRR model, and this method is also applicable for the NPI method for the American option pricing as will be shown in Section 3.1. In Section 3.2, we prove the rational trading theory, 'Never early exercise an American call option without dividend', is not valid according to our method. Then we compare the performance of the NPI method to the performance of the CRR model in two extreme scenarios using the same setting as that in Chapter 2. In Section 3.4, we conclude the results and offer some future study potentials.

3.1 NPI for American option pricing

Due to the early exercise possibility, there is no closed form formula for American option pricing based on the binomial tree model. In Section 1.1, we define American option pricing from the best exercise time aspect. Here we give a different but equivalent definition representing the idea of the backward pricing strategy. Let $h_t(x)$ denote the instant value of the American option at time t, $0 \leq t \leq T$, given $S_t = x$. Then $h_t(x) = x - K_c$ for a call option and $h_t(x) = K_p - x$ for a put option. $V_t(x)$ is the option value at time $t \in \{0, 1, \ldots, T-1\}$ given $S_t = x$. The American

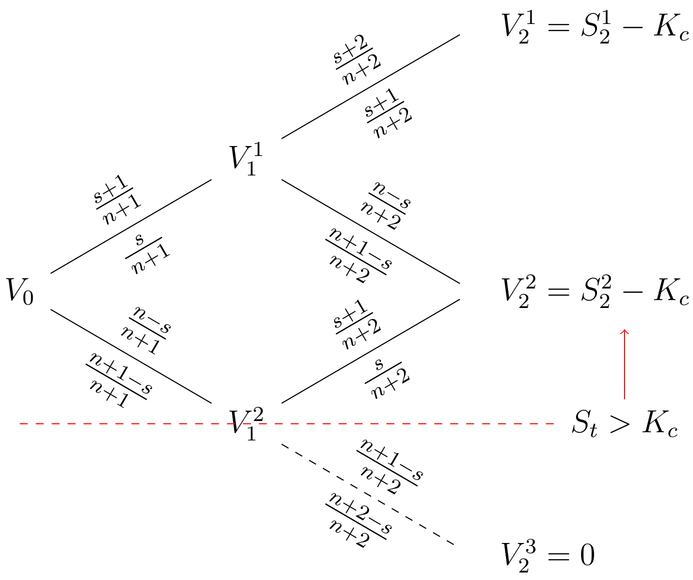


Figure 3.1: The binomial tree based on the NPI method for an American call option

option value at time t is

$$V_t(x) = \max\{h_t(x), B(t, t+1)V_{t+1}(S_{t+1}|S_t = x)\} \tag{3.1}$$

$$V_T(x) = \max\{h_T(x), 0\} \tag{3.2}$$

Here $B(t, t+1)$ is the discount factor between times t and $t+1$. By recursion, we obtain the value of the American option at the initial time, $V_0(x)$, which is the predicted price of this American option.

3.1.1 American call option

Figure 3.1 displays the backward optimization pricing procedure. Suppose there are n historical stock prices available and among them s increased. The call option value in node i at time t is V_t^i. From the tree we can tell than there are $T+1$ levels from level 0 to level T and in each level the number of nodes is the level number plus one, thus $t \in \{0, \ldots, T\}$ and $i \in \{1, \ldots, t+1\}$. We start to evaluate the call option at maturity $V_T^i = \max\{0, S_T^i - K_c\}$ with $i \in \{1, \ldots, T+1\}$ where S_T^i the stock price in case i at maturity T. Rolling back to evaluate this call option for each node i from time $T-1$ to 0 on the basis of the definition $\overline{V_t^i} = \max\left\{S_t^i - K_c, (1+r)^{-1}[\overline{P_t^i}\ \overline{V_{t+1}^i} + (1-\overline{P_t^i})\overline{V_{t+1}^{i+1}}]\right\}$ for the upper value and $\underline{V_t^i} = \max\left\{S_t^i - K_c, (1+r)^{-1}[\underline{P_t^i}\ \underline{V_{t+1}^i} + (1-\underline{P_t^i})\underline{V_{t+1}^{i+1}}]\right\}$ for the lower value, where $\overline{P_t^i}$ is the upper probability for the node i at time t derived from Equation (1.17), and $\underline{P_t^i}$ is the lower probability for the node i at time t from Equation (1.16). S_t^i is the underlying asset for node i at time t, and r is the discount rate. Based on the general formula for the American option pricing, Equations (3.1) and (3.2), the formulae for each node in the binomial tree based on the backward NPI pricing method for an American call option, are formulated below.

The maximum buying price of an American call option

$$\begin{aligned}
\underline{V_t^i}_{\{i=1\ldots t+1\}} &= \max\left\{[S_t^i - K_c]^+, (1+r)^{-1}\left[\underline{P_t^i}\ \underline{V_{t+1}^i} + (1-\underline{P_t^i})\underline{V_{t+1}^{i+1}})\right]\right\} \\
&= \max\left\{[S_t^i - K_c]^+, (1+r)^{-1}\left[\frac{s+t-i+1}{n+t+1}\underline{V_{t+1}^i} + \frac{n-s+i}{n+t+1}\underline{V_{t+1}^{i+1}}\right]\right\} \\
\underline{V_T^i}_{\{i=1\ldots T+1\}} &= \max\{0, S_T^i - K_c\}
\end{aligned} \tag{3.3}$$

The minimum selling price of an American call option

$$\begin{aligned}
\overline{V_t^i}_{\{i=1\ldots t+1\}} &= \max\left\{[S_t^i - K_c]^+, (1+r)^{-1}\left[\overline{P_t^i V_{t+1}^i} + (1-\overline{P_t^i})\overline{V_{t+1}^{i+1}}\right]\right\} \\
&= \max\left\{[S_t^i - K_c]^+, (1+r)^{-1}\left[\frac{s+t-i+2}{n+t+1}\overline{V_{t+1}^i} + \frac{n-s+i-1}{n+t+1}\overline{V_{t+1}^{i+1}}\right]\right\} \\
\overline{V_T^i}_{\{i=1\ldots T+1\}} &= \max\{0, S_T^i - K_c\}
\end{aligned} \tag{3.4}$$

Similarly, we can mathematically describe the backward optimization method based on the NPI method for the American put option as well.

3.1.2 American put option

The binomial tree for the American put option is displayed in Figure 3.2. V_t^i with $t \in \{0, \ldots, T\}$ and $i \in \{1, \ldots, t+1\}$ is the put option value in case i at time t. Similar to the call option pricing procedure, we start to evaluate the put option at maturity $V_T^i = \max\{0, K_p - S_T^i\}$ with $i \in \{1, \ldots, T+1\}$ and S_T^i the stock price in

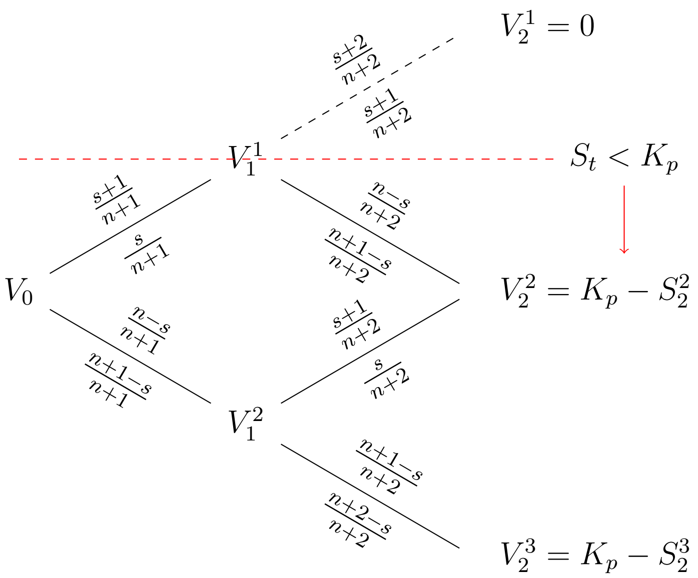


Figure 3.2: The binomial tree of the American put option (in the money)

case i at maturity T. To evaluate this put option for each case i from time $T-1$ to 0, we use formulas $\overline{V_t^i} = \max\left\{K_p - S_t^i, (1+r)^{-1}[\underline{P_t^i}\overline{V_{t+1}^i} + (1-\underline{P_t^i})\overline{V_{t+1}^{i+1}}]\right\}$ for the upper value and $\underline{V_t^i} = \max\left\{K_p - S_t^i, (1+r)^{-1}[\overline{P_t^i}\underline{V_{t+1}^i} + (1-\overline{P_t^i})\underline{V_{t+1}^{i+1}}]\right\}$ for the lower value, which are derived from Equations (3.1) and (3.2). For upper probabilities, we have the formula as Equation (1.17) $\overline{P_t^i} = \frac{s+t-i+2}{n+t+1}$, and for lower probabilities, we have the formula as Equation (1.16) $\underline{P_t^i} = \frac{s+t-i+1}{n+t+1}$. This leads to the following results.

The maximum buying price of an American put option

$$\begin{aligned}
\underline{V_t^i}_{\{i=1\ldots t+1\}} &= \max\left\{[K_p - S_t^i]^+, (1+r)^{-1}\left[\overline{P_t^i}\underline{V_{t+1}^i} + (1-\overline{P_t^i})\underline{V_{t+1}^{i+1}}\right]\right\} \\
&= \max\left\{[K_p - S_t^i]^+, (1+r)^{-1}\left[\frac{s+t-i+2}{n+t+1}\underline{V_{t+1}^i} + \frac{n-s+i-1}{n+t+1}\underline{V_{t+1}^{i+1}}\right]\right\} \\
\underline{V_T^i}_{\{i=1\ldots T+1\}} &= \max\{0, K_p - S_T^i\}
\end{aligned} \tag{3.5}$$

The minimum selling price of an American put option

$$\begin{aligned}
\overline{V_t^i}_{\{i=1\ldots t+1\}} &= \max\left\{[K_p - S_t^i]^+, (1+r)^{-1}\left[\underline{P_t^i}\overline{V_{t+1}^i} + (1-\underline{P_t^i})\overline{V_{t+1}^{i+1}}\right]\right\} \\
&= \max\left\{[K_p - S_t^i]^+, (1+r)^{-1}\left[\frac{s+t-i+1}{n+t+1}\overline{V_{t+1}^i} + \frac{n-s+i}{n+t+1}\overline{V_{t+1}^{i+1}}\right]\right\} \\
\overline{V_T^i}_{\{i=1\ldots T+1\}} &= \max\{0, K_p - S_T^i\}
\end{aligned} \tag{3.6}$$

We have implemented this backward method for American options in the statistics software R in Appendix B.1. The program user inputs to specify the option, that is stock price S_0, upward movement factor u, downward movement factor d, discount rate r, time steps between initial time and the maturity, strike price K, number of historical data n, the number of upward movements in history s. Then the program will ask the type and the trading position of this option, after that a figure like Figure 3.3 will be generated. At each node, there are two values with three digits after the decimal (values with fewer decimal digits are exact results after programming), one outside the parenthesis is the stock price S_t and the one in the parenthesis is the option value $V_t(S_t)$. The result in the parenthesis at the initial time is the price of this option, and the nodes in oval are the case supposed to be exercised early.

3.2 Early exercise of an American option

Merton[57] showed that for an American call option without dividends, the option stopping time is its expiry time, meaning that it is not optimal to exercise an American call option early. This section will discuss the reason for this phenomenon and check if it holds for the NPI method.

In the binomial tree, we used the backward optimization method to calculate the American option price. At each node, instant value $h_t(S_t = S)$ is compared to the discounted holding value $H_t(S_t = S)$, and the greater value is taken as the

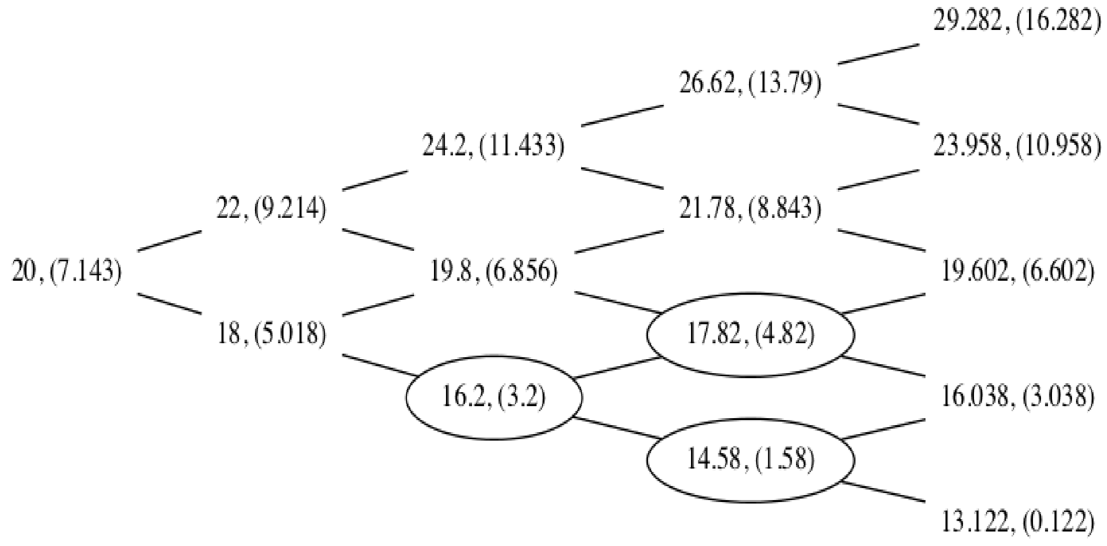


Figure 3.3: Example for option pricing in R

value of this node at time t, $V_t(S_t = S) = \max\{h_t(S_t = S), H_t(S_t = S)\}$. The holding value at time t is equal to the discounted expected value at time $t+1$, $H_t(S_t = S) = B(t, t+1)V_{t+1}(S_{t+1}|S_t = S)$. To start our study of the NPI method, some examples are presented to help us understand the pricing method heuristically.

3.2.1 Examples

Example 3.2.1

According to the different moneyness of the option, options can be categorized in three situations as discussed in Section 1.1, in the money, at the money, and out of the money. In the money options are ones have a positive payoff at the initial time, for call option the strike price lower than the initial stock price ($K_c < S_0$), for put options the strike option is higher than the initial stock price ($K_p > S_0$). At the money means that the decided strike price equals to the initial stock price ($S_0 = K_c$). Out of the money call options have a higher strike price than the initial stock price ($K_c > S_0$), while out of the money put options have a lower strike price than the initial stock price ($K_p < S_0$). In this example, there is an American call option with maturity $T = 2$, which is at the money.

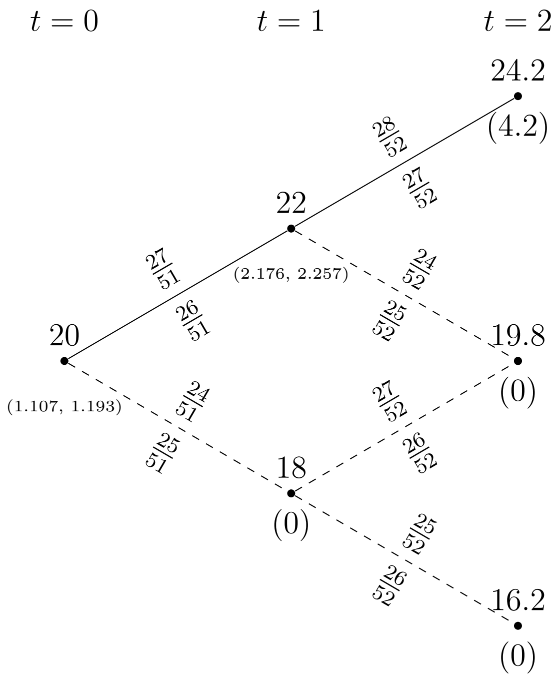


Figure 3.4: The binomial tree of Example 3.2.1

The binomial tree of the stock price and option value at each node $V_t(S_t = S)$ (in the parenthesis, maximum buying value on the left and minimum selling value on the right) are listed in Figure 3.4. A stock with initial price $S_0 = 20$ moves up by factor $u = 1.1$ or down by factor $d = 0.9$. There are 50 historical data available in the market, and among them, 26 are up, then the upper and lower probabilities of each movement are calculated based on Equations (1.17) and (1.16) and displayed in Figure 3.4. According to the criterion of discount rate settlement, the discount rate in our method is set as the expected return of the stock price. On the basis that our method has an interval of expected values, there exist an interval of expected returns. Furthermore, because of the variability of probability at each time step, the expected return interval at each time step varies. We assume in this example that the discount rate r is equal to the lower expected return of the stock price during the period from time 0 to time 1. Thus, examples in this section are under the assumption that the investor has a lower expectation for the stock price. Accordingly, the discount rate is equal to $r = u\frac{s}{n+1} + d\frac{n-s+1}{n+1} - 1 \approx 0.002$.

Clearly, in this example, only the top path with all upwards movements has a

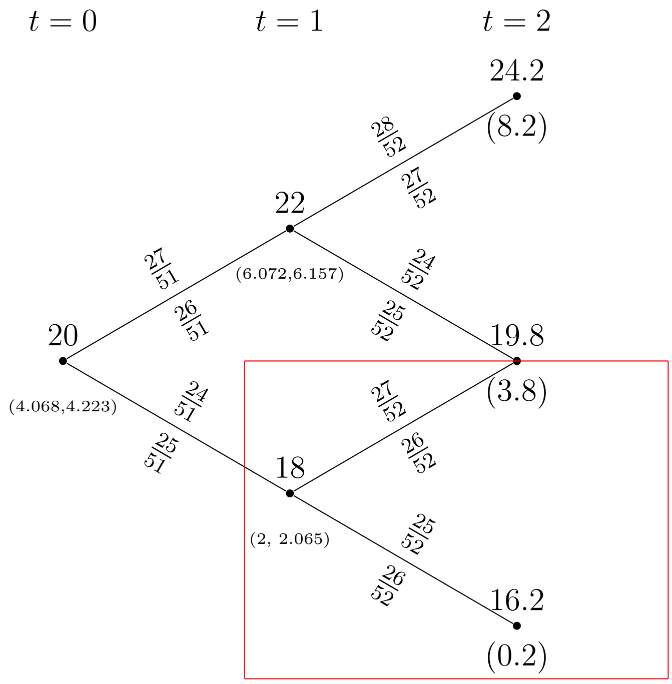


Figure 3.5: The binomial tree of Example 3.2.2

positive payoff. Then for this call option referring to the NPI method, it should not be exercised early, for at each node $i \in \{1, \ldots, t+1\}$ of each time step $t \in \{0, \ldots, 2\}$ the instant value $h_t(S_t^i)$ is lower than the holding value $H_t(S_t^i)$. There is only one path, $20 \to 22 \to 24.2$, in the tree is included in the pricing procedure that is shown as the solid line with positive values in the parenthesis at each node of the path. So the maximum buying price and minimum selling price are $\underline{V_0} = (1+0.002)^{-2}(24.2-20)\frac{27}{52}\frac{26}{51} \approx 1.107$ and $\overline{V_0} = (1+0.002)^{-2}(24.2-20)\frac{28}{52}\frac{27}{51} \approx 1.193$, respectively, which are identical to these prices for the corresponding European call option based on the NPI method.

Example 3.2.2

We aim to price an American call option based on the same underlying asset as that in Example 3.2.1 as well as the same historical data, but now we consider strike price $K_c = 16$. As a result of the lower strike price, this American call option is in the money at the initial time, and all paths are taken into account for the pricing procedure. Based on the definition of the American call option, we can check whether there is any possibility of prematurely exercise this call option. It turns out that the lowest branch of the binomial tree from time 1 to time 2, highlighted in the red square in Figure 3.5, can be exercised early. Since the lower discounted expected payoff at time 2, which is equal to $(1+0.002)^{-1}(3.8\frac{26}{52}+0.2\frac{26}{52}) \approx 1.996$, is less than 2 the instant payoff at time 1, while the upper discounted expected payoff at time 2, which is equal to $(1+0.002)^{-1}(3.8\frac{27}{52}+0.2\frac{25}{52}) \approx 2.065$, is greater than 2, the instant payoff at time 1. The NPI investor, as the option holder, prefers to exercise this option early but does not expect the option buyer to exercise this option before its maturity as the option writer, for the instant payoff is lower than the holding value at this time. Therefore for the same option, the NPI investor is willing to sell at a price as a European option but willing to buy it at a higher price as an American option.

Example 3.2.3

In this example, the American call option is still based on the same underlying asset as in Examples 3.2.1 and 3.2.2, but with a longer maturity $T = 4$ and lower strike price $K_c = 13$. At time $t = 4$, when it comes to the last branch in the tree, both upper and lower discounted expected values, 1.577 and 1.523, are lower than the instant value 1.58 at time $t = 3$ shown in the red square. Therefore, in this case, if the NPI investor is the option holder, it is optimal to exercise this American call option early, and if the NPI investor is the option writer, the buyer is expected to exercise the option early. So the option is sold at a higher price.

Similarly, this method can also confirm that for a non-dividend American call option it is possible to gain more profit when it is exercised prematurely than exercising at maturity.

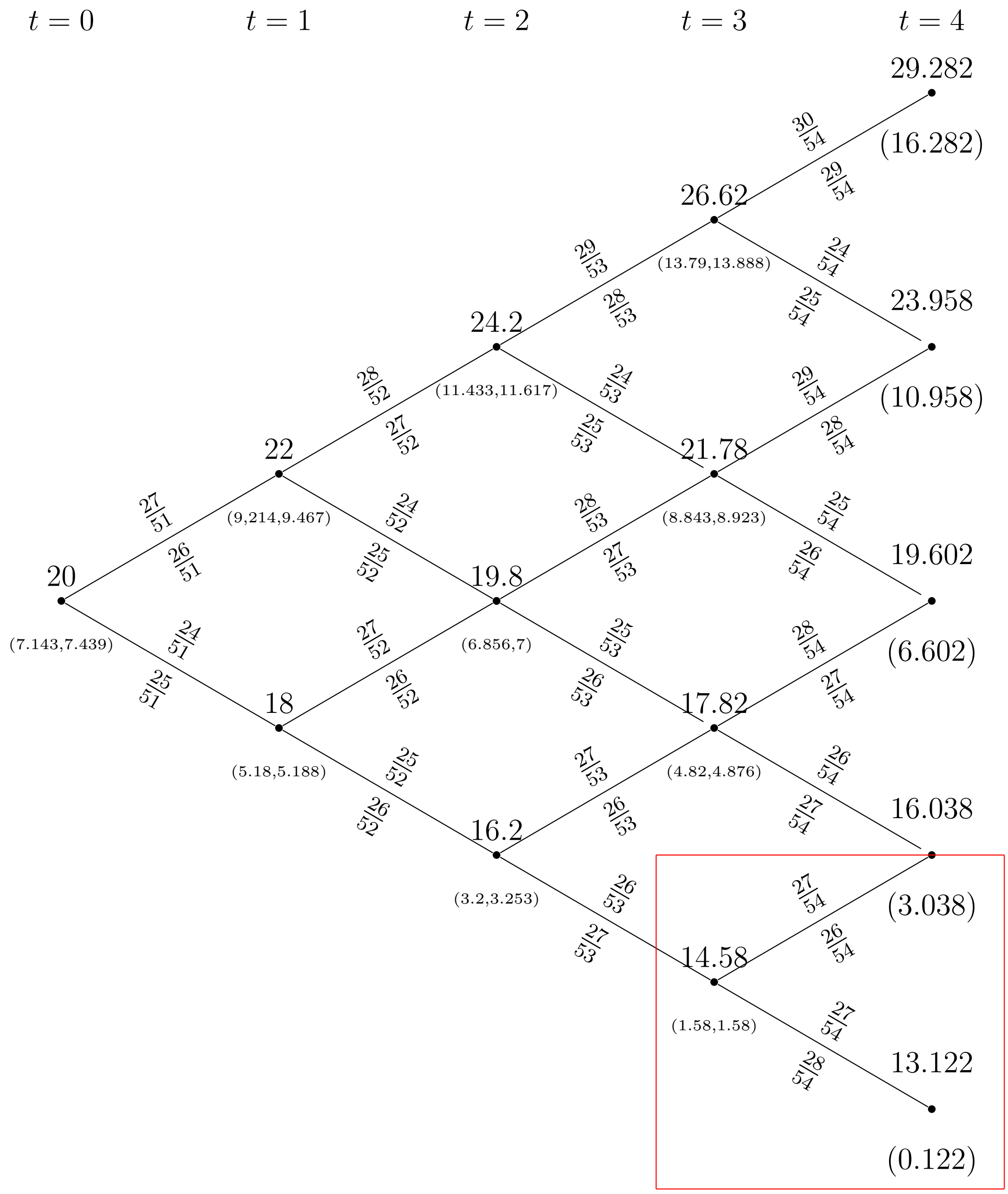


Figure 3.6: The binomial tree of Example 3.2.3

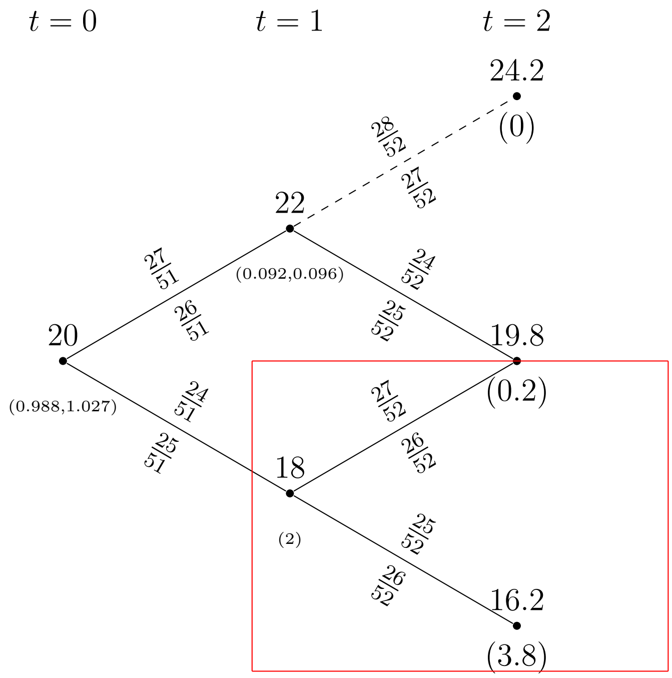


Figure 3.7: The binomial tree of Example 3.2.4

Example 3.2.4

Based on the definition of the American put option Equation (1.4), this at the money put option based on the same underlying asset as the other examples will be exercised prematurely in the case of the last branch at time $t = 2$. For the put option, based on the definition of put option payoff $[K_p - S_t]^+$, the lower stock price is, the higher instant payoff is. But due to the discount procedure, the holding value $H(S_t)$ can be lower than the instant value that is easier to be encountered than a call option. And in this example for both the buying and selling positions the early exercise is expected to happen, then as an option writer the NPI investor would be assigned to this put option early exercise payoff before its maturity so that this option would be sold at a higher price.

3.2.2 Early exercise of an American call option

After the example study, we know that an American call option can be exercised early from the NPI perspective. As for the NPI method, there are two bounds, upper and lower boundaries. An investor's trading position will decide which boundary supposed to be concerned. As an option holder, the investor should focus on the maximum buying price, while the option seller is supposed to pay attention to the minimum selling price. Then to study the condition of early exercise, we can simplify the procedure by only focusing on either the maximum buying price or the minimum selling price according to two trade positions. As we discussed in Section 1.3, the discount rate in our method can be the expected return of the stock or the expected return of the option if the market is completed. By assuming the discount rate is a constant value r, we can set the discount factor at the beginning of the call option. Firstly, we discuss to exercise an American call option prematurely, then the NPI investor is supposed to hold an American call option and willing to buy the underlying asset in the future if there exist some profits. Therefore, only the lower NPI expected value, lower expected stock price, lower expected option values, and lower expected return need to be considered. If we set the discount rate at time t equal to r, then at time $t+1$ there will be two different circumstances of stock lower expected return $\underline{r_{t+1}}$, higher than r or lower than r, where $\underline{r_{t+1}} = u\underline{P_t}(S_t) + d(1 - \underline{P_t}(S_t)) - 1$ and $\underline{E}(S_{t+1})(1 + \underline{r_{t+1}})^{-1} = S_t$, for the lower expectation of the stock at time $t+1$ should be equal to the stock price at time t times its lower expected return during this time period. Therefore, the discounted expected stock price $\underline{E}(S_{t+1})(1+r)^{-1}$ at time $t+1$ is not always equal to the stock buying price S_t at t, while in risk-neutral evaluation we always have $E(S_{t+1})(1 + r_f)^{-1} = S_t$.

Referring to Figure 3.8, there is a stock price S_t^i with $t \in \{0, \dots, T\}$ and $i \in \{1, \dots, t+1\}$ at every node in the binomial tree, and we also have option value of each node V_t^i calculated with backward optimization method following the early exercise condition. To get the early exercise condition for an American call option holder, we compare the option instant value $h_t(S_t = S_t^i)$ to the holding value $H_t(S_t = S_t^i)$ at time t for each node i. If $S_t - K_c > H_t(S_t = S_t^i)$, it is optimal to exercise this call option, otherwise holding it is wiser. Here $H_t(S_t = S_t^i)$ is computed based on

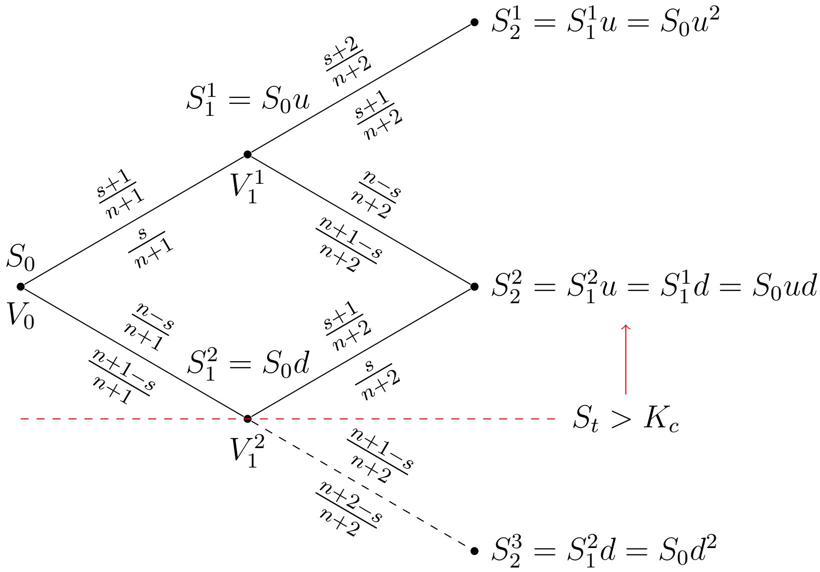


Figure 3.8: The binomial tree of the underlying asset and its American call option instant payoffs based on the NPI method

the option value of two nodes at time $t+1$, $V_{t+1}(S_{t+1} = S_t^i u)$ and $V_{t+1}(S_{t+1} = S_t^i d)$. Therefore, before the comparison we need to consider the exercise condition of two nodes at time $t+1 \neq T$, which consist of three circumstances: two nodes are exercised at time $t+1$, one is exercised while the other is better to be held, and two nodes are held. For the first circumstance, both nodes at time $t+1$ are exercised, let us discuss whether the inequality between the instant value and holding value still holds or not. Here we use the difference between the discounted the payoff of the American call option $B(t,t+1)V_{t+1}(S_{t+1}|S_t) = \underline{E}[S_{t+1} - K_c]^+(1+r)^{-1}$ at t and the instant value of the American call option $h(S_{t+1}) = S_t - K_c$ to do the comparison.

$$
\begin{aligned}
&[\underline{E}[S_{t+1} - K_c]^+](1+r)^{-1} - (S_t - K_c) \\
&\geq [\underline{E}(S_{t+1} - K_c)](1+r)^{-1} - (S_t - K_c) \\
&= [\underline{E}(S_{t+1} - K_c)](1+r)^{-1} - (\underline{E}[S_{t+1}](1+\underline{r_{t+1}})^{-1} - K_c) \\
&= \underline{E}[S_{t+1}]((1+r)^{-1} - (1+\underline{r_{t+1}})^{-1}) - K_c((1+r)^{-1} - 1) \\
&\overset{(b)}{\geq} K_c((1+r)^{-1} - (1+\underline{r_{t+1}})^{-1} - (1+r)^{-1} + 1) \\
&= K_c(1 - (1+\underline{r_{t+1}})^{-1}) > 0
\end{aligned} \tag{3.7}
$$

Here, r is the non-negative discount rate set at the open contract time, and $\underline{r_{t+1}}$ is lower expected return of the stock price at time $t+1$. And for all rates we assume they are positive. This inequality (b) holds as an American call option can be exercised under the condition $S_t - K_c > 0$, which is definitely followed. Since both stock prices in the one-step binomial tree at time $t+1$ follow the conditions, $S_t u - K_c > 0$ and $S_t d - K_c > 0$ with factors u and d movement factors of the stock price, then $S_t - K_c > 0$ is always true. And another condition of (b) that needs to be followed is $(1+r)^{-1} - (1+\underline{r_{t+1}})^{-1} > 0$, meaning $\underline{r_{t+1}} > r$.

Actually $\underline{r_{t+1}} > r$ is not the exact condition for preventing NPI American call option early exercise. We would like to reveal the condition for stopping the NPI American call option early exercise, and the condition is

$$
\begin{aligned}
&[\underline{E}[S_{t+1} - K_c]^+](1+r)^{-1} \geq \underline{E}(S_{t+1} - K_c)](1+r)^{-1} > (S_t - K_c) \\
&\Leftrightarrow \underline{E}[S_{t+1}] - K_c > (S_t - K_c)(1+r) \\
&\Leftrightarrow S_t(1+\underline{r_{t+1}}) - K_c > (S_t - K_c)(1+r) \\
&\Leftrightarrow (1+\underline{r_{t+1}}) > \frac{(1+r)(S_t - K_c) + K_c}{S_t} \\
&\Leftrightarrow \underline{r_{t+1}} > (1 - \frac{K_c}{S_t})r \qquad (3.8) \\
&\Leftrightarrow \underline{P_t}(S_t) > \frac{(1+r-d)S_t - rK_c}{(u-d)S_t} \qquad (3.9)
\end{aligned}
$$

We can express the condition for holding this call option at S_t not only as $\underline{r_{t+1}} > (1 - \frac{K_c}{S_t})r$ but also as a condition on the lower probability at S_t, $\underline{P_t}(S_t) > \frac{(1+r-d)S_t - rK_c}{(u-d)S_t}$. This is derived due to the relationship between the lower expected stock return and the lower probability, $1 + \underline{r_{t+1}} = u\underline{P_t}(S_t) + (1 - \underline{P_t}(S_t))d$.

For the circumstance that the option of one node at $t+1$ is optimal to be held while the other is exercised early, there exist two different situations; the first one is that the upward node is optimal to be held while the downward node is optimal to be exercised early. In this situation, the upward node contains the option value, which is the holding value at time $t+1$ represented as $H_{t+1}(S_t u)$, which is greater than the instant value at time $t+1$, $H_{t+1}(S_tu) > h_{t+1}(S_tu) \Leftrightarrow H_{t+1}(S_tu) > S_tu - K_c$. The downward node value is the instant value $h_t(S_td) = [S_td - K_c]^+$. The condition of this situation now becomes

$$\begin{aligned}
&(1+r)^{-1}[\underline{P_t}(S_t)H_{1+t}(S_tu) + (1-\underline{P_t}(S_t))[S_td - K_c]^+] \\
&\geq (1+r)^{-1}[\underline{P_t}(S_t)(S_tu - K_c + a) + (1-\underline{P_t}(S_t))(S_td - K_c)] > (S_t - K_c) \\
&\Leftrightarrow 1 + \underline{r_{t+1}} > \frac{(1+r)(S_t - K_c) + K_c - \underline{P_t}(S_t)a}{S_t} \\
&\Leftrightarrow \underline{r_{t+1}} > r(1 - \frac{K_c}{S_t}) - \frac{\underline{P_t}(S_t)a}{S_t} && (3.10) \\
&\Leftrightarrow \underline{P_t}(S_t) > \frac{(1+r-d)S_t - rK_c}{(u-d)S_t + a} && (3.11)
\end{aligned}$$

where a is the difference between the discounted expected value and instant value at time $t+1$ for upward node, $a = H_{t+1}(S_tu) - h_{t+1}(S_tu) = H_{t+1}(S_tu) - (S_tu - K_c)$. The value a is a positive value which depends on all future paths related to the node where $S_{t+1} = S_tu$ from time $t+1$ to maturity.

Another possible situation in this circumstance is that the upward node is optimal to be exercised, but the downward node is optimal to be held. In order to find the condition of holding the option at time t, we compare the discounted expected value at time t, $H_t(S_t) = (1+r)^{-1}[\underline{P_t}(S_t)(S_tu - K_c) + (1-\underline{P_t}(S_t))H_{t+1}(S_td)]$, and the instant value at time t, $h(S_t) = S_t - K_c$. This leads to the condition.

$$\begin{aligned}
&(1+r)^{-1}[\underline{P_t}(S_t)(S_tu - K_c) + (1-\underline{P_t}(S_t))H_{t+1}(S_td)] > (S_t - K_c) \\
&\Leftrightarrow (1+r)^{-1}[\underline{P_t}(S_t)(S_tu - K_c) + (1-\underline{P_t}(S_t))(S_td - K_c + b)] > (S_t - K_c) \\
&\Leftrightarrow 1 + \underline{r_{t+1}} > \frac{(1+r)(S_t - K_c) + K_c - (1-\underline{P_t}(S_t))b}{S_t} \\
&\Leftrightarrow \underline{r_{t+1}} > r(1 - \frac{K_c}{S_t}) - \frac{(1-\underline{P_t}(S_t))b}{S_t} && (3.12) \\
&\Leftrightarrow \underline{P_t}(S_t) > \frac{(1+r-d)S_t - rK_c - b}{(u-d)S_t - b} && (3.13)
\end{aligned}$$

where b is a positive value equal to the difference between the discounted expected value from time $t+2$ and the instant value at node $S_{t+1} = S_t d$, $b = H_{t+1}(S_t d) - h_{t+1}(S_t d) = H_{t+1}(S_t d) - (S_t d - K_c)$ with $S_t d - K_c < 0$. As if $S_t d - K_c < H_{t+1}(S_t d)$ then $S_t u - K_c < H_{t+1}(S_t u)$ unless $S_t d - K_c < 0$.

The last circumstance is pretty clear, because both future nodes at time $t+1$ are optimal to be held, surely the node at time t is optimal to be maintained as well. For circumstances investigated above it is supposed to be held that at time t the American call option has a positive instant payoff. Otherwise, the investor has to keep it for future time steps. We formulate this result as a theorem.

Theorem 1

If $\underline{r_{t+1}} > r$, then the American call option should be held.

If an American call option is exercised at time t, then $\underline{r_{t+1}} < r$.

Proof

As $(1 - \frac{K_c}{S_t})r > r(1 - \frac{K_c}{S_t}) - \frac{\underline{P_t}(S_t)a}{S_t}$ and $(1 - \frac{K_c}{S_t})r > r(1 - \frac{K_c}{S_t}) - \frac{(1-\underline{P_t}(S_t))b}{S_t}$, we can conclude that if $\underline{r_{t+1}} > (1 - \frac{K_c}{S_t})r$, the call option should be held at time t. Moreover, the upper boundary of $(1 - \frac{K_c}{S_t})r$ is r, then if $\underline{r_{t+1}} > r$, the American call option should be held. On the contrary, if an American call option is exercised at time t, we know that $\underline{r_{t+1}}$ does not follow the holding condition, at least lower than the upper boundary of the holding condition r. Thus, if an American call option is exercised at time t, then $\underline{r_{t+1}} < r$. Note that these conditions in Theorem 1 are sufficient conditions, but not necessary conditions.

For the American call option selling position, replacing $\underline{r_{t+1}}$ in all conditions of different circumstances, Equations (3.8), (3.10) and (3.12), with $\overline{r_{t+1}}$ leads us to the early exercise conditions.

3.2.3 Early exercise of an American put option

Similarly, the NPI method can also confirm that for a non-dividend American

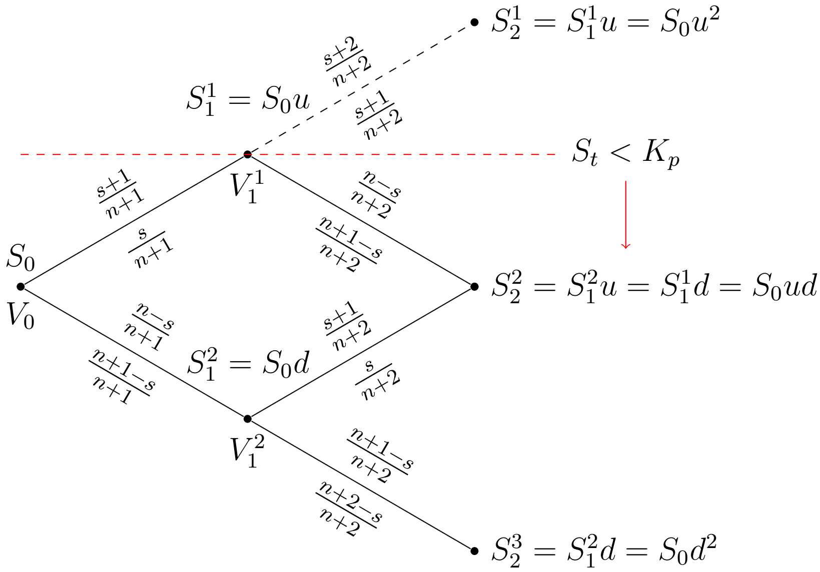


Figure 3.9: The binomial tree of the underlying asset and its American put option instant payoffs based on the NPI method

put option it is possible to gain more profit when it is exercised prematurely than at maturity. Here the NPI investor's position is buying a put option and willing to sell the underlying asset at a higher price in the future. So the lower NPI expected option value of the put option and the minimum selling stock price should be in this comparison. As for the call option, we only focus on the one step binomial tree instead of the whole tree, because of NPI probabilities for each step in the tree change with the data.

The binomial tree of the American put option is drawn in Figure 3.9, with stock price S_t^i, option value V_t^i, $t \in \{0, \ldots, T\}$ and $i \in \{1, \ldots, t+1\}$, at each node and upper and lower probabilities calculated based on Equations (1.17) and (1.16) for every one-step path. For the put option, there are three circumstances for the option of two nodes after stock price movements in the one step tree, the upward node $V_{t+1}(S_{t+1} = S_t^i u)$ and the downward node $V_{t+1}(S_{t+1} = S_t^i d)$: the option for both two nodes is optimal to be exercised early, one is excised prematurely while the other one is better to be held, and the option for both nodes are worth to be held. For the first circumstance, referring to the definition of the put option early exercise, as long as the discounted expected value at time t, $B(t, t+1)V_{t+1}(S_{t+1}|S_t) = \underline{E}[K_p -$

$S_{t+1}]^+(1+r)^{-1}$, is greater than the instant value at time t, $h(S_t) = K_p - S_t$, then it is optimal to be held. Since the option holder is going to sell the stock at the minimum selling price at the exercise time, $\underline{E}(K_p - S_{t+1}) = K_p - \underline{E}(S_{t+1}) = K_p - S_t(1 + \overline{r_{t+1}})$. Here $\overline{r_{t+1}}$ is related to the upper NPI probability, $\overline{r_{t+1}} = \overline{P_t}(S_t)u + (1 - \overline{P_t}(S_t))d - 1$. The condition for holding this option is,

$$\underline{E}[K_p - S_{t+1}]^+(1+r)^{-1} \geq \underline{E}(K_p - S_{t+1})(1+r)^{-1} > K_p - S_t$$

$$\Leftrightarrow K_P - S_t(1 + \overline{r_{t+1}}) > (K_p - S_t)(1+r)$$

$$\Leftrightarrow 1 + \overline{r_{t+1}} < \frac{S_t(1+r) - K_p r}{S_t}$$

$$\Leftrightarrow \overline{r_{t+1}} < (1 - \frac{K_p}{S_t})r \tag{3.14}$$

$$\Leftrightarrow \overline{P_t}(S_t) < \frac{(1+r-d)S_t - rK_p}{(u-d)S_t} \tag{3.15}$$

K_p is the strike price of this American put option. Since $(1 - \frac{K_P}{S_t})r < 0$, unless the stock does has a minimal expected return, in this circumstance the condition for holding a put option until the maturity is harder to achieve than holding a call option until the maturity.

For the second circumstance, for one node the option is exercised at time $t+1$ and for the other note is optimal to hold the option. Similar to the American call option, for this circumstance we have two different situations, the option of the upward node is exercised early, and the downward one is not, or the other way around. For the first situation, then the holding value at time t is $H_t(S_t) = (1+r)^{-1}[\overline{P_t}(S_t)[K_p - S_t u]^+ + (1 - \overline{P_t}(S_t))H_{t+1}(S_t d)]$, where $[K_p - S_t u]^+$ is the instant value at the upward node, and $H_{t+1}(S_t d)$ is the holding value at the downward node.

Then the comparison is,

$$
\begin{aligned}
&(1+r)^{-1}[\overline{P_t}(S_t)[K_p - S_t u]^+ + (1-\overline{P_t}(S_t))H_{t+1}(S_t d)] \\
&> (1+r)^{-1}[\overline{P_t}(S_t)(K_p - S_t u) + (1-\overline{P_t}(S_t))(K_p - S_t d + v)] > K_p - S_t \\
&\Leftrightarrow 1 + \overline{r_{t+1}} < \frac{S_t(1+r) - K_P r + (1-\overline{P_t}(S_t))v}{S_t} \\
&\Leftrightarrow \overline{r_{t+1}} < (1 - \frac{K_P}{S_t})r + \frac{(1-\overline{P_t}(S_t))v}{S_t} \qquad (3.16) \\
&\Leftrightarrow \overline{P_t}(S_t) < \frac{(1+r-d)S_t - rK_p + v}{(u-d)S_t + v} \qquad (3.17)
\end{aligned}
$$

Here $v = H_{t+1}(S_t d) - h_{t+1}(S_t d) = H_{t+1}(S_t d) - (K_p - S_t d)$. If stock price at time t, S_t, is the same as in the first circumstance, then it is clear that this condition is not as strict as that concluded from the first circumstance. Another situation in this circumstance is that the option of the downward node is optimal to be exercised early, and the upward one is optimal to be held. The comparison between the holding value $H_t(S_t) = (1+r)^{-1}[\overline{P_t}(S_t)H_{t+1}(S_t u) + (1-\overline{P_t}(S_t))[K_p - S_t d]^+]$ and the instant value $K_p - S_t$.

$$
\begin{aligned}
&(1+r)^{-1}[\overline{P_t}(S_t)H_{t+1}(S_t u) + (1-\overline{P_t}(S_t))[K_p - S_t d]^+] \\
&> (1+r)^{-1}[\overline{P_t}(S_t)(K_p - S_t u + w) + (1-\overline{P_t}(S_t))(K_p - S_t d)] > K_p - S_t \\
&\Leftrightarrow 1 + \overline{r_{t+1}} < \frac{S_t(1+r) - K_P r + \overline{P_t}(S_t)w}{S_t} \\
&\Leftrightarrow \overline{r_{t+1}} < (1 - \frac{K_P}{S_t})r + \frac{\overline{P_t}(S_t))w}{S_t} \qquad (3.18) \\
&\Leftrightarrow \overline{P_t}(S_t) < \frac{(1+r-d)S_t - rK_p}{(u-d)S_t - w} \qquad (3.19)
\end{aligned}
$$

where w is a constant positive value, which represents the difference between the holding value at $S_t u$ and the negative instant value, $w = H_{t+1}(S_t u) - h_{t+1}(S_t u) = H_{t+1}(S_t u) - (K_p - S_t u)$.

For the last circumstance, when the option for two nodes at time $t+1$ are all optimal to be held, of course at time t we should not do anything towards this option.

All these circumstances are settled based on the fact that the instant value at time t is positive. For those $K_p - S_t \leq 0$, there is no doubt that the investor should wait for further opportunities. For an American put option seller, the conditions can be formulated by replacing $\overline{r_{t+1}}$ with $\underline{r_{t+1}}$.

Theorem 2

If $\overline{r_{t+1}} < (1 - \frac{K_p}{S_t})r$, then the American put option should be held at time t.

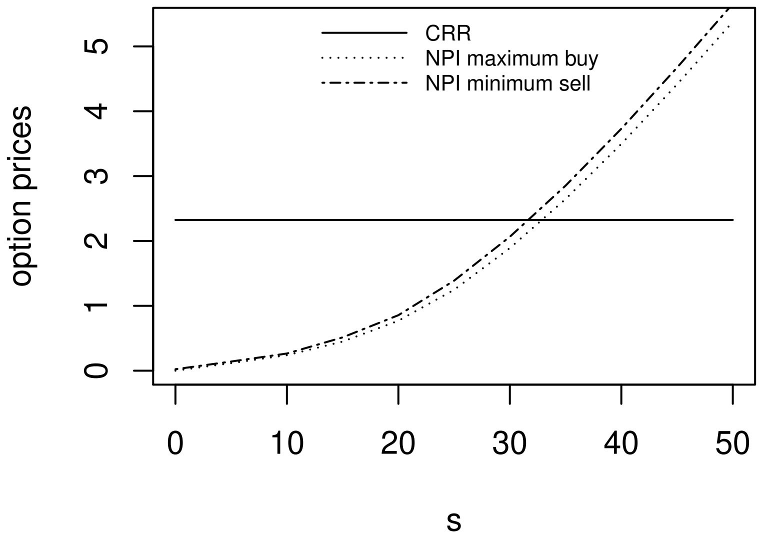
Proof

For an American put option, $(1 - \frac{K_p}{S_t})r$ is the lower boundary of all holding conditions, $(1 - \frac{K_p}{S_t})r < (1 - \frac{K_p}{S_t})r + \frac{(1-\overline{P_t}(S_t))v}{S_t}$ and $(1 - \frac{K_p}{S_t})r < (1 - \frac{K_p}{S_t})r + \frac{\overline{P_t}(S_t))w}{S_t}$ with constant positive values w and v. Thus, if the current upper expected return at time t is greater than $(1 - \frac{K_p}{S_t})r$, then it is optimal to hold this put option till the next time step. Note that the condition in Theorem 2 is a sufficient condition, but not a necessary one.

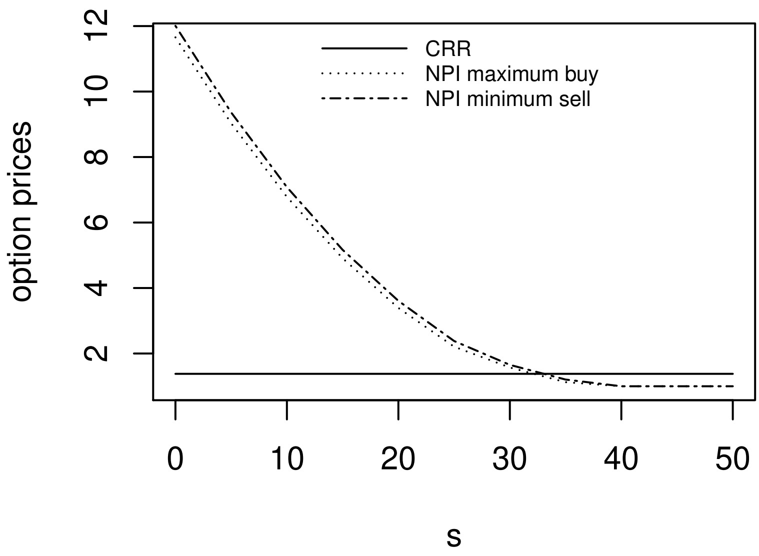
3.3 Comparison of CRR and NPI for American options

It is interesting to compare the CRR model and the NPI method for American option pricing. Following the procedure of the comparison for European options in Section 2.3, the performance study is justified by calculating the profit and loss of an investor using the NPI method and trading with the only other investor using the CRR model in two scenarios, the CRR model perfectly right or substantially wrong about the market. We firstly plot American option prices from the CRR model and the NPI model with fixed n but varying s.

In Figure 3.10, we study the comparison based on the same underlying asset ($S_0 = 20$, $K = 21$, $u = 1.1$, $d = 0.9$, $q = 0.65$, $n = 50$, $m = 4$). Here we need to mention that the CRR interest rate is equal to 0.03, which is calculated from



(a) Option prices with varying s for the American call option



(b) Option prices with varying s for the American put option

Figure 3.10: Comparison between the CRR model and the NPI method

the CRR model $r_{CRR} = qu + (1-q)d - 1$. While for the NPI discount rate is calculated based on s and n, $r = \frac{s}{n}u + (1 - \frac{s}{n})d - 1$. In Figure 3.10, for the call option when s varies from 0 to 50 the NPI maximum buying and minimum selling prices are getting higher, while the CRR price is a constant value intersecting with these two NPI prices. The two intersections are around the value $nq = 32.5$. If s is lower than the left intersection, then both NPI prices are lower than the CRR price, and if s is higher than the right intersections, both those prices are higher than the CRR price. Then for the case when s is in between the two intersections, the CRR price is in between the NPI prices. For the put option, the pattern of the NPI prices is opposite to that in the call option graph. The maximum buying and minimum selling prices go down along with s increasing. There are also two intersections between the CRR price and the two NPI prices around 32.5. When s is in the interval of two intersections, the CRR price is in the interval of the NPI prices. If s is lower than the left intersection value of s the NPI prices are higher than the CRR price, while if s is higher than the right intersection value of s the NPI prices are lower than the CRR price. The plot of the put option decreases to 1 instead of 0, because when s is close to n, the optimal exercise time is zero with a positive payoff equal to 1 in this example.

Since the American option can be early exercise, the expected option price is different from the corresponding European option, which leads to a different performance of the NPI method. As an important factor of the performance study, the optimal exercise time of the American option named as *stopping time* is studied in the following section.

3.3.1 Stopping times

Before the profit and loss calculation, it is necessary to study the different stopping times, the exercise times, of these two methods. Since the profit and loss contain two parts. One is from the price, and the other part is from the payoff.

However, different exercise times give us different option payoffs, for the stock price at the exercise time changes with the time. Then we would like to investigate the stopping times of both the NPI and the CRR price. In order to make the comparison, there are some values needed to be inputted: initial stock price S_0, upward movement factor u, downward movement factor d, predictive future time steps m, the constant probability q in the CRR model, strike price K, option type: the call or put option, and option trading position: buying or selling. The detailed steps of this simulation are listed below:

1. Simulate N paths of stock price movements. To do that, the indicator of upward movement $I^i_{t\ (i\in\{1\ldots N\},t\in\{1\ldots m\})}$ is needed.

   – For the NPI method, based on the historical data, we generate the indicator of upward movement

   $$I^i_{t\ (i\in\{1\ldots N\},t\in\{1\ldots m\})} = \begin{cases} 1 & \text{upward movement} \\ 0 & \text{downward movement} \end{cases}$$

   $I^i_{t\ (i\in\{1\ldots N\},t\in\{1\ldots m\})} \sim Bin(1, p = \frac{s+\theta_t}{n+t})$ for buying the stock and $I^i_{t\ (i\in\{1\ldots N\},t\in\{1\ldots m\})} \sim Bin(1, p = \frac{s+\theta_t+1}{n+t})$ for selling the stock, where θ_t is the cumulated number of I^i_t from time zero to time t.

   – In terms of the CRR model $I^i_{t\ (i\in\{1\ldots N\},t\in\{1\ldots m\})} \sim Bin(1, q)$.

   – The stock price at each step is $S^i_t = S^i_{t-1} u^{I^i_t} d^{(1-I^i_t)}$ with $S^i_0 = S_0$.

2. Calculate the instant value of each step, $h_t = S^i_t - K_c$ for call option and $h_t = K_p - S^i_t$ for put option. For the CRR model there is only one stock price in this calculation. Since there are two prices generated from the NPI method, the stock price in this calculation is chosen according to the input of the trading position. $S^i_t = S^i_{t-1} u^{I^i_t} d^{(1-I^i_t)}$, for buying a call option and selling a put option $I^i_{t\ (i\in\{1\ldots N\},t\in\{1\ldots m\})} \sim Bin(1, p = \frac{s+\theta_t}{n+t})$, for selling a call option and buying a put option $I^i_{t\ (i\in\{1\ldots N\},t\in\{1\ldots m\})} \sim Bin(1, p = \frac{s+\theta_t+1}{n+t})$.

3. Calculate the expected holding value of the option at time t, which is the discounted expected option value at time $t+1$ based on the NPI backward pricing method for an American option. As the option value at $t+1$ is the maximum of the instant value h_{t+1} and the expected holding value H_{t+1}, the expected holding value at t is $H_t = B(t, t+1) \max\{h_{t+1}, H_{t+1}\}$, where $B(t, t+1)$ is the discount factor from t to $t+1$.

4. Compare the instant value to the holding value from the initial time, and stop at the first time τ when the instant value is greater than the holding value, then τ is the optimal time for exercise.

In this simulation, because we want to compare the stopping time between the CRR model and the NPI method, we study the same American option based on the same underlying asset. With the same information towards both methods, $\frac{s}{n}$ is equal to q. The discount rate is the expected stock return $r = r_{CRR}$. Here we try to explain the stopping time comparison between the CRR model and the NPI method in the light of examples. According to Theorem 1, $\underline{r_{t+1}} > r$ is the condition to hold the option. To see the early exercise call option based on the NPI prediction, specific call options with $\underline{r_{t+1}}$ disobeyed Theorem 1 are studied in the examples. The first example is buying an in the money American call option with parameters as $K_c = 13$ and $T = 4$, on the basis of an underlying stock, $S_0 = 20$, $u = 1.1$, $d = 0.9$, $q = \frac{s}{n} = 0.52$, $s = 26$ and $n = 50$.

Figure 3.11 clearly shows us for the NPI method the optimal exercise time can be time 1, 3 or 4 depending on different paths of the underlying stock price, while for the CRR model, this call option is optimal to be exercised at maturity as it claimed in the rational trading theory. In the 20000 simulations of the NPI method, 9667 times stop at time 1, 2396 times stop at time 3, and 7937 times stop at time 4. The reason that time 2 is skipped is clearly shown in the binomial tree. As we can see from Figure 3.12, this American call option has a higher instant value than its corresponding holding at time 1, 2 and 3. However, if it attempts to reach early

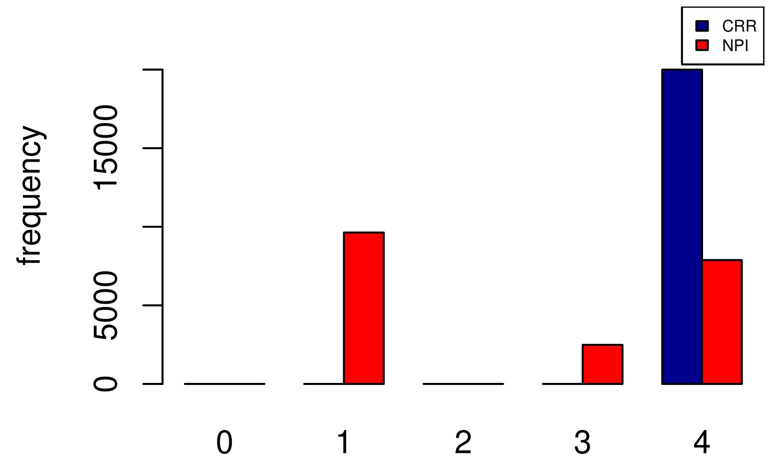


Figure 3.11: Stopping time from both the CRR model and the NPI method (20000 times simulation)

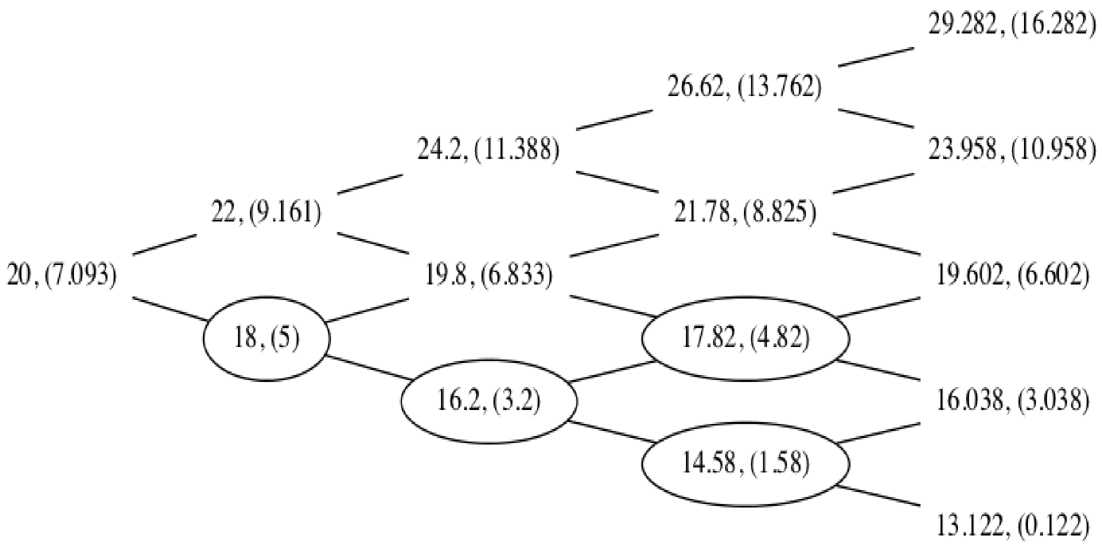


Figure 3.12: Binomial tree plot for the NPI American call option

exercise node at time 2, it is supposed to encounter the early exercise node at time 1. In this circumstance, the investor will choose to exercise at time 1 rather than time 2. We also calculate the average stopping time for the NPI method in this example,

Method \ Time	0	1	2	3	4	5
NPI	0	0	3214	3877	0	12606
CRR	0	0	3214	3822	0	12964

Table 3.1: Stopping time from both CRR and NPI method (20000 times simulation)

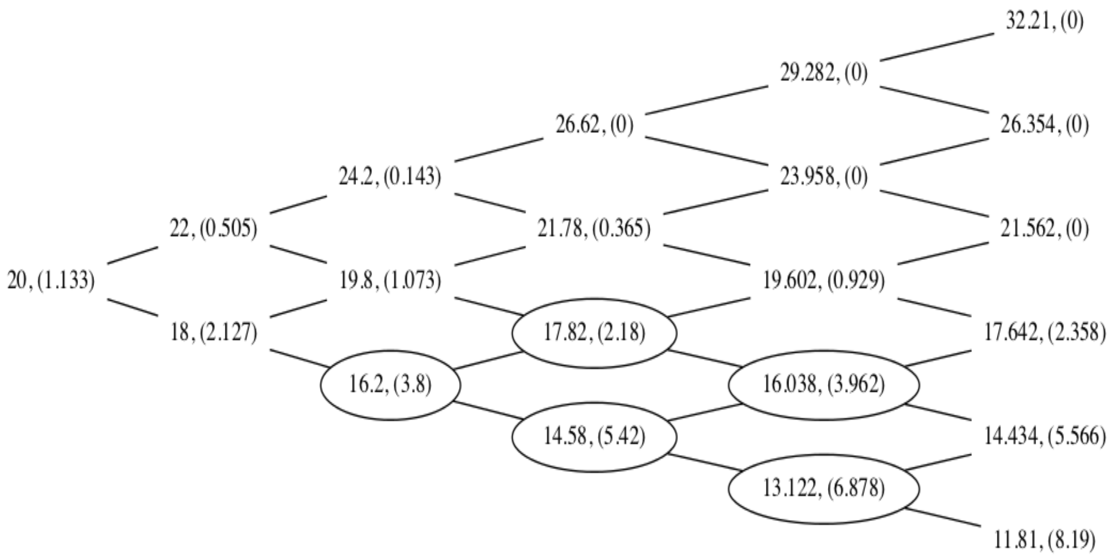


Figure 3.13: Binomial tree plot for the NPI American put option

which is 2.40605. Regarding to selling the same American call option, since the probability of upward movement for each time step is higher than the probability for buying this option, we expect the average stopping time for the NPI method to be greater than the time of buying the call option, which is shown by the result from the 20000 times simulation of selling a call option, the average stopping time is 3.89335.

An example of the put option comparison is also interesting. This time we choose to study the stopping time of selling an American put option. In order to make sure that the early exercise of the American put option happens, in this example we price a put option with the inputs as $S_0 = K_p = 20$, $u = 1.1$, $d = 0.9$, $m = 5$, $n = 500$ and $q = \frac{s}{n} = 0.6$. From Table 3.1, we can see that for both the CRR and the NPI methods this American put option is optimal to be exercised before the maturity under some circumstances, and the stopping time is either 2, 3 or 5. However, the

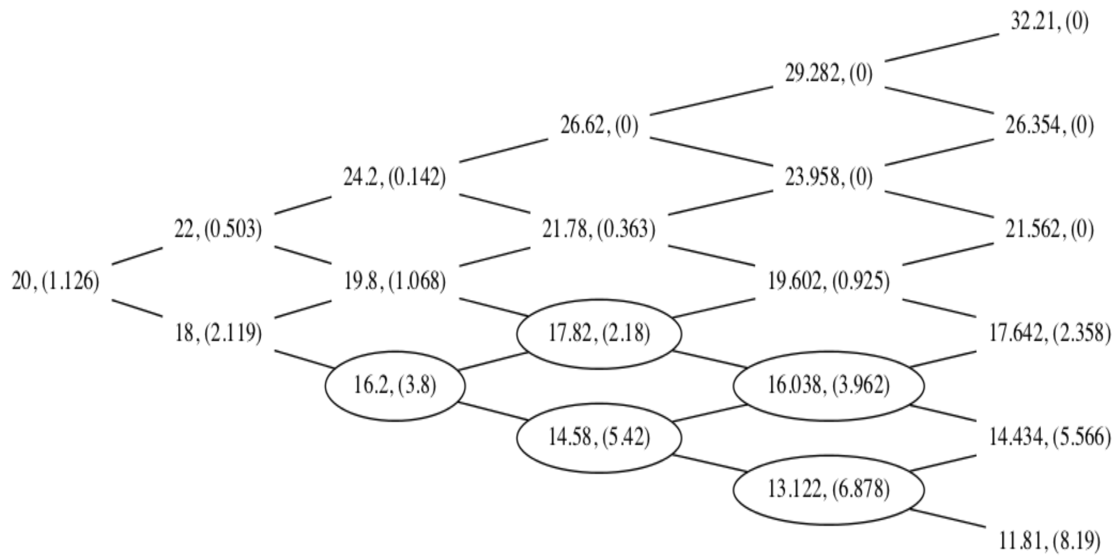
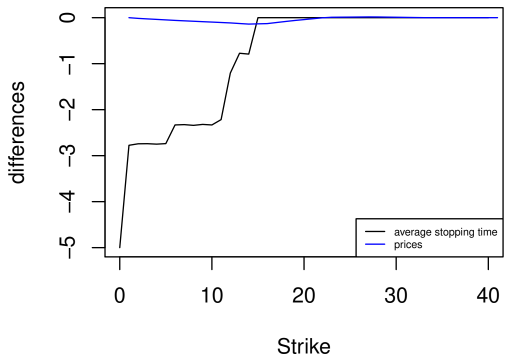


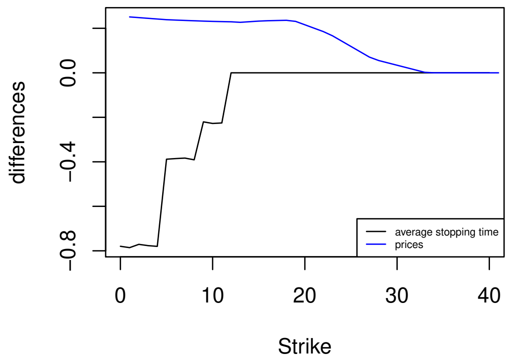
Figure 3.14: Binomial tree plot for the CRR American put option

NPI stopping time is more at time 3 and less at time 5 than the CRR model, and both methods have the same times of stopping time 2. As shown in Figures 3.13 and 3.14, the early exercise node are the same for both methods, while the option price from the NPI method is higher than that from the CRR method. The NPI method keeps learning from the data, then it assigns more probability to the path which has a lower stock price with a higher payoff for the put option. Thus, when it comes to the simulation, the NPI method would have a higher probability to encounter the early exercise case, and our results support this. The average stopping time of the CRR model is 4.1357, whereas the average stopping time of the NPI method is 4.1302. This leading to a higher NPI option price 1.133 than the CRR option price 1.126. We also simulate the buying position's average stopping time, which is equal to 4.1359.

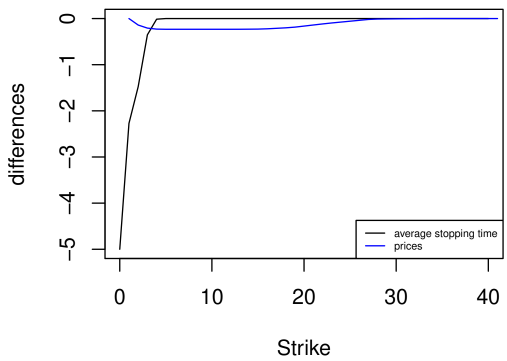
In order to see the influence of the stopping time towards the price with varying moneyness between the two methods, we also plot the differences of average stopping time and option prices for the two methods based on different moneyness character, in the money, at the money and out of the money, in Figures 3.15 and 3.17. In this simulation, the information of the stock is $S_0 = 20$, $u = 1.1$ and $d = 0.9$. As we use



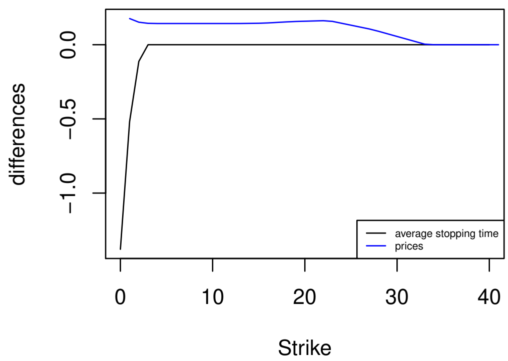
(a) $n = 50$ $q = 0.52$(buy)



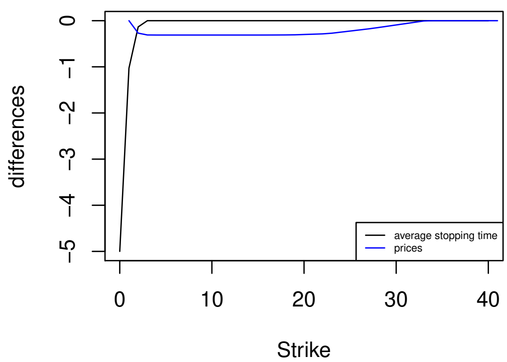
(b) $n = 50$, $q = 0.52$ (sell)



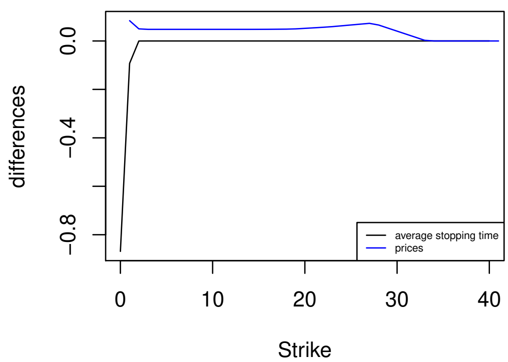
(c) $n = 50$, $q = 0.7$ (buy)



(d) $n = 50$, $q = 0.7$ (sell)

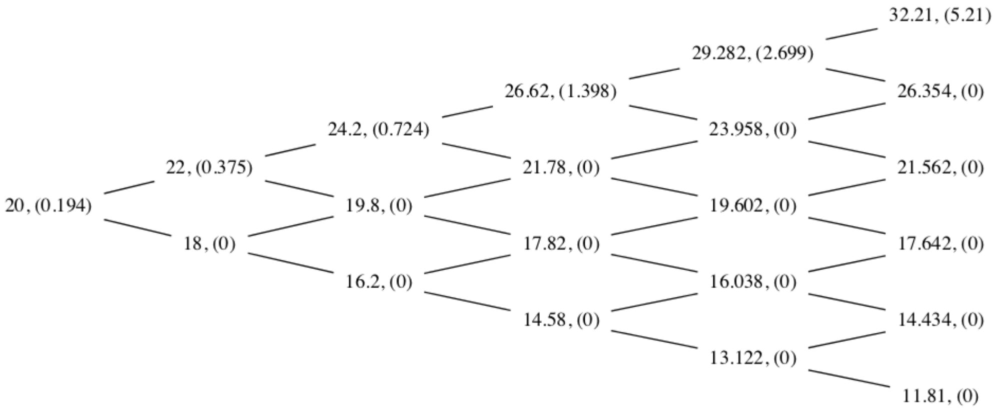


(e) $n = 50$, $q = 0.9$ (buy)

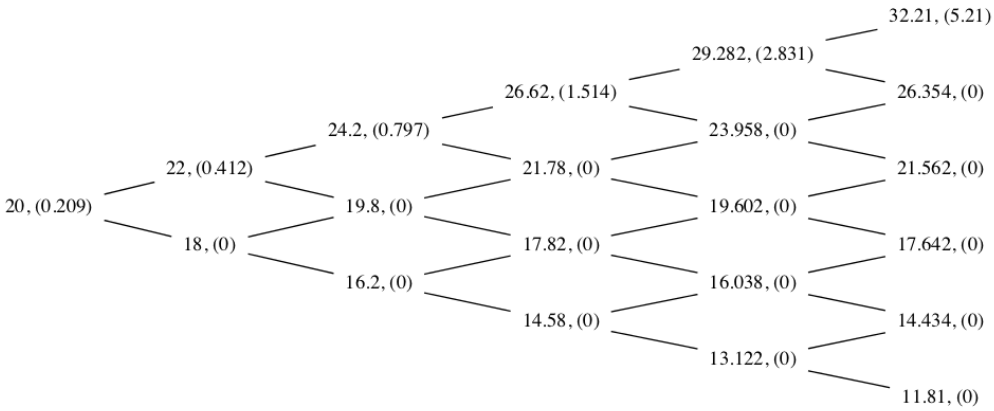


(f) $n = 50$, $q = 0.9$ (sell)

Figure 3.15: Differences of stopping times and prices between the two methods (Call option)



(a) CRR call option price with $K = 27$



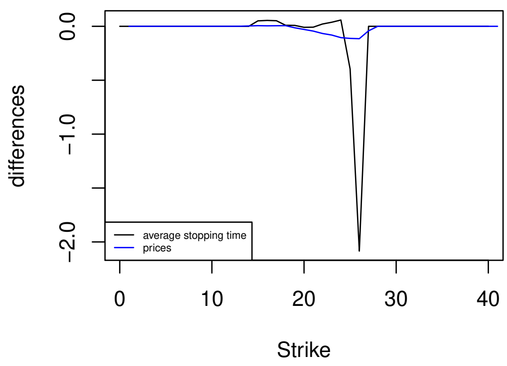
(b) NPI maximum buying price of a call option with $K = 27$

Figure 3.16: Binomial trees of specific cases

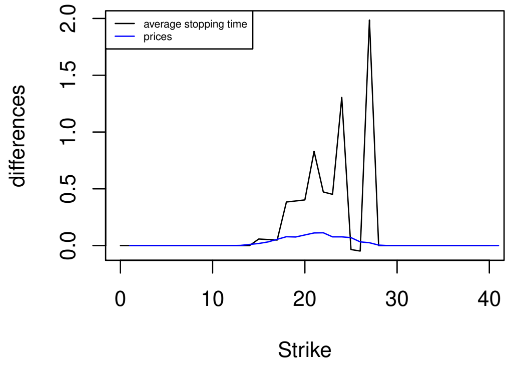
the predictive value from the CRR model as the benchmark, we use the expected stock price return from the CRR model as the discount rate $r = r_{CRR}$. First, we assume the historical data is $n = 50$, and the proportion of increasing prices is $\frac{s}{n} = q$. The American options in this simulation are the ones with five future time steps. In this simulation, we vary the strike price from 0 to 40, and for each strike price, we run 10000 stock paths to calculate the average stopping time. Then we compare two methods by calculating the differences in the average stopping time and option prices.

As acknowledged, the CRR model always follows the rational trading theory ”Never early exercise an American call option”, then we know that all American call options are held until the maturity based on the CRR model, with 5 as the average stopping time in this simulation. We can conclude from Figure 3.15 that for call options, when the call option is deep in the money $0 < K_c << S_0$, the NPI method will predict this call option to be early exercised, no matter buying or selling position. As the strike price getting larger, at some point, the call option is expected to be exercised at maturity upon both methods. Here, the option prices from the two methods are different, while the average stopping time is the same. But if the strike price gets too large, this option is predicted to be exercised in both methods, so both the differences of average stopping time and option prices are zero. The maximum buying price is always lower than or equal to the CRR price for in the money and at the money call options, but it can be higher than the CRR option when it is out of the money. For example, in the case q is equal to 0.52 and K is 27, Figure 3.16 shows that because of the large strike price, the only stock path holding a positive payoff is the first path. The NPI method adjusts the probability along with the data. Then as predictive future time steps get longer, the probability of upwards movements gets higher, making the expectation of the option payoff greater than the expected option payoff from the CRR model. After the discount procedure, the NPI has a higher option price than that from the CRR model. When the call option is deep out of the money, the call option has a zero payoff, so the costs from the two methods equal zero. However, due to different option trading positions, we can see from Figure 3.15 that a seller who uses the NPI method would expect the same call option to be held longer than as a call option buyer. As the NPI call option seller, the upper probability of each upwards movement is used, which makes it easier to reach the criteria to be held for NPI call options. Then we can see that the earliest average stopping time of selling a deep in the money option $S_0 >> K_c$ is around time 4.2 while the one of buying the same call option is time 0. The circumstance for selling this call option with same option price from both methods is when this call option is out of the money with zero payoff, then in this circumstance, both the CRR model and the NPI method will generate a zero option price. Whereas, the situation that this call option is expected to be early exercised, the difference in the prices is greater than that without early exercise but with none zero option price. Conclusively, the disagreement of option prices are raised by two parts, one is the different probability in the two methods, and the other part is the different stopping time. As we are varying q from 0.52 to 0.7 till 0.9, the early exercise case is getting harder to be achieved, only happening to the call options really deep in the money. The differences of prices also get smaller with larger q, due to larger q leads to a more substantial s. In terms of each predictive time step, the variability of NPI probability is lower comparing to small s.

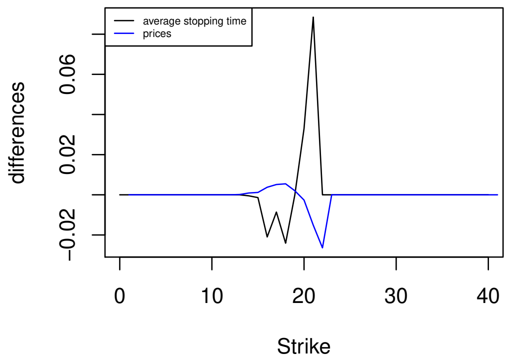
For the put option, from Figure 3.17, as it is also possible that the put option to be predicted as the early exercise option in the CRR model, the patterns of stopping time and price differences are different from that of the call option. Generally, for both seller and buyer, deep in the money and out of the money put options have the same option prices and stopping time on the basis of the two methods, but the reasons are different. Deep out of the money put options are with a zero payoff leading to a zero option price, while deep in the money put options are expected to be exercised at the initial time, with an option price $K_p - S_0$. If a put option is at the money, both patterns of different stopping time and prices fluctuate. To illustrate the comparison detailedly, let us focus on the case of buying the option with $q = 0.52$. In this case, there are two apparent fluctuations of stopping time; one is that strike price is around 18, the NPI stopping time is slightly later than the CRR stopping time. The other fluctuation dramatically happens when the strike price is 26, which the NPI buying put option is expected to be exercised at two-time steps earlier. We can find the reason in their corresponding binomial trees in Figures 3.18 and 3.19. Figures 3.18 (a) and (b) are the binomial trees of buying the same put option with $K = 18$ based on the CRR model and the NPI method, respectively. The stock price becomes either 14.58 at time 3 or 16.038 at time 4 causing an early exercise situation grounded on both pricing procedures. However,



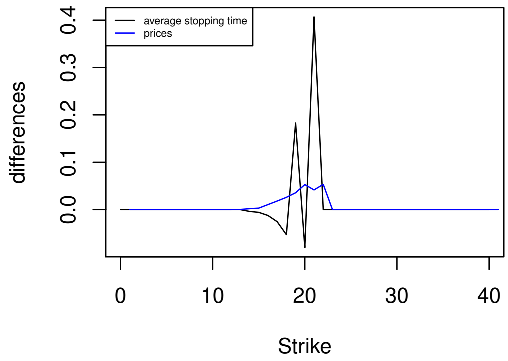
(a) $n = 50,\ q = 0.52$ (buy)



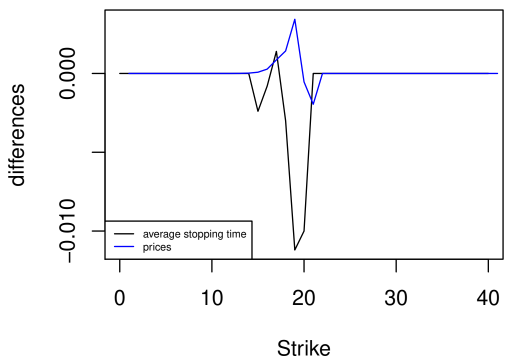
(b) $n = 50,\ q = 0.52$ (sell)



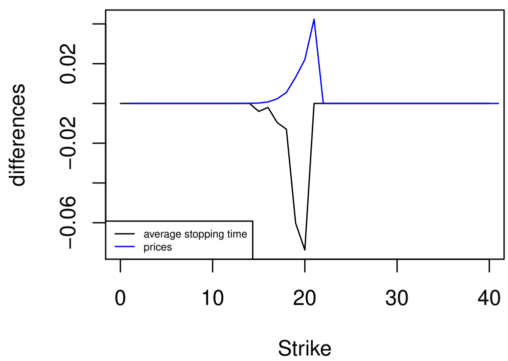
(c) $n = 50,\ q = 0.7$ (buy)



(d) $n = 50,\ q = 0.7$ (sell)

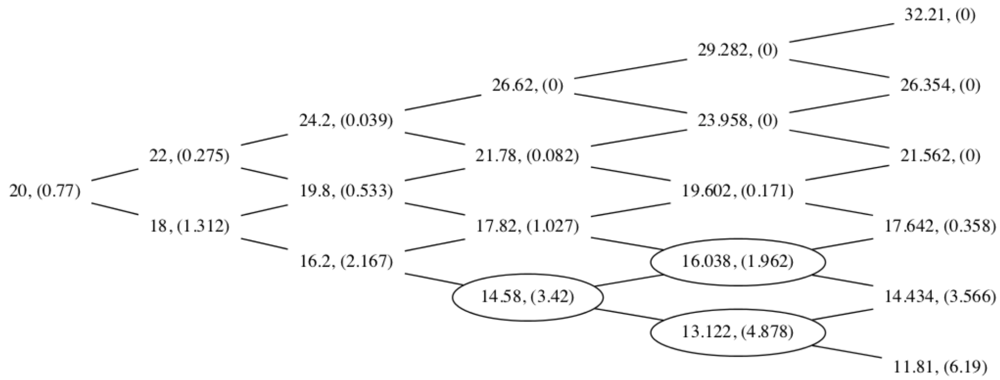


(e) $n = 50,\ q = 0.9$ (buy)

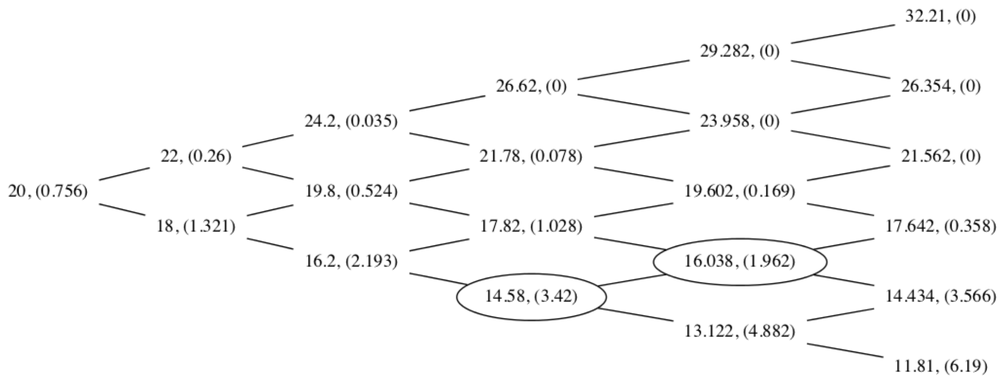


(f) $n = 50,\ q = 0.9$ (sell)

Figure 3.17: Differences of stopping time and prices between the two methods (Put option)

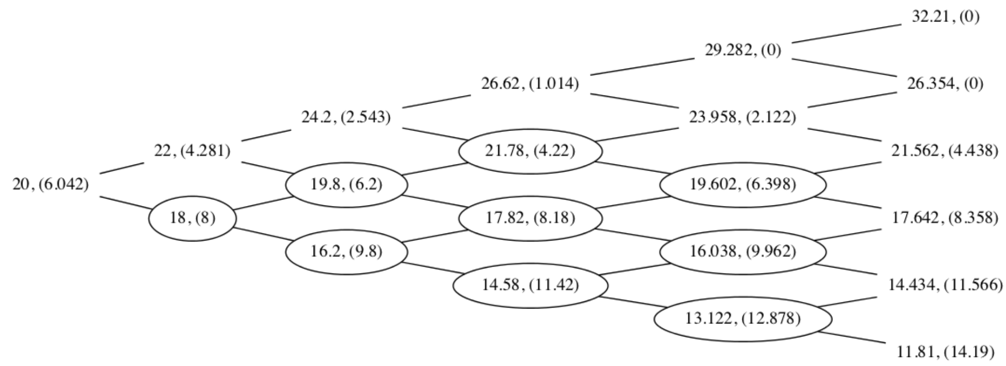


(a) CRR put option price with $K = 18$

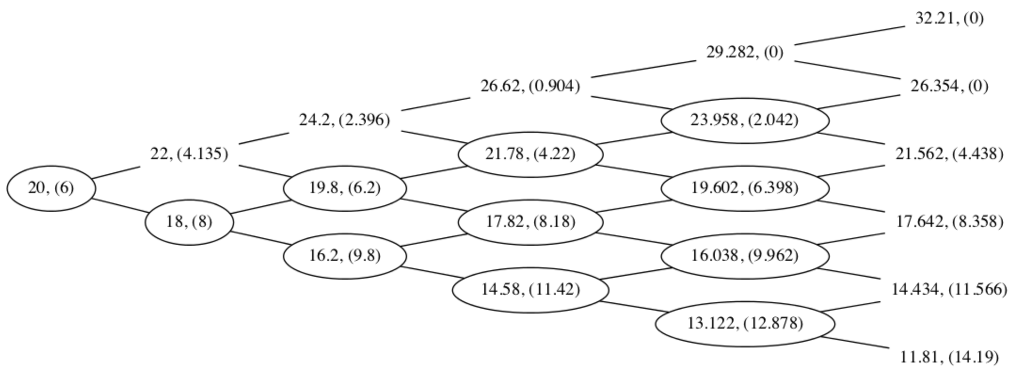


(b) NPI maximum buying price of a put option price with $K = 18$

Figure 3.18: Binomial trees of specific cases



(a) CRR put option price with $K = 26$

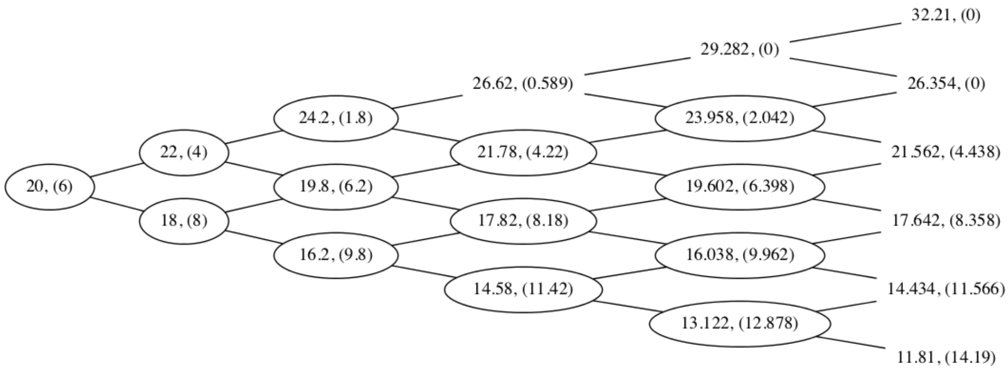


(b) NPI maximum buying price of a put option price with $K = 26$

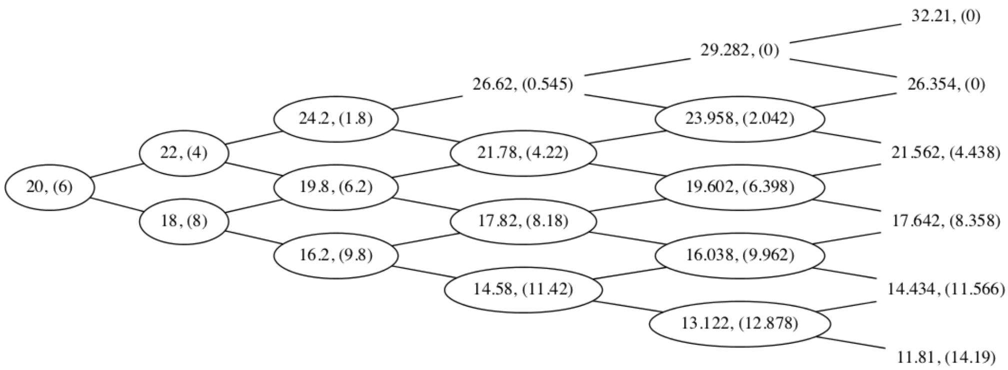
Figure 3.19: Binomial trees of specific cases

to price the NPI maximum buying price, an upper probability is assigned to the path with a higher stock price in each step binomial tree, so it is easier to reach 16.038 at time 4 based on the NPI method. Thus, the average NPI stop time will be slightly later but no more than one-time difference compared to the CRR results, for early exercise moments in the binomial tree is the same in both binomial trees. The story is different for buying a put option with $K = 26$. From Figures 3.19 (a) and (b) we can see that the early exercise moments in the binomial tree are different for these two methods. In the CRR model, there are three early exercise situations, $S_1 = 18.8$ at time 1, $S_2 = 19.8$ at time 2 and $S_3 = 21.78$ at time 3, while the NPI method predicts the same buying put option to be exercised at the initial time. The average stopping time is quite different between the two methods around two-time steps. From the aspect of option price differences, under the assumption $\frac{s}{n} = q$, along with large q, the relevant difference between the probabilities at each time does not vary a lot at each time step compared to a smaller q. Thus, the option payoff binomial tree for both methods verge to be similar. For example, we plot the option binomial trees with $K = 26$ in two different probabilities $q = 0.7$ and $q = 0.9$ in Figure 3.20. As we discussed before, in the case of $q = 0.52$, the difference between two methods predictions holds the largest value for the option with $K = 26$, while Figure 3.20 shows as q increases binomial trees of the two methods for this $K = 26$ put option are looking to resemblance exercised early at the initial time. In the light of put option holding or exercise conditions, intuitively it is not hard to understand, large q is barely possible to fit the criteria for holding the option. As long as early exercise happened, the current option values for the two option methods are the same at that time. As more nodes holding the early exercise situation, the binomial trees of two methods are similar, leading to smaller differences of average stopping time and prices.

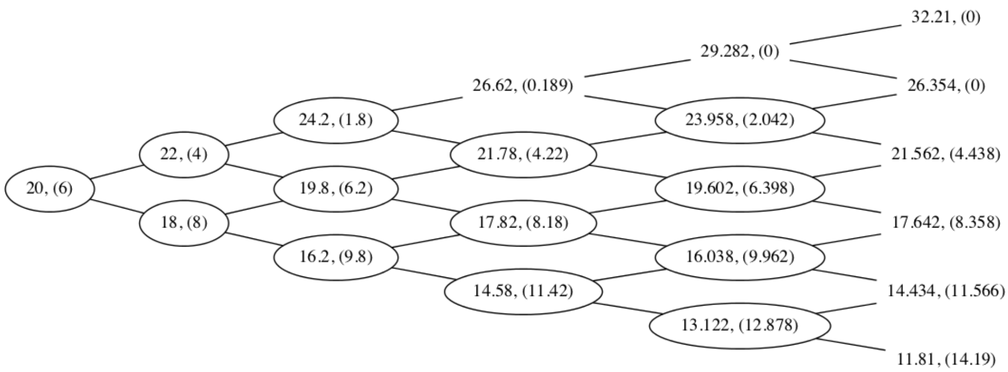
The next simulation leads us to the study of the amount of historical data n influences. This time we assume $n = 252$, the number of trading days in one calendar year, other than this, all other inputs are the same as that in the last simulation shown in Figures 3.15 and 3.17. The simulation outcomes are listed in Figures 3.21 and 3.22. In terms of call options, we are told from the figure that it is harder to encounter the early exercise situation for different moneyness options with larger n, only when this call option is deep in the money, $K_c << S_0$. The reason is for larger n, the variability of the probability in each time step is lower, which $\overline{P_t} = \frac{s+t-i+2}{n+t} \approx q$ and $\underline{P_t} = \frac{s+t-i+1}{n+t} \approx q$ with $i \in \{1 \dots t+1\}$ and $t \in \{0 \dots T-1\}$ are true, then only the circumstance that the strike price is really lower than the initial stock price will trigger the early exercise. It is also visible that for the situation of options with the same average stopping time predicted to be exercised at maturity from two methods, the option price differences are smaller than that with a small n. The reason is that a larger n narrows down the gap between the maximum buying



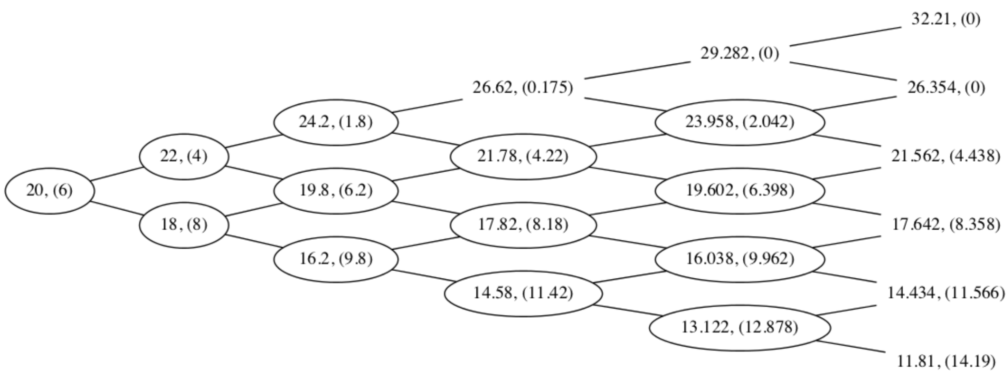
(a) CRR put option price with $K = 26$ $q = 0.7$



(b) NPI maximum buying price of a put option with $K = 26$ $q = 0.7$

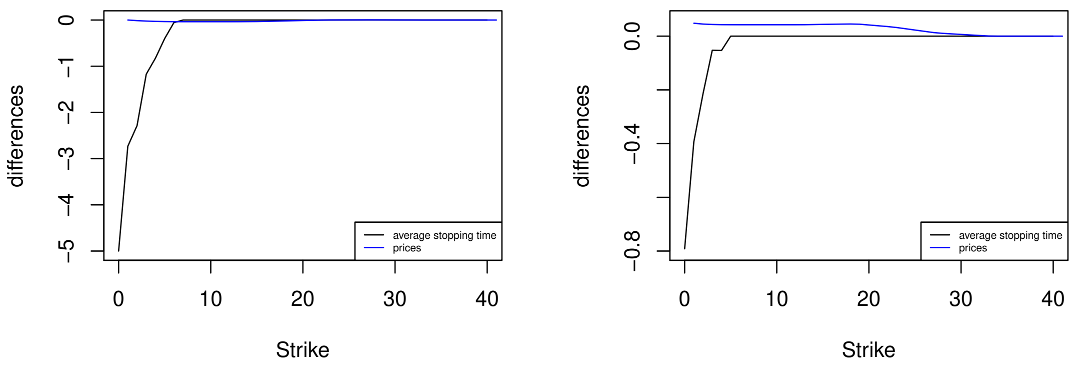


(c) CRR put option price with $K = 26$ $q = 0.9$

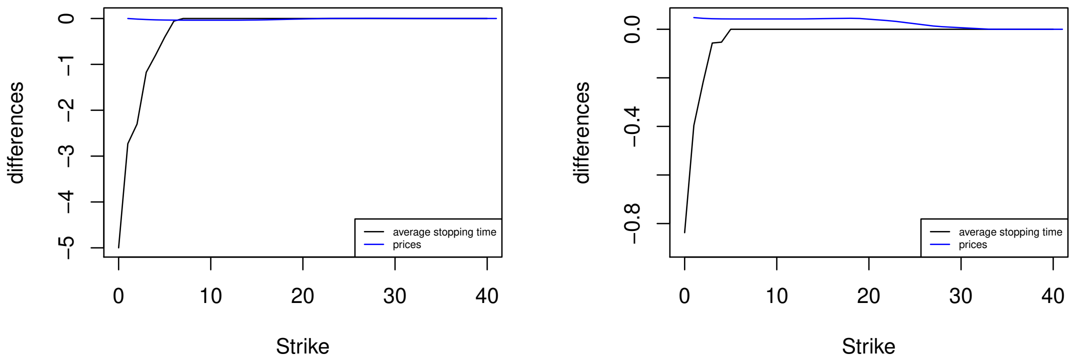


(d) NPI maximum buying price of a put option with $K = 26$ $q = 0.9$

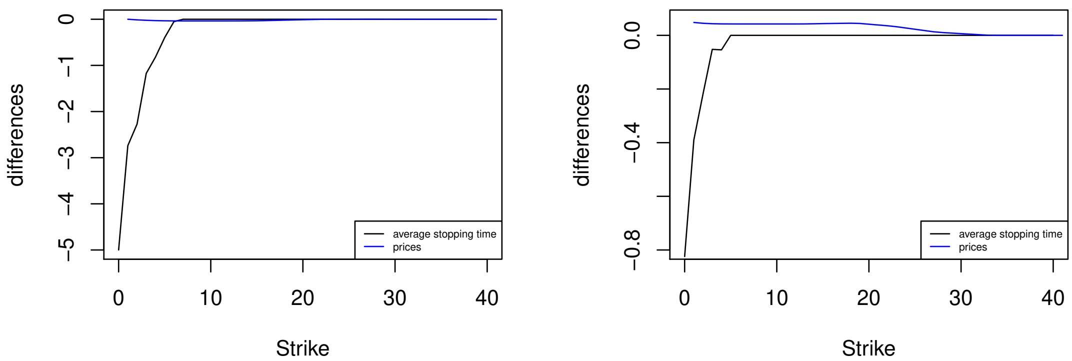
Figure 3.20: Binomial trees of specific cases to compare q influences



(a) Call with $n = 252$, $q = 0.52$ (buy and sell)

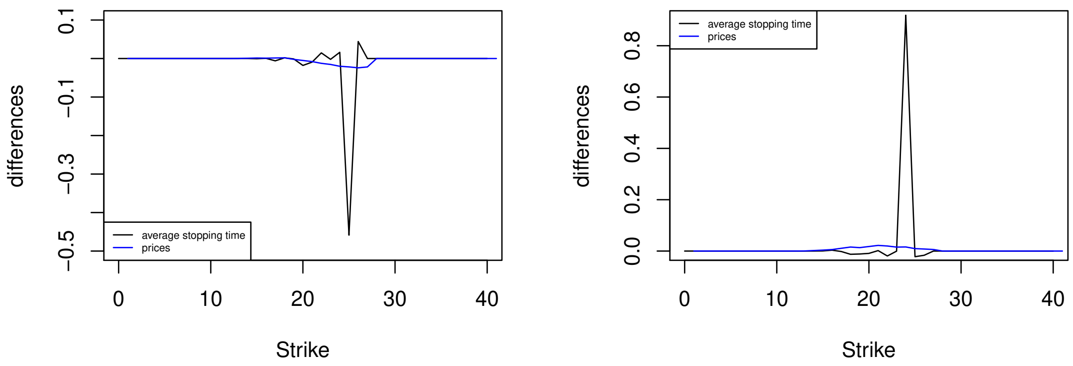


(b) Call with $n = 252$, $q = 0.7$ (buy and sell)

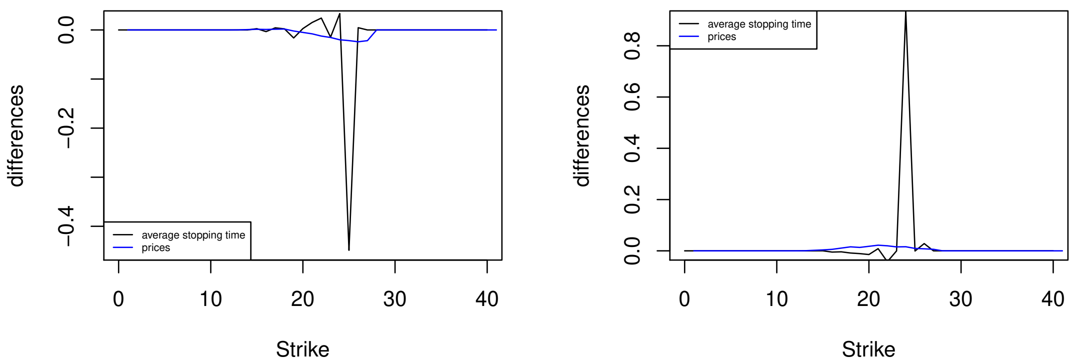


(c) Call with $n = 252$, $q = 0.9$ (buy and sell)

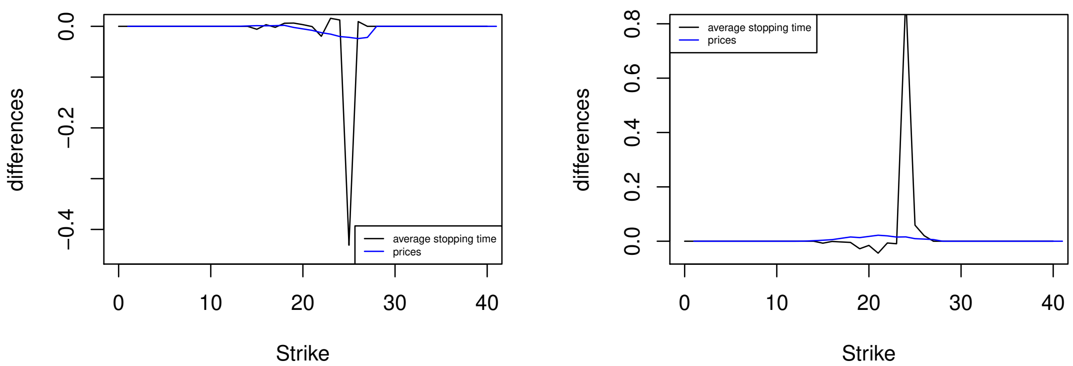
Figure 3.21: Differences of stopping time and prices between the two methods



(a) Put with $n = 252$, $q = 0.52$ (buy and sell)



(b) Put with $n = 252$, $q = 0.7$ (buy and sell)



(c) Put with $n = 252$, $q = 0.9$ (buy and sell)

Figure 3.22: Differences of stopping time and prices between the two methods

price and the minimum selling price and makes the outcome closer to the CRR result. $\frac{s}{n} = q$, the CRR price locates in between the maximum buying price and the minimum selling price, then the smaller the gap of two boundary prices is the smaller the differences between prices for two methods are. For put options, the fluctuation also happens when options are around at the money shown in Figure 3.22. However, the highest fluctuation of average stopping time differences is less than one-time step, meaning the larger n makes the binomial trees' exercise moments from the two methods similar to each other. Furthermore, the larger n gives a resist effect of varying q to both call and put options. It is easier to understand that larger n makes less diversity between the two methods when the assumption is $\frac{s}{n} = q$, then surely the pattern of differences from these two methods is supposed to be smoother and less fluctuated.

All simulation results in this section show that the stopping time influences on the prediction of option price for both two methods. Different stopping time based on the CRR model and the NPI method enhances the differences of the outcomes from these two methods. The effect of the historical data is also considered, finding that sufficient historical data can ease the influence of the stopping time on the prediction from two methods. Thus, when the performance of the NPI method is investigated, the impact of the stopping time needs to be considered.

3.3.2 Profit and loss

In this section, we investigate the performance of the NPI method by calculating the profit and loss in a circumstance that an investor using the NPI method trades with an investor using the CRR model. Inspired by the scenarios in European option study in Chapter 2, we assume:

1. There are only two investors in the market; one uses the CRR model while the other one uses the NPI method.

2. The trade is triggered if the CRR price is higher than or equal to the minimum

selling price or lower than or equal to the maximum buying price from the NPI method. And the trading price is always the NPI price, because we want to know the worse situation that the NPI investor can encounter.

We study the profit and loss (P&L) in two extreme scenarios: one is that the CRR model correctly predicts what happens to the future market, while the other is based on the CRR investor uses the wrong assumption about the market. According to the paths of stock price simulated following the steps shown in Section 3.3.1 as well as the option prices from two methods, we get the NPI profit and loss for each path based on different scenarios. More details of the calculations are presented below:

Scenario 1: The CRR assumptions are correct

In the market, there is a real probability of upward movement p, and in this scenario, the CRR assumed upward movement probability q is equal to p. When the NPI maximum buying price is higher than or equal to the CRR price, $\underline{V_0} \geq V_0^{CRR}$, the NPI investor will buy this American call or put option at this maximum buying price $\underline{V_0}$. At the stopping time, the NPI investor will get the option payoff calculated based on the CRR model, $V_\tau^{CRR} = \max_q \{0, S_T - K_c\}$ for the call option, because in the CRR model an American call option will never be early exercise, and $V_\tau^{CRR} = \max_q \{0, K_p - S_\tau\}$ for the put option. Due to different stopping time, a risk-free time value from τ to maturity T is used to calibrate the payoff, which means we assume the payoff from an early exercise option will be invested in a risk-free product until the maturity. Then for a call option, the profit and loss of the NPI investor is,

$$P\&L_c = V_\tau^{CRR}(1 + r_f)^{T-\tau} - \underline{V_0} = \max_q \{0, S_T - K_c\} - \underline{V_0} \tag{3.20}$$

And for a put option, the profit and loss of the NPI investor is,

$$P\&L_p = V_\tau^{CRR}(1 + r_f)^{T-\tau} - \underline{V_0} = \max_q \{0, K_p - S_\tau\} (1 + r_f)^{T-\tau} - \underline{V_0} \tag{3.21}$$

If the CRR price falls in the interval of NPI prices, $\underline{V_0} < V_0^{CRR} < \overline{V_0}$, then there is no trade. The NPI investor would like to sell the American option if the CRR price is higher than or equal to the minimum selling price $V_0^{CRR} \geq \overline{V_0}$. The option price $\overline{V_0}$ is the profit and will be invested in a risk-free product under our assumption. However, the loss of the NPI investor, $V_\tau^{CRR} = \max_q\{0, S_T - K_c\}$ for the call option and $V_\tau^{CRR} = \max_q\{0, K_p - S_\tau\}$ for the put option, occurs when this option is exercised. Then for a call option, the profit and loss formula is,

$$P\&L_c = \overline{V_0}(1 + r_f)^T - V_\tau^{CRR} = \overline{V_0}(1 + r_f)^T - \max_q\{0, S_T - K_c\} \tag{3.22}$$

For a put option, the profit and loss formula is,

$$P\&L_p = \overline{V_0}(1 + r_f)^T - V_\tau^{CRR} = \overline{V_0}(1 + r_f)^T - \max_q\{0, K_p - S_\tau\} \tag{3.23}$$

Scenario 2: the CRR assumptions are wrong

For this scenario, the CRR assumptions are wrong, which means $q \neq p$. As an option buyer, the NPI investor will buy this option when the CRR price is lower than or equal to the maximum buying price and exercise it at the optimal stopping time τ. However, as the CRR assumptions are wrong, meaning instead of q the probability of stock price upward movement is p, the NPI investor will get the payoff as $V_\tau^p = \max_p\{0, S_T^p - K_c\}$ for the call option at the exercise time and $V_\tau^p = \max_p\{0, K_p - S_\tau^p\}$ for the put option. Here S_τ^p is simulated follows the CRR simulation steps, so actually, V_τ^p is the option payoff calculated from the CRR model with probability p instead of q, and we assume this value as a real compensation of the option exercise from the market in this scenario. For time value calibration, we assume both the early exercise payoff and the earned option price will be invested in the risk-free product until the maturity. So for a call option, the NPI profit and loss formula is

$$P\&L_c = V_\tau^p(1+r_f)^{T-\tau} - \underline{V_0} = \max_p\{0, S_T^p - K_c\} - \underline{V_0} \tag{3.24}$$

and for a put option

$$P\&L_p = V_\tau^p(1+r_f)^{T-\tau} - \underline{V_0} = \max_p\{0, K_p - S_\tau^p\}(1+r_f)^{T-\tau} - \underline{V_0} \tag{3.25}$$

If $\underline{V_0} < V_0^{CRR} < \overline{V_0}$, there is no transaction between the NPI investor and the CRR investor. When the CRR price is higher than or equal to the minimum selling price, $V_0^{CRR} \geq \overline{V_0}$, the NPI investor prefers to sell this option at the minimum selling price and save it in a risk-free account. When the CRR investor exercise this option, the NPI investor will pay the option payoff $V_\tau^p = \max_p\{0, S_T^p - K_c\}$ for the call option and $V_\tau^p = \max_p\{0, K_p - S_\tau^p\}$ for the put option. Then the NPI profit and loss for a call option can be formulated.

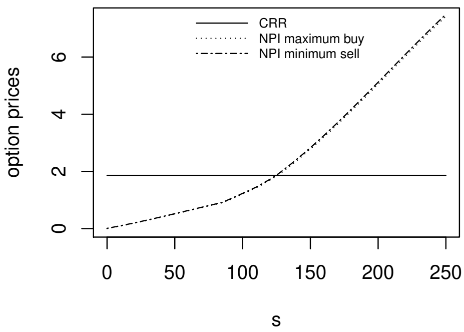
$$P\&L_c = \overline{V_0}(1+r_f)^T - V_\tau^p = \overline{V_0}(1+r_f)^T - \max_p\{0, S_T^p - K_c\} \tag{3.26}$$

And the formula of the NPI P&L for a put option is,

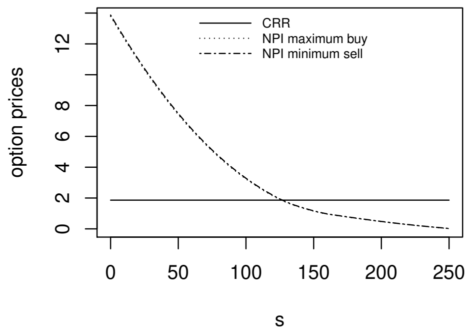
$$P\&L_p = \overline{V_0}(1+r_f)^T - V_\tau^p = \overline{V_0}(1+r_f)^T - \max_p\{0, K_p - S_\tau^p\} \tag{3.27}$$

Example 3.3.1

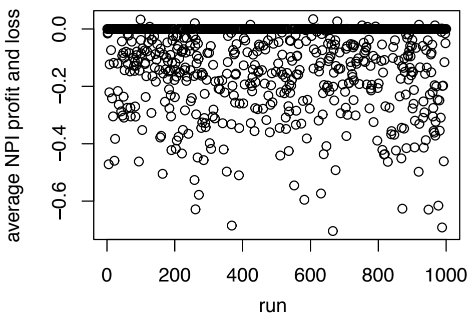
In this example, we calculate the profit or loss of the NPI investor trading with the CRR investor in Scenario 1 and Scenario 2. By investigating the NPI profit and loss in these two scenarios, we study the performance of the NPI method for American option pricing.



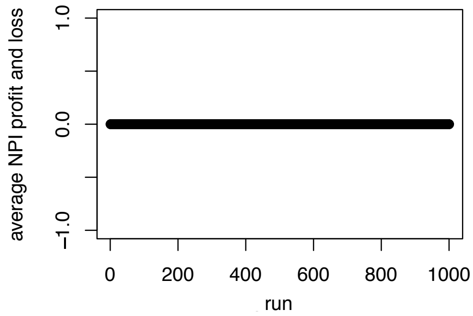
(a) NPI and CRR option prices (call)



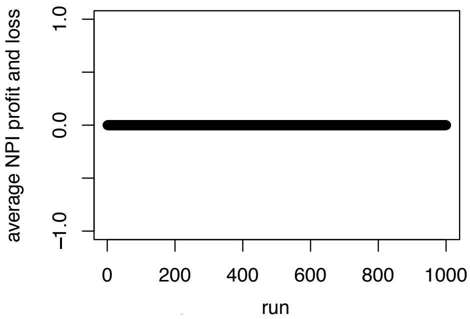
(b) NPI and CRR option prices (put)



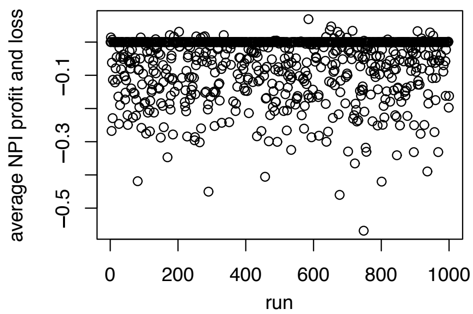
(c) NPI P&L for buying call option



(d) NPI P&L for selling call option



(e) NPI P&L for buying put option



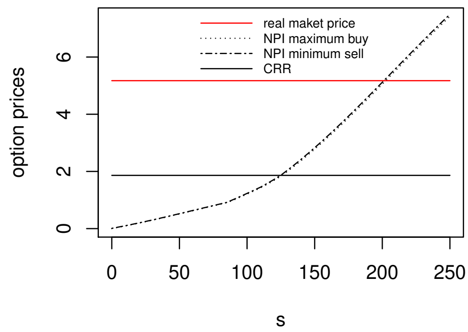
(f) NPI P&L for selling put option

Figure 3.23: NPI profit and loss for Example 3.3.1 in Scenario 1 ($q = 0.5$, $n = 252$, $p = 0.8$)

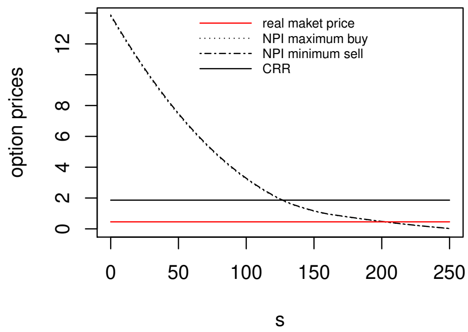
We randomize s according to two different scenarios. As s represents the number of increased prices in the historical data, s follows the binomial distribution. Two scenarios are distinguished by the different binomial distributions of s, which in the first scenario $s \sim Bin(n, q)$ whereas in the second scenario $s \sim Bin(n, p)$. Based on each randomized s, an average of profit and loss for the NPI investor is generated, which is the average value from N paths profit and loss. For our simulation, we randomly generate 1000 s values, and for each s, $N = 10000$ stock price paths are simulated to be used as underlying asset prices. In our first example, Scenario 1 follows $s \sim Bin(252, 0.5)$ with $q = p = 0.5$, and Scenario 2 follows $s \sim Bin(252, 0.8)$ with $q = 0.5$ but $p = 0.8$. The underlying asset is still the same asset as in the other examples in Chapter 3, with $S_0 = 20$, $u = 1.1$ and $d = 0.9$. We decide to investigate in the at the money options by simulation, $K = S_0 = 20$. According to the stopping time study, we know that for the put option it has the most different result from two methods when it is at the money. So we choose an at the money option as our study example. The option is an American option with $T = 5$ and discount rate $r = qu + (1 - q)d - 1$ for the CRR price and $r = \frac{s}{n}u + (1 - \frac{s}{n})d - 1$ for NPI prices. In this example, we assume the interest rate that investment of NPI investor before the maturity in the risk-free market is 0.002. For an NPI investor, buying the call option or selling the put option in this example is a wise choice, because the expectation of the stock price is positive leading to a positive expected payoff for the call option and none expected payoff for the put option. It does not mean other trading position would not give the investor a positive payoff, but these two trading actions are safer than others. We calculate the average $P\&L$ for each randomized s, all outcomes are demonstrated as Figures 3.23 and 3.24.

Figures 3.23 (a) and (b) list the CRR price, the maximum buying price and the minimum selling price with varying s in Scenario 1. Subfigure (c), (d), (e) and (f) are the average NPI profit and loss for each randomized s in different trading positions and scenarios. This four subfigures in Figure 3.23 can tell us that in Scenario 1, the NPI investor faces a high chance to lose the money in the option trade. For a call option, a longing position is possible to lead to a loss, because the NPI investor is willing to buy this option with a higher price, which depending on the value of s, but the investor can only get the payoff at maturity upon the CRR model with q. The payoff is lower than the maximum buying price leading to a loss, and the NPI loses the chance to invest the payoff into another market, because he cannot get the payment before the maturity. Another possible situation for longing a call option earns no profit or loss, as s is small, making either the CRR price higher than the NPI maximum buying price or the discount rate negative, both two situations prevent the investor from buying this call option. Selling this call option leads to no profit or loss, due to that s is either too small causing a negative discount rate or too large making the CRR price lower than the minimum selling price. Then neither situation triggers the selling this call option action.

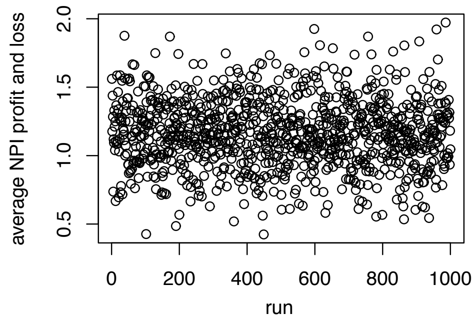
Figures 3.23 (e) and (f) show that selling a put option is not a profitable investment as well, and buying this put option is meaningless. For this put option buyer, when s is small, the discount rate is negative, while when s is large enough for a positive discount rate, but the CRR price is higher than the maximum buying price. According to the second graph in Figure 3.23 (b), when s is large the NPI investor is in this game, meaning the NPI investor sells the put option at a lower price than the CRR price and invests the price in a deposit account with a low risk-free rate. However, from the CRR model the put option is possible exercised by the option buyer, and when this situation happens, the chance that the deposit money with risk-free interest cannot cover the payoff is quite high. So the NPI investor will lose an amount of money at the most time in the put option exercise situation, except when the selling price is equal or slightly lower than the CRR price, and this situation is infrequent, and profit is very little. The other time, the NPI investor is not in the game, when the CRR price is lower than the maximum selling price, or the small s results in a negative discount rate.



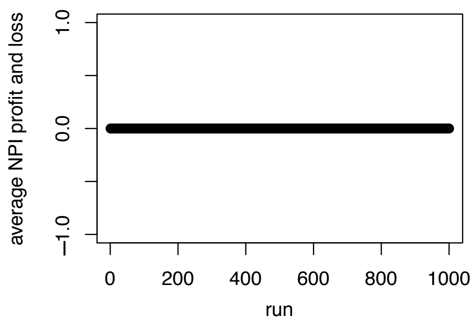
(a) NPI and CRR option prices (call)



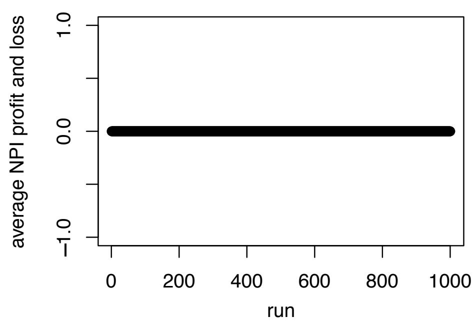
(b) NPI and CRR option prices (put)



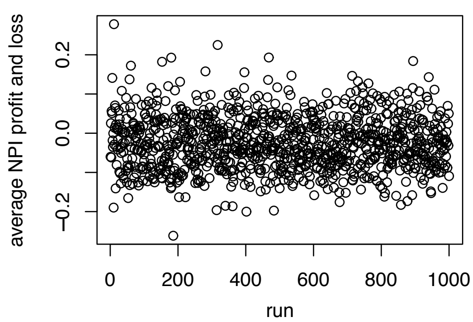
(c) NPI P&L for buying call option



(d) NPI P&L for selling call option



(e) NPI P&L for buying put option



(f) NPI P&L for selling put option

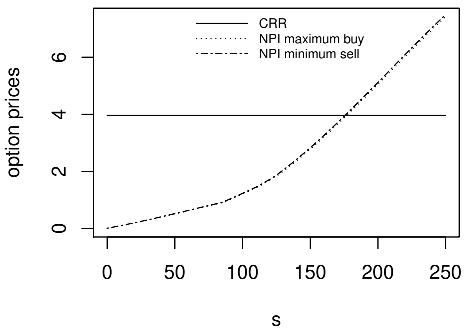
Figure 3.24: NPI profit and loss for Example 3.3.1 in Scenario 2 ($q = 0.5$, $n = 252$, $p = 0.8$)

In terms of Scenario 2, the trading environment gets better for the NPI investor. Even though the NPI investor cannot trade with the CRR investor, selling a call option or buying a put option as in Scenario 1, the reason is different. In this example n is large enough and $s \sim Bin(252, 0.8)$, then $E(s) = 252 \times 0.8 \approx 202$ that is the randomized s expectation and it is the main reason for the absent of selling call and buying put option trade. However, the absence of this two trading position is a good trade movement for the NPI investor, because now the real market probability is 0.8 meaning that the underlying asset price is expected to rise, then neither selling the call option nor buying the put option is a wise trade. The NPI investor gets more involved in the game as a call option buyer and a put option seller. Discussing the results shown in Figures 3.24 (c) and (f) along with the prices comparison in Figures 3.24 (a) and (b) reveal that as a call option buyer, the NPI investor buys the call option at an equal or higher price than the CRR investor but gets a higher payoff calculated based on real market probability p, meaning the CRR price undervalues this call option. At maturity, the NPI investor will get a payoff that is sufficient to cover the price paid. As a put option seller, although the NPI investor sells the put option at a lower price than the CRR price, due to the wrong assumption of p in the CRR model, the payoff the NPI investor needs to pay to the option buyer and the probability of the exercised put option are lower than the CRR prediction. Therefore, the NPI investor either gains some profit when this put option is not exercised or pays less amount of payoff than that in Scenario 1. Generally, in this example, the NPI model performs better than the CRR model in Scenario 2. Owing to the q we set in this example is equal to 0.5, the limitation holding a zero discount rate for the CRR model, we would like to investigate a more general case.

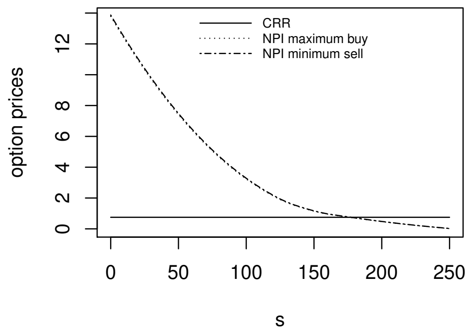
Example 3.3.2

In Scenario 1, as we can see from Figure 3.25, the NPI investor is involved in all four trading actions in this scenario, which are buying and selling both the call and put options. The call option is studied first. Since $p = q = 0.7$, randomized s the expectation of it is 176. As long as randomized s makes the maximum buying price higher than the CRR price, the trade of buying a call option is triggered. As acknowledged, the CRR call option will never be early exercised, so the NPI investor pays a higher or same price and gets the CRR payoff at maturity. In the Figure 3.25 (c), it shows that the NPI investor can earn some profit from buying the call option, meaning all the payoff can cover the price paid. But due to the sufficient number of n in the simulation, this result is not comprehensive. It is possible to encounter a negative value of profit and loss when s is very large, even though this situation is rare.

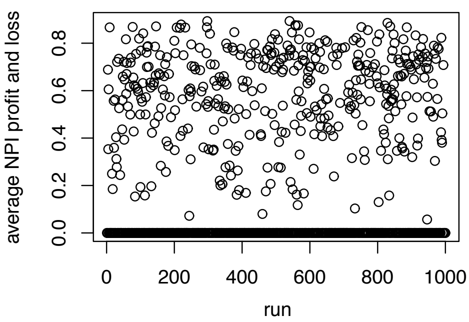
For selling call options, when s makes the minimum buying price lower than the CRR price, and it is higher than 126, making sure a non-negative discount rate for the NPI method, there is a trade of selling the call option. The NPI investor sells a call option at a lower or equal price to the CRR price and saves this money into the bank account until the maturity paying to the CRR investor if the call option is exercised at maturity. From Figure 3.25 (d), it is obvious that the average NPI $P\&L$ for each randomized s is negative, and there is a gap in the loss graph. To explain this gap, we need to tease out the trading procedure of NPI investor trading with the CRR investor at the CRR price, since all other cases cause a higher loss than the bound of the gap. In this trade, with $\overline{V_0} = V_0^{CRR}$, the NPI investor sells this call option and puts the money $\overline{V_0} = V_0^{CRR}$ into the bank account earning the profit from it as $V_0^{CRR} \times 1.002^5 = 4.0038$. However, the NPI investor needs to pay the CRR investor at maturity, for the exercised call option asks a payoff $V_0^{CRR} \times 1.04^5 = 4.8228$, where the discount rate is $r = 0.7 \times 1.1 + 0.3 \times 0.9 - 1 = 0.04$. Then the minimum loss for the NPI call option



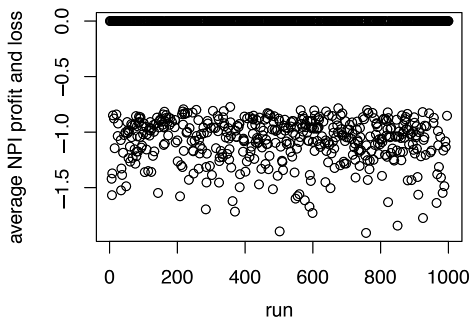
(a) NPI and CRR option prices (call)



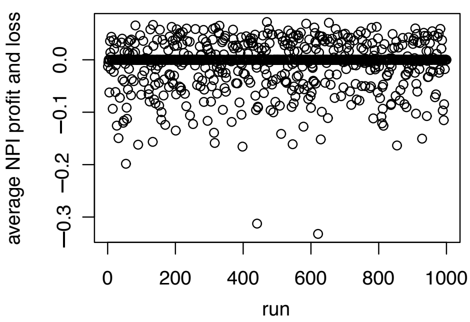
(b) NPI and CRR option prices (put)



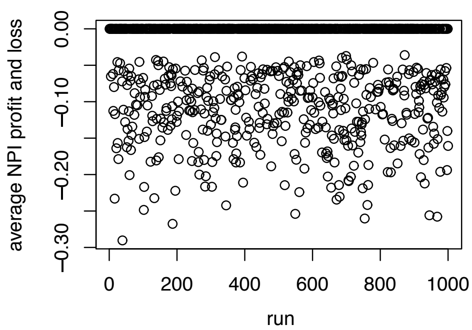
(c) NPI P&L for buying call option



(d) NPI P&L for selling call option



(e) NPI P&L for buying put option



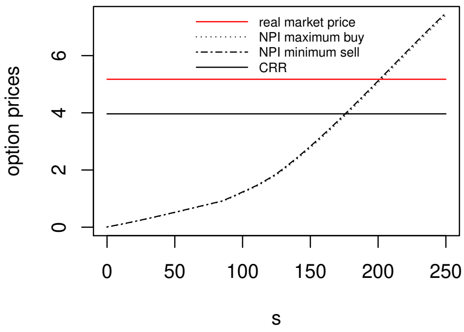
(f) NPI P&L for selling put option

Figure 3.25: NPI profit and loss for Example 3.3.2 in Scenario 1 ($q = 0.7$, $n = 252$, $p = 0.8$)

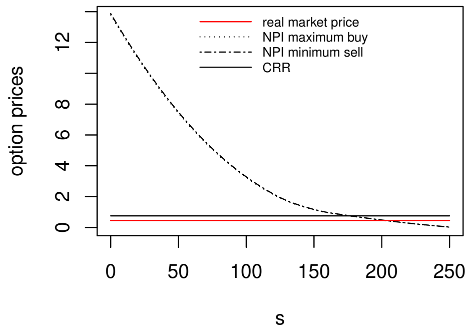
seller is $L = V_0^{CRR} \times 1.04^5 - V_0^{CRR} \times 1.002^5 = 0.8190$.

The NPI investor will face a profit or a loss for buying a put option. As long as randomized s lower than or equal to the intersection between the maximum buying price and the CRR price, and it is higher than 126, there is the action of buying the put option. If the put option is purchased at the highest price when $s = 126$, and the buyer gets the CRR payoff eventually that is not enough to cover the cost. For increasing s, the buying price gets lower. When the payoff acquired by the NPI investor is higher than the price paid, the NPI investor has a profit. Figure 3.25 (f) for selling a put option looks similar to the one for selling a call option, with a smaller variety due to the lower CRR price and NPI minimum selling price. When randomized s leads a lower or equal minimum selling price than the CRR price, the NPI investor sells the put option saving the price in the bank account, and at maturity, it is possible to face amount of payment to the CRR buyer, when this put option is exercised. As the stock price is expected to rise, the exercise of the put option is hard to take place. However, because the CRR investor is using the right probability, the NPI investor sells the put option at an undervalued price, so the payment occurs in some cases.

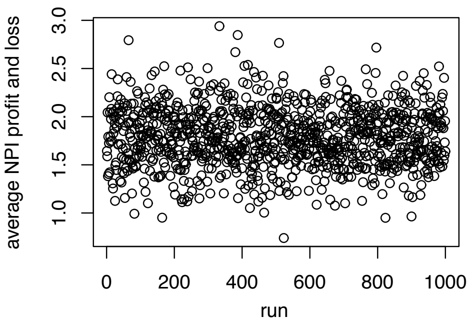
In Scenario 2, consider the values $p = 0.8$, $q = 0.7$, so we simulate $s \sim Bin(252, 0.8)$ with $E(s) = 252 \times 0.8 = 202$. According to the P&L graphs in Figure 3.26, we can be told that in this scenario the NPI investor only plays a role in buying the call option and selling the put option, which is two safe trading positions. For buying the call option, although the NPI investor is paying a higher price than that in Scenario 1, the investor can also get a higher payoff at maturity leading to a positive and higher profit than in Scenario 1. However, from the graph of selling a call option, we can see that there is no profit or loss because in this simulation we randomize 1000 $s \sim Bin(252, 0.8)$. The expectation of s is 202, with which value that all actions of buying the call option occur. As n is large enough to make sure s never leading to a lower minimum selling price than the CRR price,



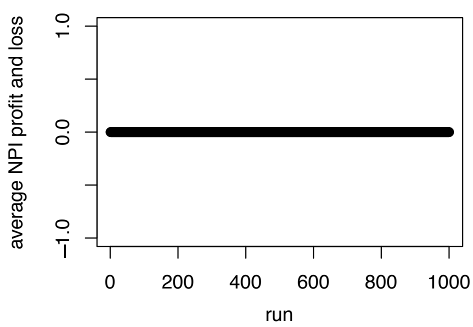
(a) NPI and CRR option prices (call)



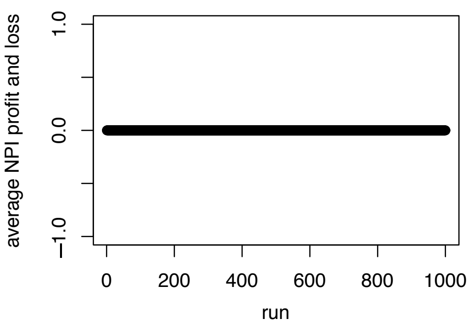
(b) NPI and CRR option prices (put)



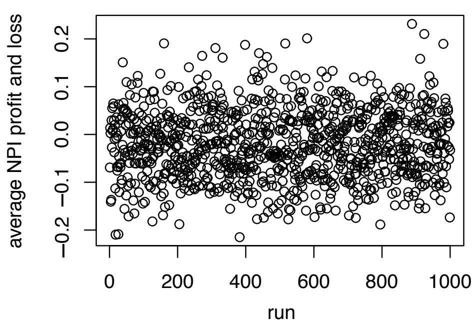
(c) NPI P&L for buying call option



(d) NPI P&L for selling call option



(e) NPI P&L for buying put option



(f) NPI P&L for selling put option

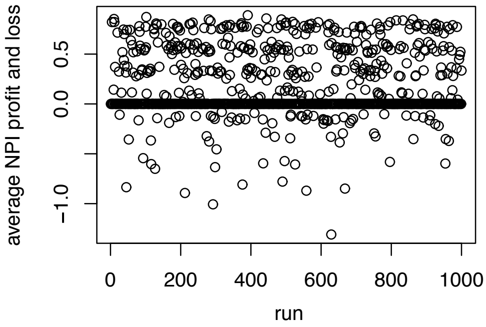
Figure 3.26: NPI profit and loss for Example 3.3.2 in Scenario 2 ($q = 0.7$, $n = 252$, $p = 0.8$)

selling the call option is never going to happen in this simulation. It does not mean selling a call option will not take place in this scenario. When n is not large enough, it can be the case, but it is still a scarce circumstance. We discuss this situation further in the next example. The same situation happens to the case of buying a put option. In our simulation, we cannot observe the case of buying a put option in Figure 3.26 (e), because in this simulation n is large enough to keep the randomized s from being too small to encounter this action. All s for selling a put option causes a successful action with either a profit or loss shown in Figure 3.26 (f). The real market will have a lower payoff for the put option, while the CRR model overvalues it. As the minimum selling price is close to the CRR price, even though the NPI investor sells the put option at a lower or comparable price, the payoff is lower than the CRR expectation, leading to a real profit. As s gets larger, the minimum selling price gets smaller. The option price gained from selling the option and its interests from the risk-free investment cannot cover the payoff at the exercise time. Then there exists a loss but lower than what happened in Scenario 1.

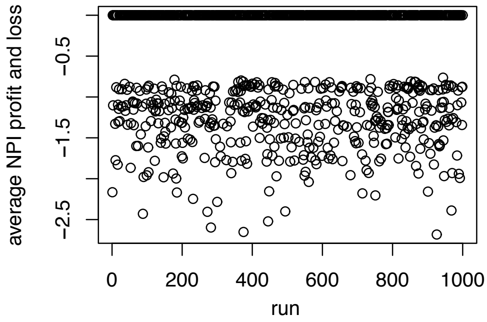
All in all, the NPI method performs better in Scenario 2 than in Scenario 1 in this simulation. First, it keeps the NPI investor away from the less safe trading action, selling the call option and buying the put option. And according to the P&L result of purchasing the call option and selling the put option, the NPI investor can make more profit and lose less in Scenario 2.

Example 3.3.3

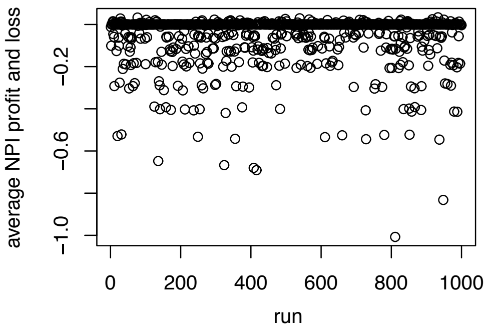
To study the influence of the historical data size, we do another simulation with a smaller amount of historical data, $n = 50$, the randomized $s \sim Bin(50, p)$ displayed in Figures 3.27 and 3.28. Generally, the average NPI profit and loss is the same as that in Example 3.4.2 with $n = 252$. The NPI method performs better than the CRR model in Scenario 2. Figure 3.27 shows that the NPI investor invests in four trade positions. Only buying the call option can offer some profit, while the other



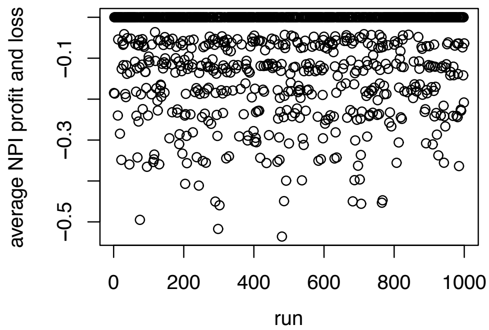
(a) buying the call opiton



(b) selling the call option



(c) buying the put option

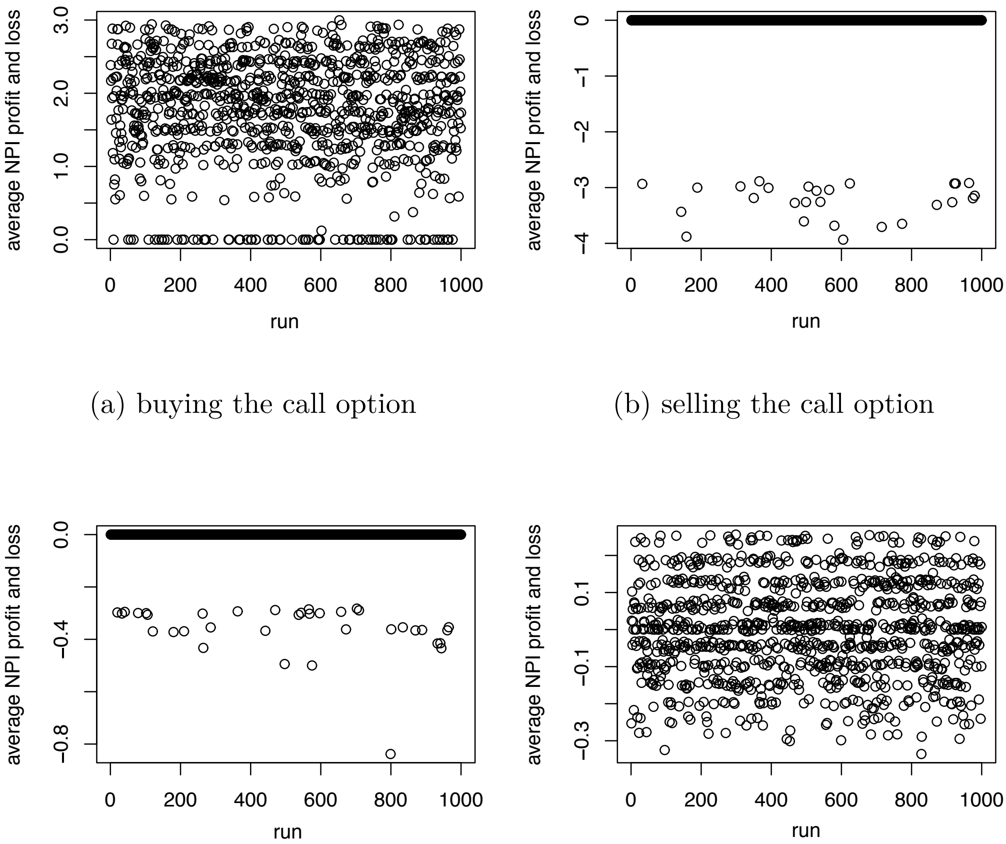


(d) selling the put option

Figure 3.27: NPI profit and loss for Example 3.3.3 in Scenario 1 ($q = 0.7$, $n = 50$, $p = 0.7$)

three positions lead to an amount of loss. From Figure 3.28, even though the NPI investor also puts the money in four positions, the chance of involvement of wrong trade positions, selling a call option and buying a put option, is far less than that shown in Figure 3.27. Also, the profit earned in Figure 3.28 is greater with a higher frequency than that in Figure 3.27. However, smaller n incurs more loss to the NPI investor.

In the first scenario, like a call option buyer, the NPI investor can face some loss. Different from Figure 3.25 (c), Figure 3.27 (a) shows that other than profit and no trade, there exists some loss as well, which confirms what we discussed in Example 3.3.1. n is small in this example, then it is possible to reach the situation that the



(a) buying the call option (b) selling the call option

(c) buying the put option (d) selling the put option

Figure 3.28: NPI profit and loss for Example 3.3.3 in Scenario 2 ($q = 0.7$, $n = 50$, $p = 0.8$)

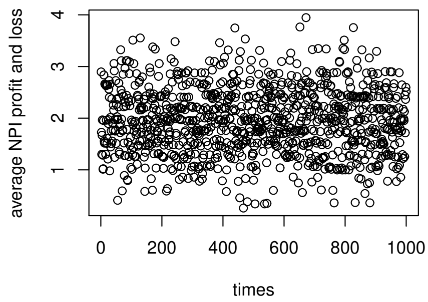
NPI investor pays the price higher than the payoff at maturity. For the same reason, when s is small the NPI investor can sell the call option at a meager price, then it eventually leads greater loss to the NPI investor as shown in Figure 3.27 (b). For the put option, because a smaller n, no matter as a buyer or seller, the NPI investor can encounter an even worse case than that with a larger n.

The same situation also happens to trade in Scenario 2. For the buying call option, the NPI investor will not get involved in the trade when the maximum buying price is lower than the CRR price. As it is shown in Figure 3.28 (a), instead of all cases are ended up with a positive payoff shown in Figure 3.26(c), here some cases hold no profit and loss meaning there is no trade. On the contrary, it is possible to sell the call option causing a loss as shown in Figure 3.28 (b). In Figure 3.28 (c), a smaller historical data also can expose the NPI put option buyer to a loss, as the payoff cannot compensate the price. It is also easier, in this case, to sell the put option at a lower price, which can lead to a more loss to the NPI investor than the loss with a more substantial n. Therefore, sufficient historical data is crucial. To be sure that $\frac{s}{n}$ does not deviate from the real probability in the market a lot, n is supposed to be large enough. Then the prediction from the NPI method is more accurate to guide the investor to a right trading decision with more profit and less loss.

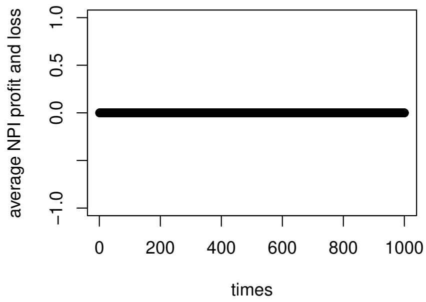
Example 3.3.4

To study the impact of the difference between p and q, here we simulate two trades in Scenario 2 between the NPI investor and the CRR investor based on 50 historical data with the same real market probability $p = 0.8$ but different q, $q = 0.6$ and $q = 0.52$.

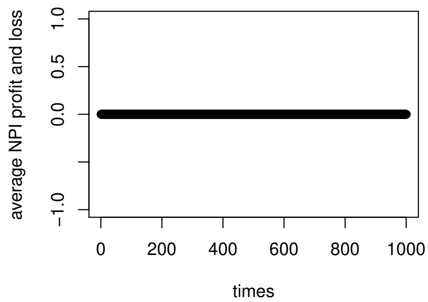
Figure 3.29 is the P&L of the NPI investor in the trade with the CRR investor who uses $q = 0.6$ to make the prediction. The NPI investor quotes at the NPI prices when the trade is happening and gets or pays the real mark payoff when the option holder exercises the option. In this trade, it is clear that the NPI investor takes part in the buying the call option and selling the put option, which are two wise actions we mentioned for the options based on this specific underlying asset. In the buying the call option action, all the trades are taken in action leading to a positive payoff. And the profit outcomes are better than that in the trade with $q = 0.7$ in the last example. When it comes to selling the put option, the NPI investor would encounter profit and loss. Comparing to the results in Figure 3.28 (d) it is not very obvious that the performance in this Example 3.4.3 is better than that in the example with $q = 0.7$. Therefore, we run another simulation, where the CRR investor uses the $q = 0.52$ to make the prediction. Here we simulate the example with $q = 0.52$ rather



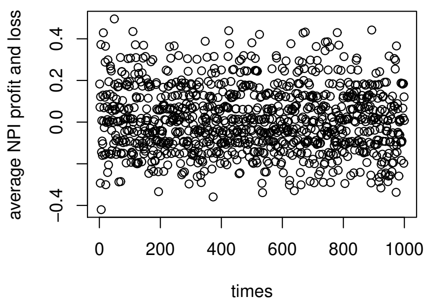
(a) buying the call option



(b) selling the call option



(c) buying the put option



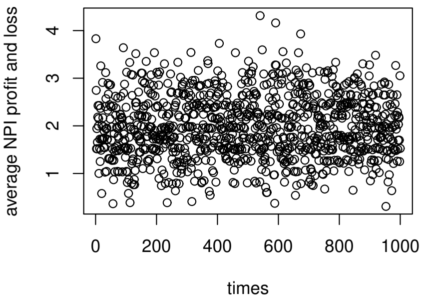
(d) selling the put option

Figure 3.29: NPI profit and loss for Example 3.3.4 in Scenario 2 ($q = 0.6$, $n = 50$, $p = 0.8$)

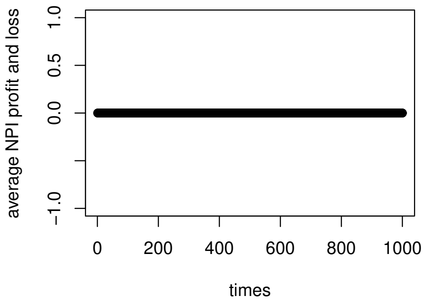
than $q = 0.5$ is to avoid making sure the none negative discount rate assumption validation.

In Figure 3.30, we display the P&L of the NPI investor in the trade with the CRR investor predicting with $q = 0.52$, while the other values stay the same. In this case, the NPI investor also participants in the trade of buying the call option and selling the put option. Be confronted with a larger difference between p and q, the P&L of the NPI investor is better than that with $q = 0.6$. The profit earning from buying the call option is slightly higher. The profit from selling the put option is also greater, and the loss in the trade is lower than those in the example with $q = 0.6$.

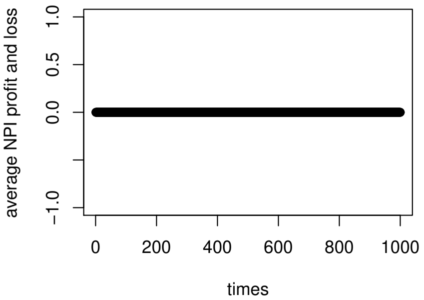
From these simulations, we conclude that insufficient n also exposes the influence of q and p the deviation. The influence reflects in two parts: the first one is that



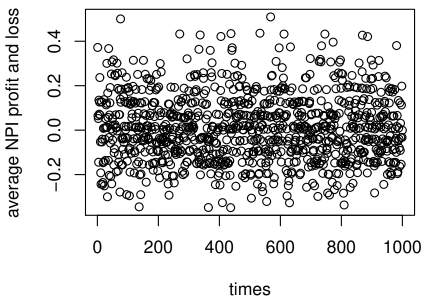
(a) buying the call option



(b) selling the call option



(c) buying the put option



(d) selling the put option

Figure 3.30: NPI profit and loss for Example 3.3.4 in Scenario 2 ($q = 0.52$, $n = 50$, $p = 0.8$)

more q and p deviation prevents the NPI investor getting involved in the unwise the trading action, selling the call option and buying the put option in our example, and also leads more NPI investor's participants in the wise trading action. Another part is that more deviation gives the NPI investor more profits in the trades. This is because the more q and p deviation means the worse the CRR investor's prediction is. In the whole trading process between the NPI investor and the CRR investor, the CRR prices are the criteria to justify whether the NPI buys or sells the option or does nothing. So when n is large enough, no matter how much q deviates from p, the $\frac{s}{n}$ is highly like to be around p, guiding the NPI investor a wise trading decision and expected a profit. However, when n is small, if q and p are different but with slight deviation, the randomized $\frac{s}{n}$ can be further away from p than q, resulting in unwise trading position and a loss. As q gets worse, the chance that the randomized

$\frac{s}{n}$ is closer to the real market p than q is higher, then the NPI investor's P&L is better.

According to the performance study of the NPI method for American option pricing, we can conclude that the NPI method does not perform as good as the CRR model under the assumption that the CRR model is right about the real market. However, the prediction of the NPI method can be improved by having sufficient historical data. When some wrong assumptions are used in the CRR model, the NPI performs better. And sufficient historical data helps the NPI investor make a better decision of more profit. A significant difference between the real market probability and the risk-neutral probability influences the result of the NPI performance, especially when the historical data is insufficient.

3.4 Concluding remarks

The NPI performances comparing to the CRR model performances		
Influencing factor	Scenario 1: The CRR model is right	Scenario 2: The CRR model is wrong
	The NPI investor gets involved in all trading positions and has a high chance to face amount of loss trading with the CRR investor.	The NPI investor trades in the safe and profitable trade positions and has a high opportunity to benefit from the trade with the CRR investor.
Increase the number of historical data	The loss of the NPI investor decreases.	The profit of the NPI investor drops. The NPI investor makes fewer mistakes of choosing wrong trade position.
Enlarge the difference between the risk-neutral probability and the real market probability		The profit of the NPI investor increases. Increase the possibility of the NPI investor trading in the right trade position.

Table 3.2: A summary of the NPI performance results

We developed the NPI method for American option pricing based on the backward optimization method for a binomial tree model. NPI is an imprecise statistical method continuously learning from the data. This property makes the NPI method for American option pricing more close to reality: for the NPI method for American option pricing, we can encounter the situation that the NPI investor would like to exercise an American call option with no dividend early and this also happens in the real market. All conditions to justify whether early exercise or holding further for both call and put options are listed in this chapter. We also studied the average stopping time and option prices comparison between the CRR model and the NPI method by simulation and found that the stopping time of American options is different between the CRR model and the NPI method, which is one of the reasons leading to different expected option prices. Then we illustrated the NPI investor's profit and loss trading with the investor who uses the CRR model in two scenarios. In Scenario 1 the CRR investor uses the right assumption about the future market, and Scenario 2 is under the assumption that the CRR investor does not use the correct assumption. The outcomes show the NPI investor gets a better payoff in Scenario 2 than Scenario 1. This conclusion is displayed in two parts, one that the NPI investor only plays roles in the safer and wiser trade position, the other one that the P&L of the NPI profit is also more optimistic than that in Scenario 1. We also study the influence of the historical data size and the p and q difference and find that the NPI method performs better when the historical data is sufficient or p and q difference is substantial. A summary table of performance study and the influence factors are displayed as Table 3.2.

There still some problems that need to be studied further. The study of the historical data size is necessary, how much historical data is sufficient enough for a relatively accurate prediction. Another challenging future study problem is to apply the NPI method to the real market to see if it can fit in the empirical market.

Chapter 4

NPI for Exotic Option Pricing

After introducing NPI to the European option and the American option, we want to see if the NPI method can be implemented for other complex types of options. The term 'Exotic option' was used by Rubinstein in 1990 [68], which is a long time after the actual product was presented. Distinguishing from vanilla options, the exotic option has flexible and complex trading features to meet the particular demands of clients. Financial engineers add additional exercise conditions to the vanilla options to make it exotic to meet their clients' demands. As a derivative financial product type, more and more new exotic options are produced by financial engineers, like the digital option, the barrier option and the look-back option. In this chapter, we first explain the concept of payoff monotonicity. Then we provide the NPI option pricing methods for three types of exotic options, the digital option in Section 4.2, the barrier option in Section 4.3 and the look-back option in Section 4.4, of which option values can be structured as a binomial tree with monotonic node values. In Section 4.5, we conclude the content of this chapter and discuss future research topics.

4.1 Payoff monotonicity

So far, we have applied the NPI method to vanilla options, the European and American options. There is a common characteristic of all vanilla options, which is a monotonic payoff in the binomial tree. It means that the payoffs of the European and the American options are a monotone function of the number of upward movements. For the European call option, its payoff is $[S_T - K_c]^+$, and for the European put option, its payoff is $[K_p - S_T]^+$. As in the binomial tree, the top node at time T has the largest value of S_T, $[S_T - K_c]^+$ has the largest value for the top node and decreases as the node moves to the bottom of the tree. While $[K_p - S_T]^+$ has the lowest value for the top node at time T and increases as the node moves to the bottom of the tree. For the American options, although for each path in the binomial tree, the exercise time τ can be different, the payoffs at τ are still monotonic. Based on the definition of the American options in Section 1.1, $[S_\tau - K_c]^+$ is the payoff for the American call option, and $[K_p - S_\tau]^+$ is the payoff for the American put option. As τ is the best time to exercise the American option to get the optimal payoff for each path in the binomial tree, $[S_\tau - K_c]^+$ has the largest value for the top node at time τ and decreases as the node move to the bottom of the binomial tree. $[K_p - S_\tau]^+$ has the lowest value for the top node at time τ and increases as the node moves to the bottom of the tree. So the payoff of the American option is also monotonic.

Applying the NPI method to an option with monotonic payoffs is less complicated than to an option with non-monotonic payoffs. For instance, when we want to calculate the upper expected payoff of an option with monotonic payoffs, we can assign the upper probability from Equation (1.17) to each one-time-step path of the upward movement in the binomial tree to get the result. Correspondingly, if the lower expected payoff is needed, we can compute it by assigning the lower probability from Equation (1.16) to each one-time-step path of the upward movement in the binomial tree. However, if the payoffs of an option are not monotonic, the upper and lower expected payoffs cannot be calculated by assigning the upper and lower probability. This does not mean we cannot use the NPI method to derive the lower and upper expected value of non-monotonic payoffs. In Section 4.4, we will introduce a way to manipulate the binomial tree according to the look-back option's definition to get a monotonic binomial tree, or we can use the idea in Chen's thesis [19], that the possible outcomes are listed first, and based on the results, we can assign the most substantial probability mass to the greatest value of interest. In the following sections, we first introduce the NPI method to the options with monotonic payoffs.

4.2 Digital option

The digital option can be dated back to the year 1978 when Beerden and Litzenberger [10] presented a pricing model to evaluate the price giving compensation based on the portfolio price level, which if the portfolio price reaches a certain level the product buyer can get the compensation otherwise he can not. It is the simplest type of exotic option, which is attractive to the market because of its lower contract entrance and lower transaction costs than other types of exotic option[61]. There are two kinds of digital options: all-or-nothing options and asset-or-nothing options. All-or-nothing options give the predetermined amount of money X to the option holder at the maturity if the option is in the money, or nothing if the option expires out of the money or at the money. The all-or-nothing digital option is a noncontinuous payoff option, of which payoff is constant and irrelevant to the underlying asset maturity price. Asset-or-nothing options pay off the underlying asset price S_T at the maturity if the option is in the money, or nothing if the option is out of the money or at the money.

Each kind of digital option can be classified as either a European option or an American option, depending on whether it can be early exercised or not. For an American digital option, the option holder can choose to exercise the option before the maturity and get the predetermined payoff, while for a European option the option holder only has the right to exercise the option at maturity.

Due to its simple feature, the digital option can be priced through both the discrete binomial tree model [30], and the continuous process model, the Black-Scholes model [12]. Thavaneswaran et al. [70] also evaluate the uncertainty of this specific option type in the pricing procedure using fuzzy set theory under the assumption that the market is risk-neutral. We show that the NPI method can also be implemented in digital option pricing.

4.2.1 All-or-nothing option

We start with the evaluation of the European all-or-nothing option. The holder of a European all-or-nothing call option has the right to exercise it at the maturity T, giving the payment X to the holder if it is in the money. To calculate the option price, we need to evaluate the expected payoff at maturity. For a m time-step call option, the option buyer will get a predetermined amount of payment X at the maturity if $S_T > K$ or nothing if $S_T \leq K$, where K is the strike price. Apply the NPI method by assuming that there are n historical data of stock prices involved in the pricing procedure, including s increasing stock prices. The stock price is a Bernoulli quantity that will either go up by the factor u or go down by the factor d. On the basis of the historical information n, we can get the boundary expected payoffs, the upper expected payoff, $\overline{E_c} = X\overline{P}(S_m > K)$, the lower expected payoff, $\underline{E_c} = X\underline{P}(S_m > K)$, of the m time-step call option ended up with stock price S_m. As acknowledged, the condition for the option exercise is $S_m - K = u^{Y(m)}d^{m-Y(m)}S_0 - K_c > 0$, then paths have positive payoffs are ones with the number of upward movements $Y(m) > \frac{\ln K_c - \ln S_0 - m \ln d}{\ln u - \ln d} =: k_c^*$. In Section 1.5, we illustrated the boundary probabilities of event $\{Y(m) \leq k^*|(n,s)\}$ as Equations (1.30) and (1.31), utilizing the conjugacy property, then we can get the NPI upper and lower probabilities for the event $S_m > K$, which are

$$\overline{P}(S_m > K) = \overline{P}(Y(m) > k_c^*) = \binom{n+m}{m}^{-1} \sum_{k=\lfloor k_c^* \rfloor + 1}^{m} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \tag{4.1}$$

$$\underline{P}(S_m > K) = \underline{P}(Y(m) > k_c^*) = \binom{n+m}{m}^{-1} \sum_{k=\lfloor k_c^* \rfloor + 1}^{m} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \tag{4.2}$$

Here k starts from $\lfloor k_c^* \rfloor + 1$ in the binomial tree of the digital option, where $\lfloor k_c^* \rfloor$ is the largest integer equal to or less than k_c^*. Applying the NPI interval probability formulas in Equations (4.1) and (4.2) to the expected payoff calculation leads to the expected option price. Below we list the maximum buying price and the minimum selling price for the European all-or-nothing call option.

The maximum buying price of the call option

$$\underline{V_c} = B(0,m)X \binom{n+m}{m}^{-1} \sum_{k=\lfloor k_c^* \rfloor + 1}^{m} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \tag{4.3}$$

The minimum selling price of the call option

$$\overline{V_c} = B(0,m)X \binom{n+m}{m}^{-1} \sum_{k=\lfloor k_c^* \rfloor + 1}^{m} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \tag{4.4}$$

where $B(0,m)$ is the discount factor between the initial time and the maturity. In Equations (4.3), we assign the maximum probability to small k, while in Equation (4.4), the maximum probability is assigned to large k to get the maximum expected price.

For the m time step put option, the expected value at maturity would be calculated based on the formulas, $\underline{E_p} = X\underline{P}(S_m < K)$ and $\overline{E_p} = X\overline{P}(S_m < K)$, since the put option buyer can get the predetermined amount money X if the maturity stock price is lower than the strike price. Referring to Equations (1.28) and (1.29), the upper and lower probabilities of event $\{Y(m) \geq k^*|(n,s)\}$ are produced. But the paths involved in the pricing process are determined by the condition $S_m - K = u^{Y(m)} d^{m-Y(m)} S_0 - K_c < 0$, then $Y(m) < \frac{\ln K_c - \ln S_0 - m \ln d}{\ln u - \ln d} =: k_p^*$. The upper and lower probabilities of the event $\{Y(m) < k^*|(n,s)\}$ are,

$$\overline{P}(S_m < K) = \overline{P}(Y(m) < k_p^*) = \binom{n+m}{m}^{-1} \sum_{k=0}^{\lceil k_p^* \rceil - 1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \tag{4.5}$$

$$\underline{P}(S_m < K) = \underline{P}(Y(m) < k_p^*) = \binom{n+m}{m}^{-1} \sum_{k=0}^{\lceil k_p^* \rceil - 1} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \tag{4.6}$$

Only paths from 0 to $\lceil k_p^* \rceil - 1$ are taken into account to ensure $Y(m) < k_p^*$, the upper limitation should be either $\lfloor k_p^* \rfloor$ if k_p^* is not a integer or $k_p^* - 1$ if k_p^* is an integer. After getting the expected payoffs and the discounted procedure, we can get pricing option formulae, the maximum buying price and the minimum selling price of the put option shown in Equations (4.7) and (4.8).

The maximum buying price of the put option

$$\underline{V_p} = B(0,m) X \binom{n+m}{m}^{-1} \sum_{k=0}^{\lceil k_p^* \rceil - 1} \binom{s+k}{k} \binom{n-s+m-k-1}{m-k} \tag{4.7}$$

The minimum selling price of the put option

$$\overline{V_p} = B(0,m) X \binom{n+m}{m}^{-1} \sum_{k=0}^{\lceil k_p^* \rceil - 1} \binom{s+k-1}{k} \binom{n-s+m-k}{m-k} \tag{4.8}$$

In Equation (4.7), to get the maximum buying price of the put option, we assign the maximum probability to large k. And to get the minimum selling price of the put option, we assign the minimum probability to large k.

Example 4.1

In this example, we exclude the discount procedure, because we would like to focus on the difference between the two option pricing methods, the CRR model and the NPI method. As the payoff of the digital option is a constant value, the discounted procedure affects the expected profit and loss of the digital option more strongly than the vanilla option.

In this example, we study a digital call option with a strike price $K = 21$, the maturity $m = 4$ and a constant payoff $X = 10$ based on the underlying asset with the initial stock price $S_0 = 20$, upward movement factor $u = 1.1$, and downward movement factor $d = 0.9$. In the CRR model, the prediction of the risk-neutral probability of upward movement is $q = 0.65$, while the NPI method predicts on the basis of $n = 50$ historical data. All these inputs are taken the same value as the corresponding values in the European option examples in Chapter 2 for consistency. Let us look at the difference of the expected price without the discount procedure, namely the expected payoff, between the two methods.

From Figure 4.1 we see that the NPI upper and lower expected payoffs change along with the increasing historical stock price s and intersect the CRR expected payoff. Denote the s value of the intersection between the NPI upper expected payoff and the CRR expected payoff as s_9, and the intersection between the NPI lower expected payoff and the CRR expected payoff as s_{10}. When the s is less than s_9, the NPI upper expected payoff is lower than the CRR expected payoff, so it is the chance to sell this call option. When $s_9 < s < s_{10}$, the CRR expected payoff is in the middle of the NPI upper expected payoff and the NPI lower expected payoff, then there is no willing for trades. In this example, the value of s leaded by $\frac{s}{n} = q = 0.65$, is in this interval. When s is greater than s_{10}, the NPI lower expected payoff is higher than the CRR expected payoff, then the investor who uses the NPI method would like to buy this call option from the CRR investor. To have a more comprehensive concept of the expected payoff, then we varying m and n and plot

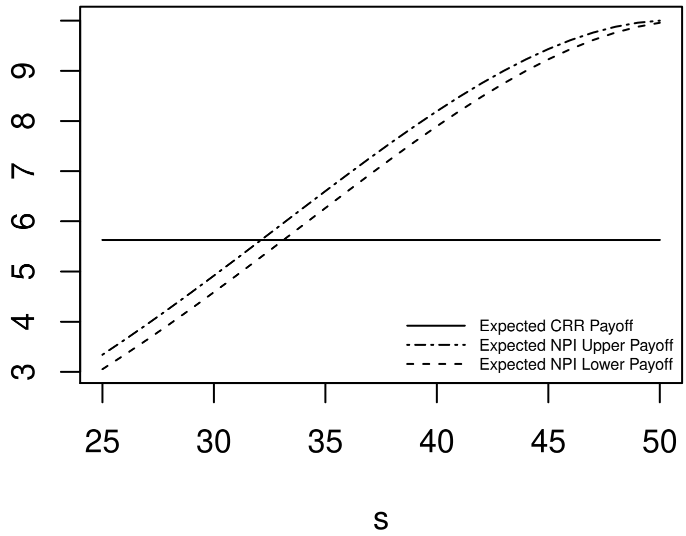
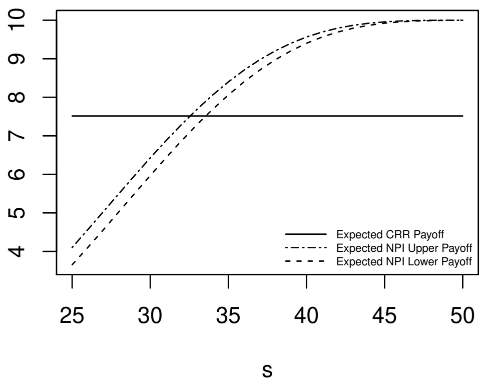


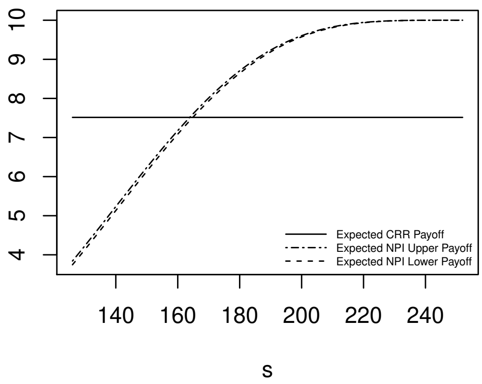
Figure 4.1: Expected payoffs of digital European call option from the NPI method and the CRR model

the expected payoff of different situations in Figure 4.2.

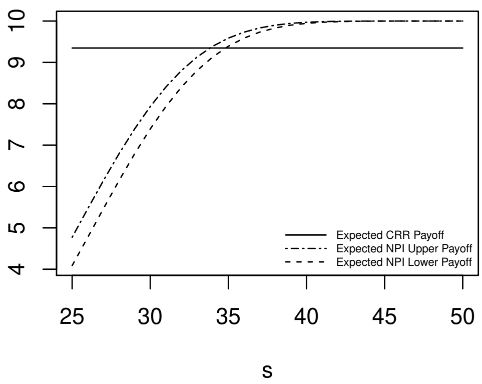
In Figure 4.2, the first column contains three plots of expected payoffs from the NPI method and the CRR model with the same value of $n = 50$ but varying $m = 10$, $m = 30$ and $m = 50$, and the second column contains three expected payoff plots with a fixed $n = 252$ but varying $m = 10$, $m = 30$ and $m = 50$. As we can see from the figure that as m increase that both the CRR and NPI expected payoffs increase as well, the NPI expected payoffs are approaching 10 quicker with larger m. With a larger m, the underlying asset has a higher probability of achieving a higher price, especially for the NPI method. As the NPI method keeps learning from the data, with a larger m, more probabilities are assigned to the upward movement. As each time of upward movement, the NPI method assigns one integer to both the numerator and the dominator of the NPI probability of the upward movement, which is greater than the probability in the last time step. Reflecting in the plots, Figures 4.2 (c) and (e) show when m changes from 30 to 50, the CRR expected



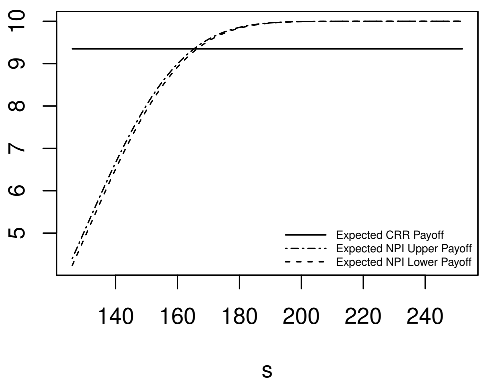
(a) $n = 50$, $m = 10$



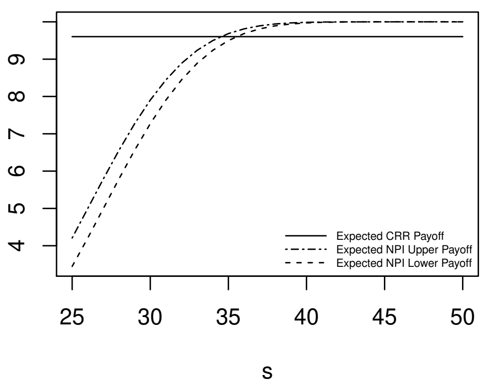
(b) $n = 252$, $m = 10$



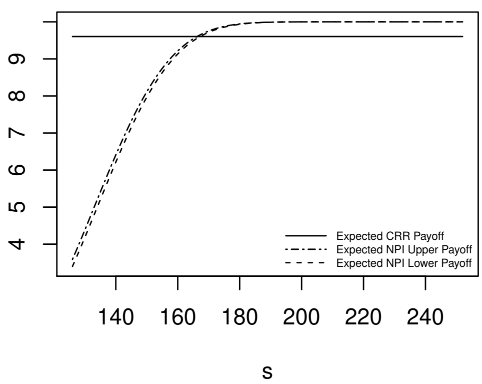
(c) $n = 50$, $m = 30$



(d) $n = 252$, $m = 30$



(e) $n = 50$, $m = 50$



(f) $n = 252$, $m = 50$

Figure 4.2: Expected payoffs of digital European call oprion from the NPI method and the CRR model

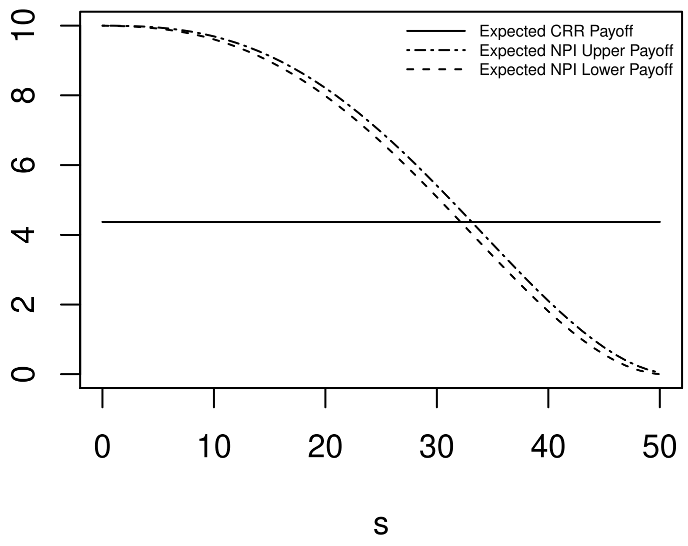


Figure 4.3: Expected payoffs of digital European put option from the NPI method and the CRR model

payoff increases from the value 9.3481 to the value 9.6036, while the NPI expected payoffs do not change the value as dramatically as the CRR expected payoff does. When it comes to the NPI method, there is no big difference between the shapes of expected payoff when $m = 30$ or $m = 50$. Comparing Figures 4.2 (a), (c), (e) to Figures 4.2 (b), (d) and (f), we are told that with larger n the gap between the NPI upper expected payoffs and the NPI lower expected payoffs are narrower than that with smaller n.

We also plot the expected payoffs from two pricing methods for a put option with the same strike price $K_p = 21$ and maturity $m = 4$ based on the same underlying asset as the call option. The results are shown in Figure 4.3. There are two intersections between the NPI expected payoffs and the CRR expected payoff, of which s is denoted as s_{11} for lower NPI expected payoff and s_{12} for upper NPI expected payoff. As s starts from 0, and corresponding expected NPI payoffs begin from 10, the constant payoff value the option holder can gain from a put option that is in the money. As s increases but less than s_{11}, the investor who uses the NPI method is willing to buy the put option from the investor who uses the CRR model to make the prediction. When s is in between s_{11} and s_{12}, there is no trade between these two investors. When s is greater than s_{12}, the NPI investor will sell this put option to the CRR investor. The whole shape of the NPI expected payoffs is decreasing as s increases, approaching to zero with the larger s. We also study the n influence of the expected NPI payoff, showing that for larger n the interval between the upper and lower expected payoff narrows down, and if we increase m, the NPI expected payoff will approach zero as s approaching to 50 or 10 as s approaches 0.

The digital option can have the feature of early exercise, which makes it an American type of all-or-nothing digital option. The option holder can choose to exercise the option anytime earlier than the maturity and get a fixed amount of payment X. To exercise a call option, the spot price S_t needs to be higher than the strike price K for the call option or lower than the strike price K for the put option.

Based on the fixed payment feature of the digital option, the earlier exercise, the better to the option holder [45]. As long as the digital option is exercised, the option holder can get the amount of money X. The earlier the investor receives the money, the earlier the investor can either spend it or invest in another financial product earning more money. Therefore, as an option holder, as long as the spot price is above the strike price, the investor is supposed to exercise it immediately, because the earlier time to get the fixed payoff the better. This exercising action is always triggered no matter what option pricing method is used.

Figure 4.4 displays an example of the American all-or-nothing digital call option through its binomial tree. The nodes in a circle is the one with a spot price higher than the strike price, so in this example the option holder is supposed to exercise this call option either at time 1 when the spot price at time 1 with the paths including S_1 reaching S_1^1 or at time 2 with the path with $S_1 = S_1^2$ and $S_2 = S_2^2$ and earns X as a

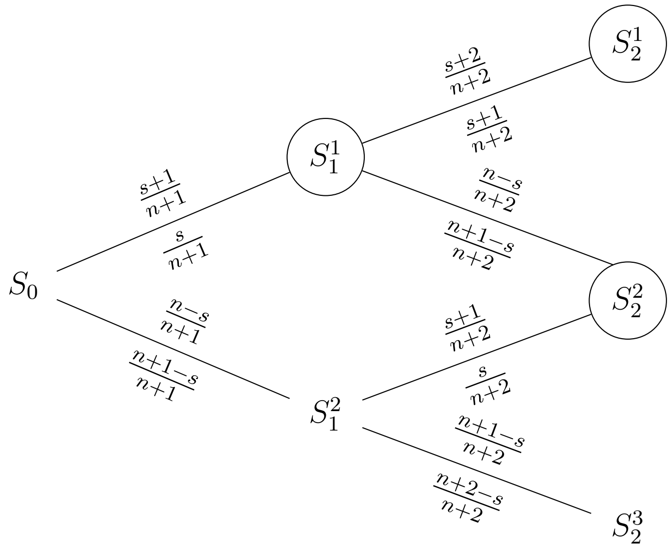


Figure 4.4: The binomial tree of the American all-or-nothing digital option based on the NPI method

payoff. Due to the American call option characteristic, there is no closed formula for the digital option as well. At each node, we justified the condition with an indicator $\mathbb{1}_{\{S_t^i > K\}}$, where t represents the time step, and i represents the case at each time step or the node of each time step in the tree. $\mathbb{1}_{\{S_t^i > K\}}$ is 1 when $S_t^i > K$ otherwise it is 0. We can get the option value at each node V_t^i with $t \in \{0 \dots m\}$, $i \in \{1 \dots t+1\}$ by taking the maximum value between the indicator $\mathbb{1}_{\{S_t^i > K\}}$ at the node to the fixed payoff X and the discounted option expected NPI option value at $t+1$ from this node $B(t, t+1)E(V_{t+1}|S_t = S_t^i)$. As the option is monotonic, the NPI probabilities for each time step is the same as the vanilla American option. According to Equations (1.12) and (1.18), we assign the upper one step probability to the upward movement path to calculate the upper expectation, $\overline{P_t^i}(S_{t+1} = S_t u) = \overline{P_t^i}$, and lower one step probability to the upward movement path when the lower expectation is computed, $\underline{P_t^i}(S_{t+1} = S_t u) = \underline{P_t^i}$. Then the backward pricing method is stated.

So far, we have illustrated the backward evaluation method. Below is the mathematical description of the backward method of the American digital call option to compute the maximum buying price and the minimum selling price.

The maximum buying price of the call option

$$\begin{aligned}
&\underline{V_t^i}_{\{t\in\{0\ldots m-1\},\ i\in\{1\ldots t+1\}\}}\\
&= \max\left\{X\mathbb{1}_{\{S_t^i>K\}}, B(t,t+1)\underline{E}(V_{t+1}|S_t=S_t^i)\right\}\\
&= \max\left\{X\mathbb{1}_{\{S_t^i>K\}}, (1+r)^{-1}\left[\underline{P_t^i}\ \underline{V_{t+1}^i} + (1-\underline{P_t^i})\underline{V_{t+1}^{i+1}}\right]\right\}\\
&= \max\left\{X\mathbb{1}_{\{S_t^i>K\}}, (1+r)^{-1}\left[\frac{s+t-i+1}{n+t+1}\underline{V_{t+1}^i} + \frac{n-s+i}{n+t+1}\underline{V_{t+1}^{i+1}}\right]\right\}\\
&\underline{V_T^i}_{\{T=m,\ i\in\{1\ldots T+1\}\}} = X\mathbb{1}_{\{S_T^i>K\}}
\end{aligned} \tag{4.9}$$

The minimum selling price of the call option

$$\begin{aligned}
&\overline{V_t^i}_{\{t\in\{0\ldots m-1\},\ i\in\{1\ldots t+1\}\}}\\
&= \max\left\{X\mathbb{1}_{\{S_t^i>K\}}, B(t,t+1)\overline{E}(V_{t+1}|S_t=S_t^i)\right\}\\
&= \max\left\{X\mathbb{1}_{\{S_t^i>K\}}, (1+r)^{-1}\left[\overline{P_t^i}\ \overline{V_{t+1}^i} + (1-\overline{P_t^i})\overline{V_{t+1}^{i+1}}\right]\right\}\\
&= \max\left\{X\mathbb{1}_{\{S_t^i>K\}}, (1+r)^{-1}\left[\frac{s+t-i+2}{n+t+1}\overline{V_{t+1}^i} + \frac{n-s+i-1}{n+t+1}\overline{V_{t+1}^{i+1}}\right]\right\}\\
&\overline{V_T^i}_{\{T=m,\ i\in\{1\ldots T+1\}\}} = X\mathbb{1}_{\{S_T^i>K\}}
\end{aligned} \tag{4.10}$$

Here $B(t,t+1)$ is the one time-step discount factor which equals $(1+r)^{-1}$ where r is the discount rate.

The American all-or-nothing put option holder gets the fixed payoff if the spot price S_t is lower than the strike price K. As the call option, the stopping time of the put option is the first time of $S_t < K$, the earliest the best. Then at each node, the indicator $\mathbb{1}_{\{S_t<K\}}$ filters the satisfied nodes. Compare $X\mathbb{1}_{\{S_t^i<K\}}$ to the discounted expected future option value $B(t,t+1)\underline{E}(V_{t+1}|S_t = S_t^i)$ and take the maximum value to be the current option value. In order to calculate the expected future option value, we use the NPI lower expected probabilities at each time step like what we have done in the single American put option evaluation in Section 3.2. For the maximum buying price of the put option, we give the lowest probability to the downward stock movement $\underline{P_t^i}(S_{t+1} = S_t d) = 1 - \overline{P_t^i}$ based on the conjugacy

property. Similarly, we assign the highest probability to the upward stock movement $\overline{P_t^i}(S_{t+1} = S_t d) = 1 - \underline{P_t^i}$ to get its minimum selling price, which has been illustrated in Section 1.4. The mathematical description of this backward option pricing method for the American digital put option is shown below.

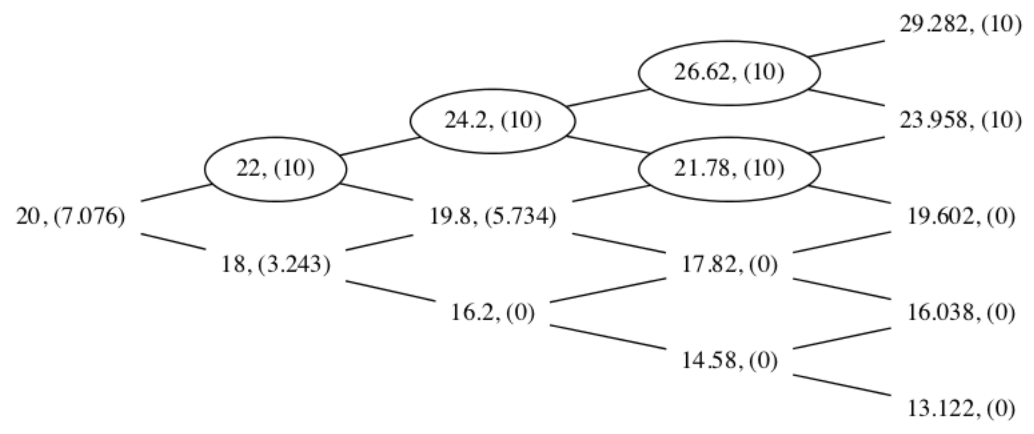
The maximum buying price of the put option

$$
\begin{aligned}
&\underline{V_t^i}_{\{t\in\{0...m-1\}\ i\in\{1...t+1\}\}} \\
&= \max\left\{X\mathbb{1}_{\{S_t^i<K\}}, B(t,t+1)\underline{E}(V_{t+1}|S_t = S_t^i)\right\} \\
&= \max\left\{X\mathbb{1}_{\{S_t^i<K\}}, (1+r)^{-1}\left[\overline{P_t^i}\ \underline{V_{t+1}^i} + (1-\overline{P_t^i})\underline{V_{t+1}^{i+1}}\right]\right\} \\
&= \max\left\{X\mathbb{1}_{\{S_t^i<K\}}, (1+r)^{-1}\left[\frac{s+t-i+2}{n+t+1}\underline{V_{t+1}^i} + \frac{n-s+i-1}{n+t+1}\underline{V_{t+1}^{i+1}}\right]\right\} \\
&\underline{V_T^i}_{\{T=m\ i\in\{1...T+1\}\}} = X\mathbb{1}_{\{S_T^i<K\}}
\end{aligned}
\tag{4.11}
$$

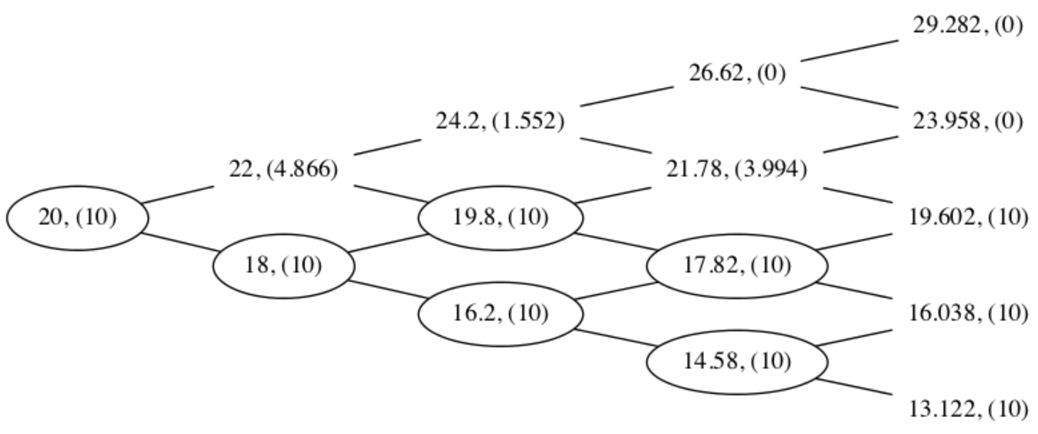
The minimum selling price of the put option

$$
\begin{aligned}
&\overline{V_t^i}_{\{t\in\{0...m-1\}\ i\in\{1...t+1\}\}} \\
&= \max\left\{X\mathbb{1}_{\{S_t^i<K\}}, B(t,t+1)\overline{E}(V_{t+1}|S_t = S_t^i)\right\} \\
&= \max\left\{X\mathbb{1}_{\{S_t^i<K\}}, (1+r)^{-1}\left[\underline{P_t^i}\ \overline{V_{t+1}^i} + (1-\underline{P_t^i})\overline{V_{t+1}^{i+1}}\right]\right\} \\
&= \max\left\{X\mathbb{1}_{\{S_t^i<K\}}, (1+r)^{-1}\left[\frac{s+t-i+1}{n+t+1}\overline{V_{t+1}^i} + \frac{n-s+i}{n+t+1}\overline{V_{t+1}^{i+1}}\right]\right\} \\
&\overline{V_T^i}_{\{T=m\ i\in\{1...T+1\}\}} = X\mathbb{1}_{\{S_T^i<K\}}
\end{aligned}
\tag{4.12}
$$

Referring to the backward pricing method based on the American style digital option, even though there is no closed formula for this type of option, we can still get the expected option price quickly by using of the R software. The R program of the pricing formulae is given in Appendix B.2.



(a) buying a call option



(b) selling a put option

Figure 4.5: American style digital options pricing example

Example 4.2

In this example, we present the results of buying a call option expected price and selling a corresponding put option price in Figure 4.5. At each node of the binomial tree, there are two values, of which the one outside the parenthesis is the asset price, and the other in the parenthesis is option value. The nodes in the oval are the nodes optimal for early exercise. In this example, the stock price starts with 20, and it will either go up by the upward movement factor $u = 1.1$ or down by the downward movement factor $d = 0.9$. For the option, we set the strike price is $K = 21$ and the maturity $m = 4$ for both call and put options. The NPI pricing procedure is based on $n = 50$ historical data and among them the successful data is $s = 30$, and the discount rate is 0.02 calculated by formula $r = \frac{s}{n}u + \frac{n-s}{n}d - 1$. After the evaluation, we know that the NPI investor will buy this call option at a price 7.076, or sell the corresponding put option at a price 10. As we can see from Figure 4.5 (b), this put option is already in the money at the initial time, so it is optimal to be exercised at the initial time. Then, of course, the minimum selling price of this put option is 10. By far, we finish setting up the option pricing method for the all-or-nothing digital option based on NPI. In the next section, we move on to study the other type of digital option, the asset-or-nothing digital option.

4.2.2 Asset-or-nothing option

Different from the all-or-nothing option, this kind of digital options pays the option buyer the underlying asset maturity price S_T rather than a fixed amount of money. The European type of asset-or-nothing option only offers the exercise option at maturity. Therefore, an option holder can either get a payment S_T if $S_T > K$ for call options or $S_T < K$ for put options or nothing if $S_T \leq K$ for call options or $S_T \geq K$ for put options.

The call option is studied at first. By the European option definition, we know that the expected price of this style of option is the discounted expectation of stock price satisfied the condition $S_T > K$. Based on the NPI method, the lower expected value is $\underline{E}(S_T|S_T > K) = S_T\underline{P}(S_T > K)$ and the upper expected value is $\overline{E}(S_T|S_T > K) = S_T\overline{P}(S_T > K)$. For a m-period call option with n historical stock price data among them s historical stock prices go up, we can define the stock price paths holding the positive payoff in the pricing procedure on the basis of the exercise condition, $S_m - K = u^{Y(m)}d^{m-Y(m)}S_0 - K_c > 0 \Leftrightarrow Y(m) > \frac{\ln K_c - \ln S_0 - m\ln d}{\ln u - \ln d} =: k_c^*$. We already know the formulae of $\underline{P}(S_T > K)$ and $\overline{P}(S_T > K)$, Equations (4.1) and (4.2). Thus, if we want to evaluate a call option with the m time step call option ended up with the stock price S_m, the maximum buying and the minimum selling prices of the call option are,

The maximum buying price of the call option

$$\underline{V_c} = B(0,m)\binom{n+m}{m}^{-1}\sum_{k=\lfloor k_c^*\rfloor+1}^{m} S_m\binom{s+k-1}{k}\binom{n-s+m-k}{m-k} \tag{4.13}$$

The minimum selling price of the call option

$$\overline{V_c} = B(0,m)\binom{n+m}{m}^{-1}\sum_{k=\lfloor k_c^*\rfloor+1}^{m} S_m\binom{s+k}{k}\binom{n-s+m-k-1}{m-k} \tag{4.14}$$

where $B(0,m)$ represents the discount factor between the initial time and maturity. To get the maximum buying price, we assign the maximum possible probability to the lowest possible value for k, then the maximum possible remaining probability to the second lowest value for k, and so on. To get the minimum selling price, we assign the maximum possible probability to the greatest value for k, then the maximum possible remaining probability to the second largest value for k, and so on.

The put option holder gets the underlying asset price at maturity S_T as his payment when $S_T < K$, then the expected payoff for him is $E(S_T|S_T < K)$. Based on the NPI method, we can have the interval expected payoffs, with lower expected value $\underline{E}(S_T|S_T < K) = S_T\overline{P}(S_T < K)$ and the upper expected value $\overline{E}(S_T|S_T <$

$K) = S_T \underline{P}(S_T < K)$. In terms of a m period put option, the paths included in the pricing procedure are the same as the all-or-nothing European put option, $Y(m) < \frac{\ln K_c - \ln S_0 - m \ln d}{\ln u - \ln d} =: k_p^*$. Then the formulae the maximum buying price and the minimum selling price of this m time step put option ended up with a maturity stock price S_m are,

The maximum buying price of the put option

$$\underline{V_p} = B(0,m)\binom{n+m}{m}^{-1} \sum_{k=0}^{\lceil k_p^* \rceil - 1} S_m \binom{s+k-1}{k}\binom{n-s+m-k}{m-k} \tag{4.15}$$

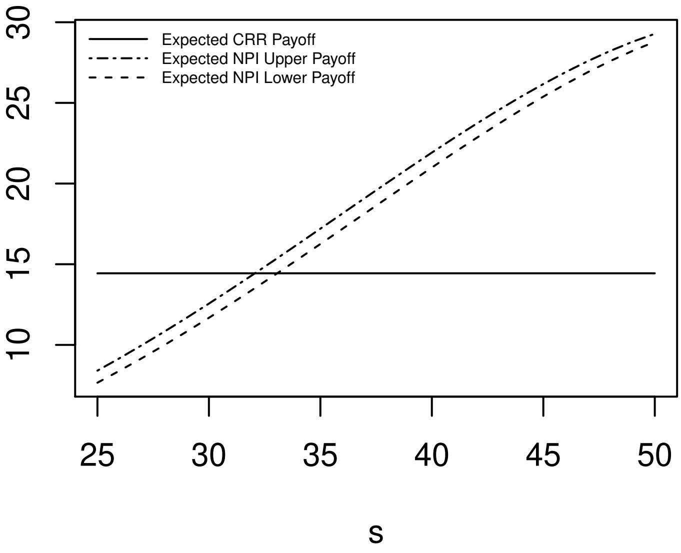
The minimum selling price of the put option

$$\overline{V_p} = B(0,m)\binom{n+m}{m}^{-1} \sum_{k=0}^{\lceil k_p^* \rceil - 1} S_m \binom{s+k}{k}\binom{n-s+m-k-1}{m-k} \tag{4.16}$$

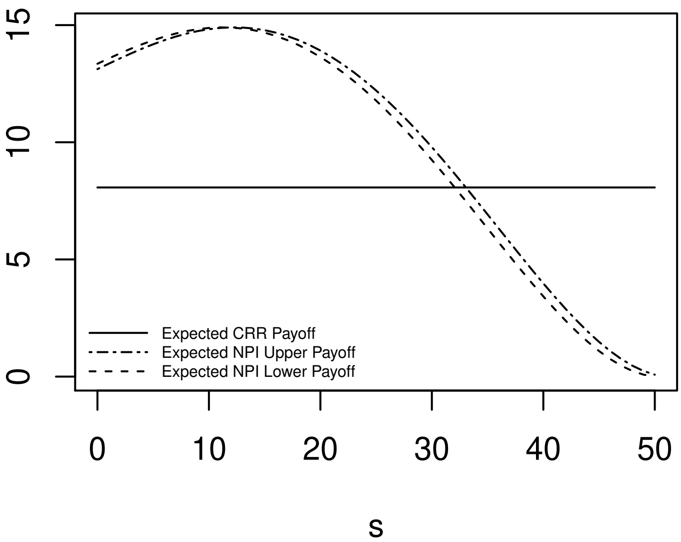
Here to get the upper boundary price, we do the same steps as for the call option, assigning the maximum possible probability to the most considerable value for k, then the maximum possible remaining probability to the second largest value for k, and so on. Also, for the maximum buying price's probability assignment is the same as that for the call option. Below we give an example of the asset-or-nothing option pricing procedure.

Example 4.3

To straightforward see the differences between the NPI method and CRR model, we do not include the discount procedure as well. Both call and put options' expected NPI payoffs and the constant CRR payoffs are plotted in Figure 4.6. The option is also based on the same underlying asset as in Example 4.1, $S_0 = 20$, $u = 1.1$, $d = 0.9$ with the same maturity $m = 4$ and strike price $K = 21$. The CRR model is pricing under the assumption $q = 0.65$, while the NPI method predicts the option price from n historical data with varying s. Figure 4.6 contradistinguishing with



(a) buying a call option



(b) selling a put option

Figure 4.6: American style digital asset-or-nothing options pricing example

Figures 4.1 and 4.3 has an obvious difference that the NPI expected payoffs are not approaching to a constant value because the payment in the asset-or-nothing option is not a constant value anymore. Increasing n can still make the upper and lower interval smaller, since more historical information supports a more accurate result.

After talking about the European option for the asset-or-nothing digital option, the application of the NPI method to the American option is also appealing. The American option gives the holder the right to exercise the option at any time in the option life period when he thinks it is optimal. The payoff is the spot price S_t at the exercise time t. For the call option, the condition to get a positive payoff is $S_t > K$. Due to the early exercise right, there is no closed formula for the option pricing procedure. Instead, we use the backward optimization method comparing the holding value and the instant value of the call option in order to get the initial price. Similar to the all-or-nothing American call option in Section 4.2.1, the indicator for the condition $\mathbb{1}_{\{S_t^i > K\}}$, with $t \in \{0, \ldots, m\}$ and $i \in \{, 1 \ldots, t+1\}$, is used to qualify the instant value of the option S_t^i following the exercise condition. The holding value of the option is the discounted NPI expected value $B(t, t+1)E(V_{t+1}|S_t = S_t^i)$ calculating from the one step binomial tree with initial node S_t^i. Here $B(t, t+1)$ is the one time-step discount factor which equals $(1+r)^{-1}$ where r is the discount rate. Take the greater value to be the current option value V_t^i, $t \in \{0, \ldots, m\}$ $i \in \{1, \ldots, t+1\}$ at time t. After rolling back from the maturity to the initial time, we can get the expected option prices, the maximum buying price, and the minimum selling price.

The maximum buying price of the call option

$$\begin{aligned}
&\underline{V_t^i}_{\{t\in\{0\ldots m-1\},\ i\in\{1\ldots t+1\}\}} \\
&= \max\left\{S_t^i \mathbb{1}_{\{S_t^i>K\}}, B(t,t+1)\underline{E}(V_{t+1}|S_t=S_t^i)\right\} \\
&= \max\left\{S_t^i \mathbb{1}_{\{S_t^i>K\}}, (1+r)^{-1}\left[\underline{P_t^i}\ \underline{V_{t+1}^i} + (1-\underline{P_t^i})\underline{V_{t+1}^{i+1}}\right]\right\} \\
&= \max\left\{S_t^i \mathbb{1}_{\{S_t^i>K\}}, (1+r)^{-1}\left[\frac{s+t-i+1}{n+t+1}\underline{V_{t+1}^i} + \frac{n-s+i}{n+t+1}\underline{V_{t+1}^{i+1}}\right]\right\} \\
&\underline{V_T^i}_{\{T=m,\ i\in\{1\ldots T+1\}\}} = S_T^i \mathbb{1}_{\{S_T^i>K\}}
\end{aligned} \tag{4.17}$$

The minimum selling price of the call option

$$\begin{aligned}
&\overline{V_t^i}_{\{t\in\{0\ldots m-1\},\ i\in\{1\ldots t+1\}\}} \\
&= \max\left\{S_t^i \mathbb{1}_{\{S_t^i>K\}}, B(t,t+1)\overline{E}(V_{t+1}|S_t=S_t^i)\right\} \\
&= \max\left\{S_t^i \mathbb{1}_{\{S_t^i>K\}}, (1+r)^{-1}\left[\overline{P_t^i}\ \overline{V_{t+1}^i} + (1-\overline{P_t^i})\overline{V_{t+1}^{i+1}}\right]\right\} \\
&= \max\left\{S_t^i \mathbb{1}_{\{S_t^i>K\}}, (1+r)^{-1}\left[\frac{s+t-i+2}{n+t+1}\overline{V_{t+1}^i} + \frac{n-s+i-1}{n+t+1}\overline{V_{t+1}^{i+1}}\right]\right\} \\
&\overline{V_T^i}_{\{T=m,\ i\in\{1\ldots T+1\}\}} = S_T^i \mathbb{1}_{\{S_T^i>K\}}
\end{aligned} \tag{4.18}$$

The put option exercise condition is $S_t < K$ by digital option definition, and an option holder can get the spot price S_t as his payment at the expiration t. Then the option value at each node V_t^i $t \in \{0\ldots m\}$ and $i \in \{1\ldots t+1\}$, at time t is the greater value between the instant value $S_t^i \mathbb{1}_{\{S_t^i<K\}}$ and the holding value $B(t,t+1)E(V_{t+1}|S_t = S_t^i)$ gained based on the NPI method. In the tree, it is obvious that the put option instant value is monotone decreasing, the lower path holding a lower payoff. Thus, to calculate the upper holding value, we should assign the upper probability to the upward movement, $B(t,t+1)\overline{E}(V_{t+1}|S_t = S_t^i) = (1+r)^{-1}\left[\overline{P_t^i}\ \overline{V_{t+1}^i} + (1-\overline{P_t^i})\overline{V_{t+1}^{i+1}}\right]$, and to get the lower holding value, we should assign lower probability to the upward movement, $B(t,t+1)\underline{E}(V_{t+1}|S_t = S_t^i) = (1+r)^{-1}\left[\underline{P_t^i}\ \underline{V_{t+1}^i} + (1-\underline{P_t^i})\underline{V_{t+1}^{i+1}}\right]$. Therefore, the maximum buying and the minimum selling prices can be written as following.

The maximum buying price of the put option

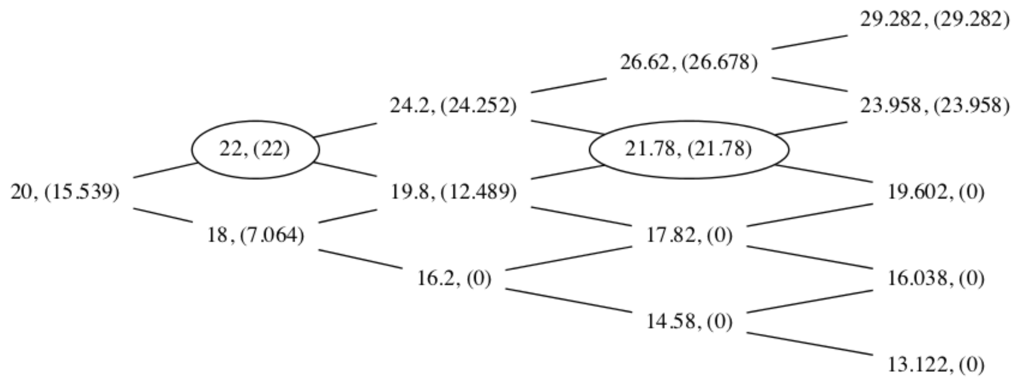
$$
\begin{aligned}
&\underline{V_t^i}_{\{t\in\{0\ldots m-1\},\ i\in\{1\ldots t+1\}\}}\\
&= \max\left\{S_t^i \mathbb{1}_{\{S_t^i<K\}}, B(t,t+1)\underline{E}(V_{t+1}|S_t=S_t^i)\right\}\\
&= \max\left\{S_t^i \mathbb{1}_{\{S_t^i<K\}}, (1+r)^{-1}\left[\underline{P_t^i}\ \underline{V_{t+1}^i}+(1-\underline{P_t^i})\underline{V_{t+1}^{i+1}}\right]\right\}\\
&= \max\left\{S_t^i \mathbb{1}_{\{S_t^i<K\}}, (1+r)^{-1}\left[\frac{s+t-i+1}{n+t+1}\underline{V_{t+1}^i}+\frac{n-s+i}{n+t+1}\underline{V_{t+1}^{i+1}}\right]\right\}\\
&\underline{V_T^i}_{\{T=m,\ i\in\{1\ldots T+1\}\}} = S_T^i \mathbb{1}_{\{S_T^i<K\}}
\end{aligned}
\tag{4.19}
$$

The minimum selling price of the put option

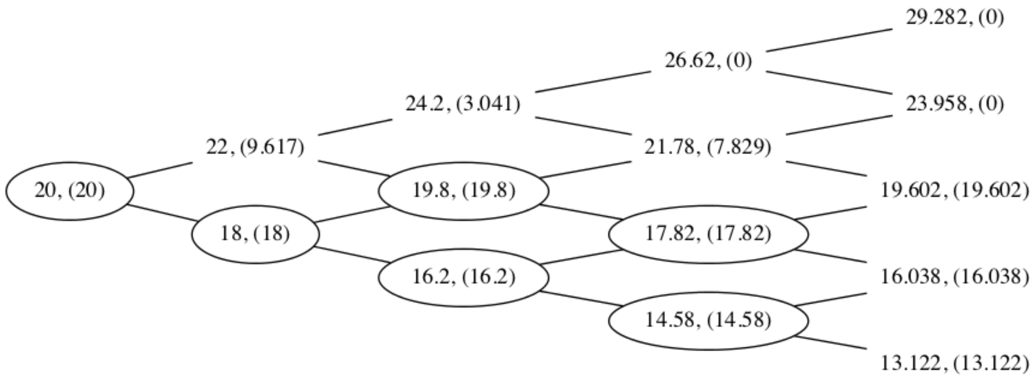
$$
\begin{aligned}
&\overline{V_t^i}_{\{t\in\{0\ldots m-1\},\ i\in\{1\ldots t+1\}\}}\\
&= \max\left\{S_t^i \mathbb{1}_{\{S_t^i<K\}}, B(t,t+1)\overline{E}(V_{t+1}|S_t=S_t^i)\right\}\\
&= \max\left\{S_t^i \mathbb{1}_{\{S_t^i<K\}}, (1+r)^{-1}\left[\overline{P_t^i}\ \overline{V_{t+1}^i}+(1-\overline{P_t^i})\overline{V_{t+1}^{i+1}}\right]\right\}\\
&= \max\left\{S_t^i \mathbb{1}_{\{S_t^i<K\}}, (1+r)^{-1}\left[\frac{s+t-i+2}{n+t+1}\overline{V_{t+1}^i}+\frac{n-s+i-1}{n+t+1}\overline{V_{t+1}^{i+1}}\right]\right\}\\
&\overline{V_T^i}_{\{T=m,\ i\in\{1\ldots T+1\}\}} = S_T^i \mathbb{1}_{\{S_T^i<K\}}
\end{aligned}
\tag{4.20}
$$

The best time to exercise the option is also an important question. In a one-step binomial tree, the criteria to justify if an investor should early exercise the option is the instant value of option S_t should be higher than the holding value $H(t) = B(t,t+1)E(V_{t+1}|S_t=S_t^i) \geq B(t,t+1)S_t(1+r_{t+1}) = B(t,t+1)\left(S_tuP_t+S_td(1-P_t)\right)$ for both upper and lower option values. Here r_{t+1} represents the expected return calculated from the NPI probability P_t, $r_{t+1}+1 = uP_t+d(1-P_t)$. Therefore, we can get the condition for holding an option

$$
\begin{aligned}
H(t) \geq (1+r)^{-1}S_t(1+r_{t+1}) > S_t \Leftrightarrow r_{t+1} > r\\
H(t) \geq (1+r)^{-1}\left(S_tuP_t+S_td(1-P_t)\right) > S_t \Leftrightarrow P_t > \frac{1+r-d}{u-d}
\end{aligned}
\tag{4.21}
$$



(a) buying a call option



(b) selling a put option

Figure 4.7: American style digital asset-or-nothing options pricing example

Now let us look at an example of asset-or-nothing digital option in American style in the R program that is included in Appendix B.3.

Example 4.4

Consider an asset-or-nothing American option is based on a stock with the initial stock price $S_0 = 20$, the upward movement factor $u = 1.1$ and downward movement factor $d = 0.9$. Both call and put options have a strike price $K = 21$ and the maturity $m = 4$. The NPI method does the prediction based on 50 historical data and 30 of them are up raising stock prices. At each node of the tree in Figure 4.7, there are two values, which the stock price is outside the parenthesis and the option value is in the parenthesis. The case of early exercise is disclosed as the node in the oval shape. Therefore, for the NPI call option buyer, his optimal exercise time is either at time 1 when $S_1 = 22$, or at time 3 when $S_1 = 18$ and $S_3 = 21.78$. So this call option buyer expected to buy this option at 15.539 as the maximum. For a NPI put option seller, the investor makes the prediction, and it turns out this put option is optimal to be exercised at the initial time. Therefore, the investor expects to sell this put option at 20, the same value as the stock price in this example.

4.3 Barrier option

As another important type of exotic option, the barrier option has an unique feature distinguishing it from the vanilla option that a barrier of the underlying asset price is predetermined. This barrier of the asset price justifies the option's validation that if the future asset price reaches the barrier, either this option expires or be valid immediately. Merton [57] first presented the down and out option in 1973. There are two classes of the barrier option, ”knock-in” and ”knock-out” barrier options. The ”knock-in” option has a barrier making the option exercisable, while the barrier of the ”knock-out” option causes the expiration of the option. And according to the initial underlying asset price, both ”in” and ”out” options are separated into ”up” and ”down” options. Therefore, there are eight types of barrier options. Cox and Rubinstein [29] illustrated this type of barrier option pricing model based on the CRR model in 1985. Rubinstein and Reiner [69] listed formulae for the eight different barrier options in a continuous time model. Boyle and Lau [17] used the binomial lattices to price the barrier option and try to find its convergence of prices of barrier options. In 1996, Reimer and Sandmann [67] explained the formulae for all types of barrier options including European style and the American style, which are all set up in the risk-neutral world. In 2006, a modified standard binomial method which can price the American type barrier option was introduced by Gaudenzi and Lepellere [35], which is more efficient and can be used in the trinomial method as well. Appolloni et al. [3] explore the binomial lattice method to evaluate the step double barrier options.

If we denote the barrier of asset price as S_b, for the knock-in options, the options are valid when the stock price is less than S_b for the down-and-in option or greater than S_b for the up-and-in option. Here we use the indicator $\mathbb{1}$ to describe the barrier, so for the down-and-out option, the barrier is denoted as $\mathbb{1}_{\{S_t>S_b,t\in(0,\ldots,T)\}}$ and for the up-and-in option, the barrier is denoted as $1-\mathbb{1}_{\{S_t<S_b,t\in(0,\ldots,T)\}}$. According to the payoffs for the call and put options, we can define the knock-in options mathematically as follows.

Knock-in options

$$\text{down-and-in}\begin{cases}[S_T-K_c]^+\left(1-\mathbb{1}_{\{S_t>S_b,t\in(0,\ldots,T)\}}\right), & \text{Call}\\[2ex] [K_p-S_T]^+\left(1-\mathbb{1}_{\{S_t>S_b,t\in(0,\ldots,T)\}}\right), & \text{Put}\end{cases}$$

For a down-and-in option, as long as the stock price during the option valid period S_t goes down and reaches the barrier value S_b, the option holder can get the payoff as $[S_T-K_c]^+$ for the call option or $[K_p-S_T]^+$ for the put option at maturity.

$$\text{up-and-in}\begin{cases}[S_T-K_c]^+\left(1-\mathbb{1}_{\{S_t<S_b,t\in(0,\ldots,T)\}}\right), & \text{Call}\\[2ex] [K_p-S_T]^+\left(1-\mathbb{1}_{\{S_t<S_b,t\in(0,\ldots,T)\}}\right), & \text{Put}\end{cases}$$

Regarding to an up-and-in option, as long as the stock price during the option valid time S_t goes up and reaches the barrier value S_b, the corresponding option is immediately valid and offers the payoff $[S_T-K_c]^+$ for the call option or $[K_p-S_T]^+$ for the put option at maturity.

For the knock-out options, the option is expired once the stock price S_t touch the barrier S_b. Thus, the down-and-out option is valid when $\mathbb{1}_{\{S_t>S_b,t\in(0,\ldots,T)\}}$, and the up-and-out option is valid when $\mathbb{1}_{\{S_t<S_b,t\in(0,\ldots,T)\}}$. So the mathematical formulae of the knock-out options are given below.

Knock-out options

$$\text{down-and-out}\begin{cases}[S_T-K_c]^+\mathbb{1}_{\{S_t>S_b,t\in(0,\ldots,T)\}}, & \text{Call}\\[2ex] [K_p-S_T]^+\mathbb{1}_{\{S_t>S_b,t\in(0,\ldots,T)\}}, & \text{Put}\end{cases}$$

For a down-and-out option, if the stock price during the option validation S_t is always greater than the barrier value S_b, then the option holder can get the payoff as $[S_T - K_c]^+$ for the call option or $[K_p - S_T]^+$ for the put option in the end.

$$\text{up-and-out}\begin{cases}[S_T-K_c]^+\mathbb{1}_{\{S_t<S_b,t\in(0,\ldots,T)\}}, & \text{Call}\\[2ex] [K_p-S_T]^+\mathbb{1}_{\{S_t<S_b,t\in(0,\ldots,T)\}}, & \text{Put}\end{cases}$$

When it comes to an up-and-out option, during the option validation, as long as the stock price S_t always holds a lower value than the barrier value S_b, the option holder can the payoff as $[S_T - K_c]^+$ for the call option or $[K_p - S_T]^+$ for the put option at maturity.

From the definition of the barrier option, we can tell that to evaluate a barrier option we need to monitor the underlying asset regularly during the option life period, and as long as the option reaches the bound either the option is valid or expired. The NPI method can also be applied to this option. For the knock-out type of option, even though there is no closed form formula, we can use the backward valuation method to get the expected option price.

Figure 4.8 displays a knock-and-out call option. As we can see, the payoff is still monotonic with the path, and the probabilities of the NPI boundary prices of the barrier option for each path is still the same as the vanilla options. However, due to the bound S_b the path included in the pricing procedure is reduced, which means that the path having the asset price greater or equal to the S_b are excluded, even though they hold a positive payoff. In this example, only the paths with all solid line in three-time steps are involved in the pricing evaluations.

The details of evaluating this type of exotic option are based on the backward

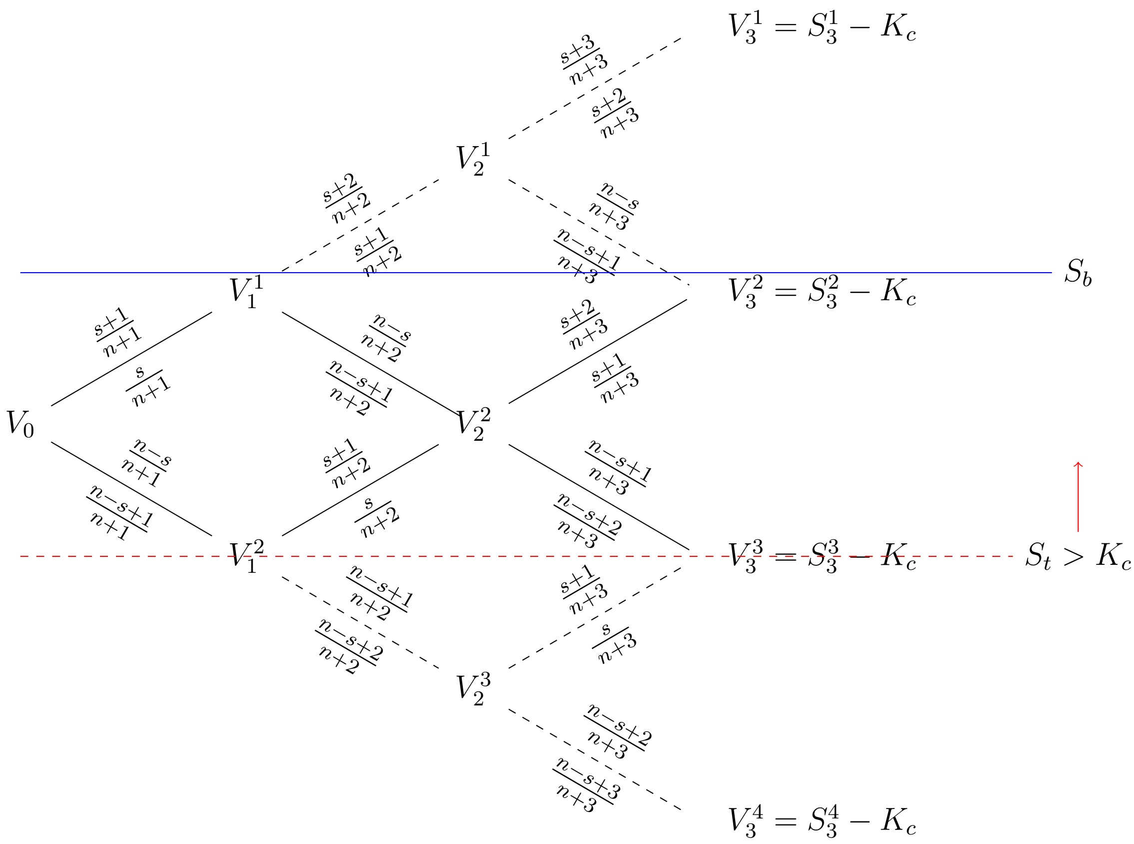


Figure 4.8: The binomial tree based on the NPI method for an up-and-out call option

valuation method. We start from the maturity payoff $[S_T - K_c]^+$ for the call option and $[K_p - S_T]^+$ for the put option, rolling back to the initial time. And at each time step, we check the condition of the "knock-in" or "knock-out" option. For example, there an up-and-out m period call option with the barrier S_b, we can get the maturity payoff at each node i as $[S_T^i - K_c]^+$ with $i \in \{1, \dots, T+1\}$ and $T = m$, and we check the underlying asset price at maturity S_T^i following the condition $S_T^i < S_b$. If not, the option value at that node is immediate equals zero. Thus, the payoff of the whole tree is $V_T^i = [S_T^i - K_c]^+ \mathbb{1}_{\{S_T < S_b\}}$. Then we move to the one time step before the maturity $T-1$, which the option value is the expectation at maturity after the discount procedure if the spot price at T is less than S_b. Otherwise, the option value equals zero, thus, $V_{T-1}^i = B(T-1, T)[S_T^i - K_c]^+ \mathbb{1}_{\{S_T < S_b\}} \mathbb{1}_{\{S_{T-1} < S_b\}}$. According to the NPI method, we can get the upper and lower expectations based on n historical stock price data with s increased prices, then these values lead us to two boundaries of the option value at time $T-1$. After doing the same procedure at every time step, we can get two initial boundary option values, named as the maximum buying price and the minimum selling price. For an up-and-out call option, the mathematical description is listed below.

The maximum buying price of the call option

$$\begin{aligned}
\underline{V}^i_{t\{t\in\{0\ldots m-1\}\ i\in\{1\ldots t+1\}\}} &= B(t,t+1)\left[\underline{P}^i_t\ \underline{V}^i_{t+1} + (1-\underline{P}^i_t)\underline{V}^{i+1}_{t+1}\right]\mathbb{1}_{\{S^i_t<S_b\}} \\
&= (1+r)^{-1}\left[\frac{s+t-i+1}{n+t+1}\underline{V}^i_{t+1} + \frac{n-s+i}{n+t+1}\underline{V}^{i+1}_{t+1}\right]\mathbb{1}_{\{S^i_t<S_b\}} \\
\underline{V}^i_{T\{T=m\ i\in\{1\ldots T+1\}\}} &= [S^i_T - K_c]^+\mathbb{1}_{\{S^i_T<S_b\}}
\end{aligned} \tag{4.22}$$

The minimum selling price of the call option

$$\begin{aligned}
\overline{V}^i_{t\{t\in\{0\ldots m-1\}\ i\in\{1\ldots t+1\}\}} &= B(t,t+1)\left[\overline{P}^i_t\ \overline{V}^i_{t+1} + (1-\overline{P}^i_t)\overline{V}^{i+1}_{t+1}\right]\mathbb{1}_{\{S^i_t<S_b\}} \\
&= (1+r)^{-1}\left[\frac{s+t-i+2}{n+t+1}\overline{V}^i_{t+1} + \frac{n-s+i-1}{n+t+1}\overline{V}^{i+1}_{t+1}\right]\mathbb{1}_{\{S^i_t<S_b\}} \\
\overline{V}^i_{T\{T=m\ i\in\{1\ldots T+1\}\}} &= [S^i_T - K_c]^+\mathbb{1}_{\{S^i_T<S_b\}}
\end{aligned} \tag{4.23}$$

For an up-and-out put option, the probability assignment is the same as what we have done to the vanilla American put option in Chapter 3. Then we control the option's validation by the barrier of the asset price $S^i_t < S_b$. There are no closed formulae for the put option as well. The mathematical description of the backward method can be written as follows.

The maximum buying price of the put option

$$\begin{aligned}
\underline{V}^i_{t\{t\in\{0\ldots m-1\}\ i\in\{1\ldots t+1\}\}} &= B(t,t+1)\left[\overline{P}^i_t\ \underline{V}^i_{t+1} + (1-\overline{P}^i_t)\underline{V}^{i+1}_{t+1}\right]\mathbb{1}_{\{S^i_t<S_b\}} \\
&= (1+r)^{-1}\left[\frac{s+t-i+2}{n+t+1}\underline{V}^i_{t+1} + \frac{n-s+i-1}{n+t+1}\underline{V}^{i+1}_{t+1}\right]\mathbb{1}_{\{S^i_t<S_b\}} \\
\underline{V}^i_{T\{T=m\ i\in\{1\ldots T+1\}\}} &= [K_p - S^i_T]^+\mathbb{1}_{\{S^i_T<S_b\}}
\end{aligned} \tag{4.24}$$

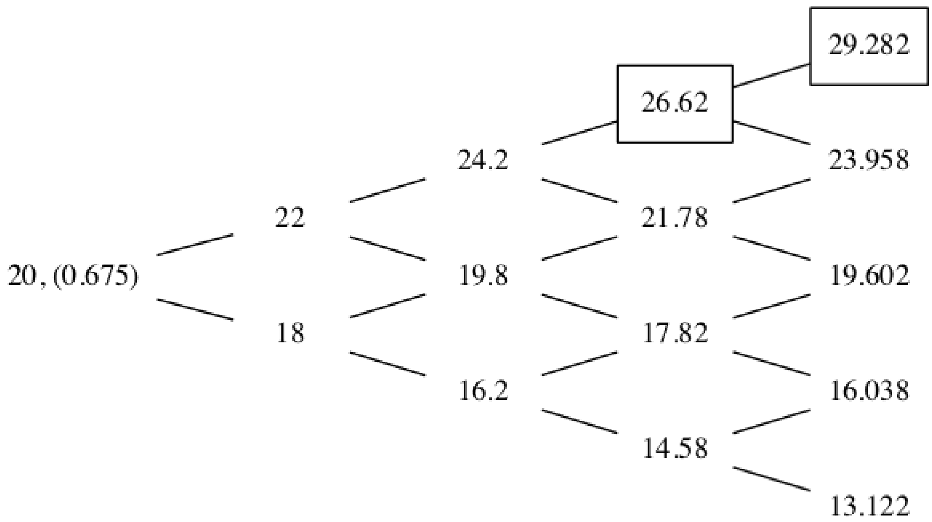


Figure 4.9: The binomial tree of an up-and-out call option

The minimum selling price of the put option

$$
\begin{aligned}
\overline{V_t^i}_{\{t\in\{0\ldots m-1\}\ i\in\{1\ldots t+1\}\}} &= B(t,t+1)\left[\underline{P_t^i}\ \overline{V_{t+1}^i} + (1-\underline{P_t^i})\overline{V_{t+1}^{i+1}}\right]\mathbb{1}_{\{S_t^i<S_b\}} \\
&= (1+r)^{-1}\left[\frac{s+t-i+1}{n+t+1}\overline{V_{t+1}^i} + \frac{n-s+i}{n+t+1}\overline{V_{t+1}^{i+1}}\right]\mathbb{1}_{\{S_t^i<S_b\}} \\
\overline{V_T^i}_{\{T=m\ i\in\{1\ldots T+1\}\}} &= [K_p - S_T^i]^+\mathbb{1}_{\{S_T^i<S_b\}} \qquad (4.25)
\end{aligned}
$$

When it comes to the down-and-out options, only the barrier of asset changes to $S_t^i > S_b$, other than that, the probability assignment and payoff are the same as the up-and-out barrier options.

Example 4.5

By using R program in Appendix B.4, here we predict an up-and-out call option with the strike price $K = 21$ based on $n = 50$ and $s = 30$ historical data. In this example, the underlying asset with an initial price $S_0 = 20$ has a barrier $S_b = 26$. Then any path reaches the barrier of the asset price is not included in the option evaluation. As the stock price is a Bernoulli random quantity, either up with the factor $u = 1.1$ or down with the factor $d = 0.9$, the asset price at each node in the binomial tree is determined. In Figure 4.9, the two nodes higher than the barrier

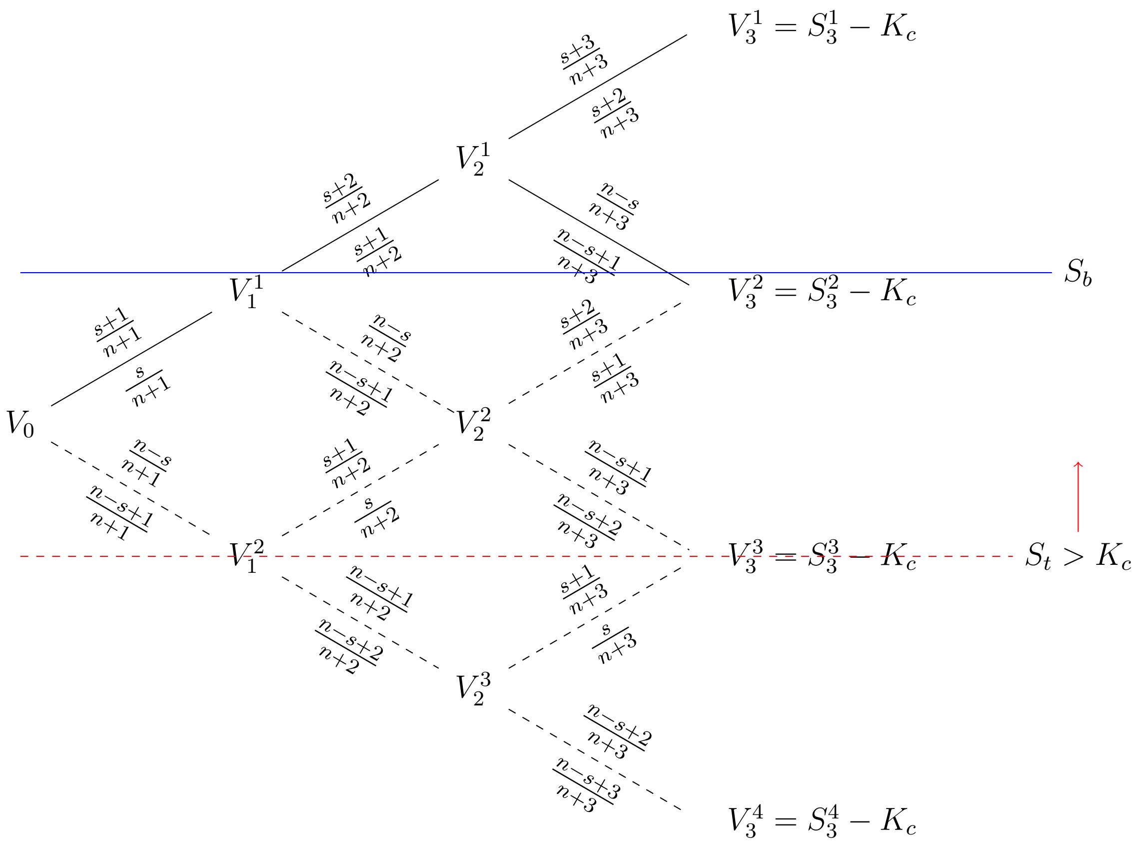


Figure 4.10: The binomial tree based on the NPI method for an up-and-in call option

are in the boxes, which are $S_3^1 = 26.62$ and $S_4^1 = 29.282$. So the paths have the nodes in the boxes not involved in the evaluation holding the value zero. Then we can get the option value at every node of the binomial tree. Here the discount rate is a constant value equal to $\frac{s}{n} = \frac{30}{50} = 0.02$. After doing the backward evaluation until the initial time, we get the expected price of this up-and-out barrier option shown as the value 0.675 in the parenthesis in Figure 4.9.

Unlike the knock-out option, the knock-in option cannot simply use the backward optimization method to make the prediction. It is easier to illustrate this in an example. Figure 4.10 shows an up-and-in call option with the barrier of the asset price S_b. Based on the definition of the up-and-in call option, as S_2^1 and S_3^1 are higher than the barrier price S_b, the only two paths in the evaluation are $V_0 \rightarrow V_1^1 \rightarrow V_2^1 \rightarrow V_3^1$ and $V_0 \rightarrow V_1^1 \rightarrow V_2^1 \rightarrow V_3^2$. However, the backward method cannot be used to do the evaluation. Unlike the knock-out barrier option, we need to know the valid path before making the prediction. If not, for example, the paths ended up with V_3^2 are $V_0 \to V_1^1 \to V_2^1 \to V_3^2$ which is valid, and $V_0 \to V_1^1 \to V_2^2 \to V_3^2$ which is invalid. But if we use backward method directly, as V_3^2 has a positive payoff, for the path $V_0 \to V_1^1 \to V_2^2 \to V_3^2$, after discounted V_2^2 should have a positive payoff as well. Doing the same step rolling back procedure until the initial time, shows clearly that this path is not included in the evaluation. Then all the works above are just a waste of time.

If at time t the stock price S_t^i is the first node of each path from the initial time qualified with the barrier condition, then we can see option value at this node as a vanilla European option with the same strike price but different maturity $T - t$. After getting all the option value at every first valid node in the tree, then we use the backward evaluation method to roll back to the initial time and get the expected option price. For the call option listed in Figure 4.10, as V_2^1 is the first node that higher than S_b then, we see it as a one-step European call option with the initial stock price S_2^1. Using Equation (2.14) for buying position and Equation (2.15) for selling equation, we can get the option value at node V_2^1. One thing we need to pay attention that for this European option the historical data is $n + t$, and the successful historical data is $s + t - i + 1$. Then apply the backward valuation method, then we obtain the expected value V_0. This pricing procedure can be described mathematically as follows.

The maximum buying price for the call option

$$\underline{V_t^i}_{\{t\in\{0\dots m-1\}\ i\in\{1\dots t+1\}\}} = \begin{cases} B(t,t+1)\left[\underline{P_t^i}\ \underline{V_{t+1}^i} + (1-\underline{P_t^i})\underline{V_{t+1}^{i+1}}\right] \\ \text{If } S_t^i < S_b \text{ and } S_{t-1}^i < S_b \\ B(t,T)\binom{m+n+t}{m}^{-1}\sum_{k=\lceil k_c^*\rceil}^{m-t}[u^k d^{m-t-k} S_t^i - K_c] \\ \times\binom{s+t-i+k}{k}\binom{n-s-t+i+m-k-1}{m-k} \text{If } S_t^i \geq S_b \text{ and } S_{t-1}^i < S_b \end{cases} \tag{4.26}$$

$$\underline{V_T^i}_{\{T=m\ i\in\{1\dots T+1\}\}} = [S_T^i - K_c]^+\left(1 - \mathbb{1}_{\{S_t<S_b, t\in(0,\dots,T)\}}\right) \tag{4.27}$$

The minimum selling price for the call option

$$\overline{V_t^i}_{\{t\in\{0\dots m-1\}\ i\in\{1\dots t+1\}\}} = \begin{cases} B(t,t+1)\left[\overline{P_t^i}\ \overline{V_{t+1}^i} + (1-\overline{P_t^i})\overline{V_{t+1}^{i+1}}\right] \\ \text{If } S_t^i < S_b\text{and } S_{t-1}^i < S_b \\ B(t,T)\binom{m+n+t}{m}^{-1}\sum_{k=\lceil k_c^*\rceil}^{m-t}[u^k d^{m-t-k} S_t^i - K_c] \\ \times\binom{s+t-i+k+1}{k}\binom{n-s-t+i+m-k-2}{m-k} \text{If } S_t^i \geq S_b \text{ and } S_{t-1}^i < S_b \end{cases} \tag{4.28}$$

$$\overline{V_T^i}_{\{T=m\ i\in\{1\dots T+1\}\}} = [S_T^i - K_c]^+\left(1 - \mathbb{1}_{\{S_t<S_b, t\in(0,\dots,T)\}}\right) \tag{4.29}$$

For an up-and-in put option, the option maturity value is $[K_p - S_T^i]^+ \times \left(1 - \mathbb{1}_{\{S_t<S_b, t\in(0,\dots,T)\}}\right)$. Otherwise, the option value is either the discounted value rolling back and discounted from the next time steps option value or equal to the European put option with the maturity $T-t$ calculated based on $n+t$ historical data among them $s+t-i+1$ are the raised stock price. The mathematical description to calculate the maximum buying price and the minimum selling price is listed below.

The maximum buying price for the put option

$$\underline{V_t^i}_{\{t\in\{0\ldots m-1\}\ i\in\{1\ldots t+1\}\}} = \begin{cases} B(t,t+1)\left[\overline{P_t^i}\ \underline{V_{t+1}^i} + (1-\overline{P_t^i})\underline{V_{t+1}^{i+1}}\right] \\ \text{If } S_t^i < S_b \text{ and } S_{t-1}^i < S_b \\ B(t,T)\binom{m+n+t}{m}^{-1}\sum_{k=\lceil k_c^*\rceil}^{m-t}[K_p - u^k d^{m-t-k} S_t^i] \\ \times\binom{s+t-i+k+1}{k}\binom{n-s-t+i+m-k-2}{m-k}\text{If } S_t^i \geq S_b \text{ and } S_{t-1}^i < S_b \end{cases} \tag{4.30}$$

$$\underline{V_T^i}_{\{T=m\ i\in\{1\ldots T+1\}\}} = [K_p - S_T^i]^+\left(1 - \mathbb{1}_{\{S_t<S_b,t\in(0,\ldots,T)\}}\right) \tag{4.31}$$

The minimum selling price for the put option

$$\overline{V_t^i}_{\{t\in\{0\ldots m-1\}\ i\in\{1\ldots t+1\}\}} = \begin{cases} B(t,t+1)\left[\underline{P_t^i}\ \overline{V_{t+1}^i} + (1-\underline{P_t^i})\overline{V_{t+1}^{i+1}}\right] \\ \text{If } S_t^i < S_b \text{ and } S_{t-1}^i < S_b \\ B(t,T)\binom{m+n+t}{m}^{-1}\sum_{k=\lceil k_c^*\rceil}^{m-t}[K_p - u^k d^{m-t-k} S_t^i] \\ \times\binom{s+t-i+k}{k}\binom{n-s-t+i+m-k-1}{m-k}\text{If } S_t^i \geq S_b \text{ and } S_{t-1}^i < S_b \end{cases} \tag{4.32}$$

$$\overline{V_T^i}_{\{T=m\ i\in\{1\ldots T+1\}\}} = [K_p - S_T^i]^+\left(1 - \mathbb{1}_{\{S_t<S_b,t\in(0,\ldots,T)\}}\right) \tag{4.33}$$

To price the down-an-in barrier option, we change the barrier of the underlying asset price to the first underlying asset lower or equal to the barrier at time t and $\left(1 - \mathbb{1}_{\{S_t>S_b,t\in(0,\ldots,T)\}}\right)$ at maturity.

Example 4.6

Example 4.6 is an up-and-in call option in buying position based on the same underlying asset. The barrier of the underlying asset is $S_b = 23$, so any path contains asset price higher or equal to 23 are included in the pricing process. In Figure 4.11, there is the binomial tree of this option. The nodes in the box are the two cases that the underlying asset price first over the barrier. Let us look at first node $S_2^1 = 24.2$. When the underlying asset price encounters this price, then the paths have this

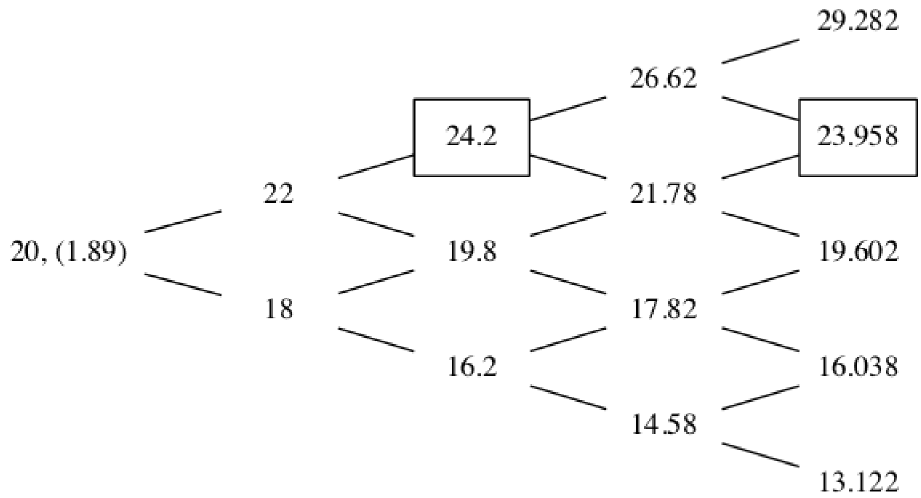


Figure 4.11: The binomial tree of an up-and-in call option

value included in the option price evaluation. So as we described the up-and-in option evaluation, we first compute the option value at this node by seeing it as a vanilla European option with the initial stock price $S_0 = 24.2$ and maturity $m = 2$. Then we roll this option value back to the initial time. The second node in the box is $S_4^2 = 23.958$. This is a maturity node, so we use the backward method to get the initial expected value. However, we would like to highlight one point that as the $S_4^2 = 23.958$ is also included in the paths containing $S_2^1 = 18$, so the only backward procedure of the paths that have S_4^2 as the first node over the barrier are $20 \to 22 \to 19.8 \to 21.78 \to 23.958$ and $20 \to 18 \to 19.8 \to 21.78 \to 23.958$. After pricing, the maximum buying price of this up-and-in barrier call option is 1.89 shown in the parenthesis in Figure 4.11.

4.4 Look-back option

We have implemented the NPI method to two relatively less complicated type of exotic options, the digital option, and the barrier option. In this section, the NPI method's application to the look-back option is presented.

'Look-back option' as one of the exotic option is introduced by Goldman, Sosin and Gatto [39]. The look-back option is classed into two types: the look-back option with the fixed strike price and the look-back option with the float strike price. The option with fixed strike price entitles the option holder to get the payoff as the difference between the maximum underlying asset price over the observation option period and the strike price for the call option $\max_{0\leq i\leq m} S(t_i) - K_c$ or the positive value of the strike price minus the minimum underlying asset price during this period for the put option $K_p - \min_{0\leq i\leq m} S(t_i)$. The one with float strike price gives the option holder the right of buying the underlying asset at the minimum underlying asset price during the option life period $\min_{0\leq i\leq m} S(t_i)$ to the call option holder or selling the underlying asset at its maximum price during this period $\max_{0\leq i\leq m} S(t_i)$. Goldman, Sosin, and Gatto [39] provided the pricing method based on the Brownian motion when they first presented this type of option. The CRR model can be used in the look-back option as well. Hull and White [44] elaborated the path-dependent option evaluation based on the binomial tree in 1993. In the same year, Amin[2] considered the generalization of the CRR model to make it suitable for path-dependent options' evaluation by adding a jump-diffusion process. Kima, Park and Qian [51] derived a binomial tree model with jump diffusion specific for the look-back option. Babbs [7] monitored the look-back option with a discrete time scheme instead of the continuous monitor based on the binomial tree. Park [62] also explored a binomial tree model with double-exponential jumps and studied its convergence. According to the definition of the look-back option with the fixed or float strike price, the payoff of the look-back option is pretty clear. However, the payoff of the look-back option can be not monotone with the path structure anymore. Let us look at the tree example in Figure 4.12 to explain the monotonicity of the option.

In Figure 4.12, there is a tree of the stock price, and at the last step, we also listed the maximum and the minimum stock prices of each movement path. The stock price starts from S_0, and at each time step, it will go either up by the factor u or down by the factor d. Generally, the monotonicity of the option value highly

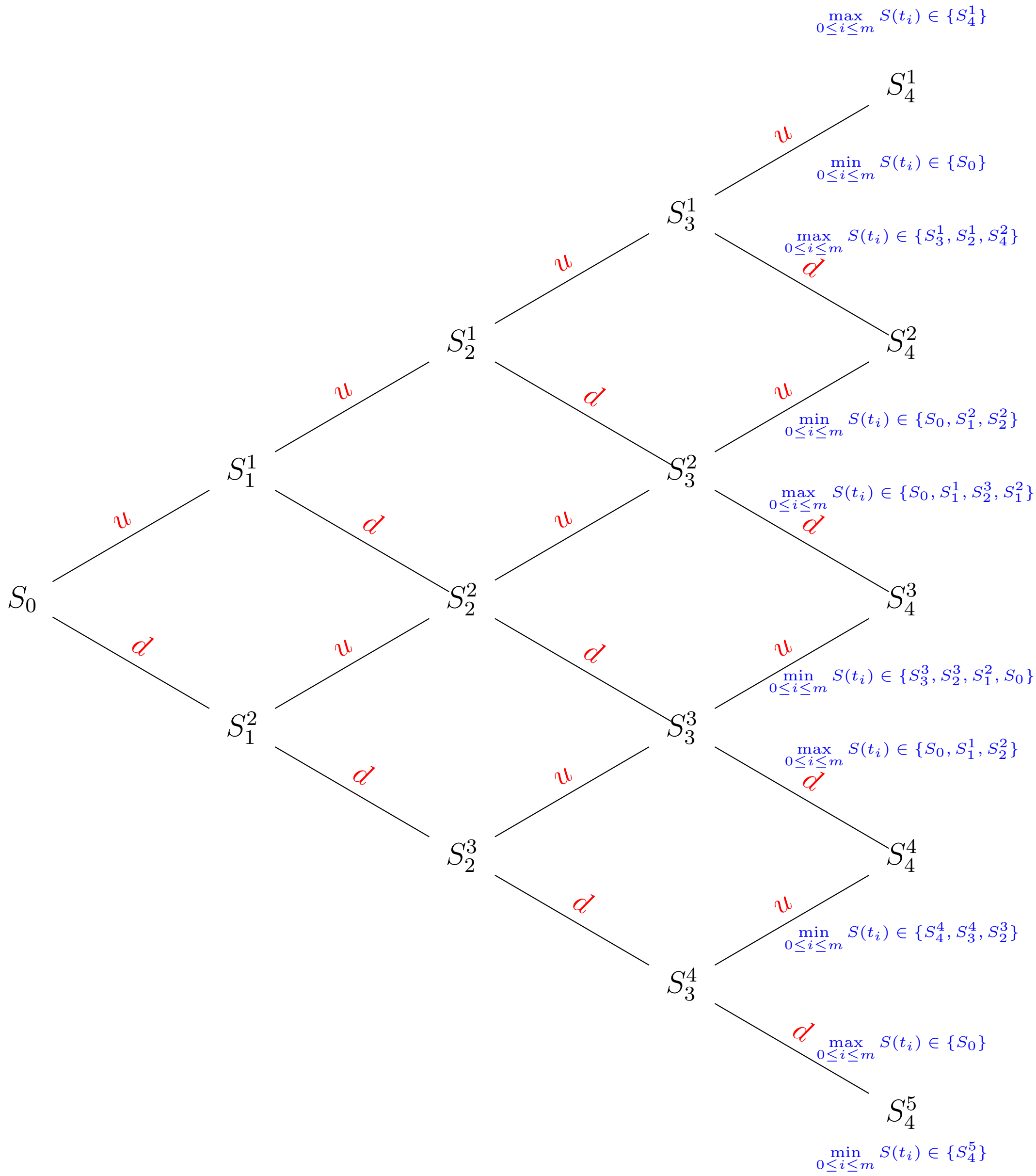


Figure 4.12: The binomial tree of the stock price with maximum and minimum stock price of each path

depends on the movement factor, $ud > 1$, $ud = 1$, or $ud < 1$. For example, in this 4-period option tree the option value is not monotone when $ud < 1$. The maximum stock price of the path $S_0 \to S_1^1 \to S_2^2 \to S_3^2 \to S_4^2$ is S_4^2, while the maximum stock price of the path $S_0 \to S_1^1 \to S_2^1 \to S_3^2 \to S_4^3$ is S_2^1, because $ud < 1$, then $S_4^2 < S_2^1$. But when it comes to the other path ends with the same maturity stock price S_4^2, $S_0 \to S_1^1 \to S_2^1 \to S_3^1 \to S_4^2$, the maximum stock price is S_3^1 which is higher than the maximum stock price of the picked path ending with S_4^3, which is equal to S_2^1.

This causes the results that look-back option payoff is not monotonic. So the option values are not monotonic in the binomial tree. When the option is monotonic, then we can use the same NPI probability assignment as other types of options. If not, we need to think about the probability structure again, which is not considered in this thesis but challenging and interesting topic for a future topic. Here instead of giving a new probability assignment, we offer a new binomial tree which is monotone inspired by the look-back option pricing model presented by Cheuk and Vorst [20] in 1997.

Cheuk and Vorst [20] presented the new binomial approach for the look-back option with float strike price. As acknowledged, the payoff of a look-back call option with a float price is defined as $S(T) - \min_{0 \leq i \leq m} S(t_i)$. Here t_0 is the initial time of the option contract, and t_m is the maturity time T. Then for any time in the binomial tree t_j, denote the minimum stock price of the option life period as $\underline{M}(t_j) = \min_{0 \leq i \leq j} S(t_i) = S(t_j)u^{-k}$, then the look-back call option value is $V(S(t_j), \underline{M}(t_j), t_j)$. Define the power of stock price upward movement factor u:

$$k = \ln[S(t_j)/\underline{M}(t_j)]/\ln(u) \tag{4.34}$$

$S(t_j) \geq \underline{M}(t_j)$, k is positive integer and $k = 0, 1, \ldots, j$, so the option value at each time step can be transferred to a function depending on the stock price $S(t_j)$ and k, i.e.

$$V(S(t_j), \underline{M}(t_j), t_j) = S(t_j) - \underline{M}(t_j) = S(t_j)(1 - u^{-k}) = S(t_j)W_{t_j}(k) \tag{4.35}$$

This claim also holds for the maturity. Hence, by defining $W_{t_j}(k) = 1 - u^{-k}$ we can construct a new binomial tree of $W_{t_j}(k)$, $k = 0, 1, \ldots, j$.

In Figure 4.13, if $k \geq 1$ at t_j and the stock price goes up to $S(t_{j+1}) = S(t_j)u$, then at time t_{j+1} the power of u is $k + 1$. If the stock price goes down, the power of u is $k - 1$. While when $k = 0$, the situation is different, which for the upward

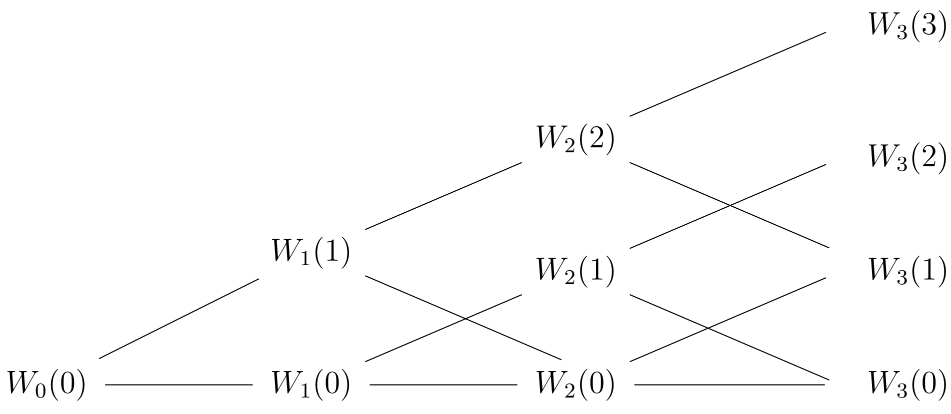


Figure 4.13: The lookback call option with float strike price

movement the power of u is 1, but the power of u for downward movement is still 0. As we can see here the binomial tree of $W_{t_j}(k)$ with the path, we can use the NPI probabilities to evaluate the option. Here for n historical observations, and s represents the number of times that the stock price went up in the previous time. Then we can get the upper and lower probabilities of upward movement from $W_{t_j}(k)$ to $W_{t_{j+1}}(k+1)$ as,

$$\overline{P}(t_j) = \frac{s+k+1}{n+t_j+1} \tag{4.36}$$

$$\underline{P}(t_j) = \frac{s+k}{n+t_j+1} \tag{4.37}$$

Then we apply these NPI probabilities to each one step path. Based on the NPI method, we can compute the expected value of $W_0(0)$, and based on the definition of the look-back call option with float stock price. We know the option price is $S(0)W_0(0)$. The backward method for each node can be formulated as :

The maximum buying price of the call option

$$\begin{aligned}
&\underline{V_0} = S(0)W_0(0) \\
&\underline{W}_{t_j}(k) = B(t_j, t_{j+1}) \left[\underline{P}(t_j)\underline{W}_{t_{j+1}}(k+1) + (1-\underline{P}(t_j))\underline{W}_{t_{j+1}}(k-1) \right]
\end{aligned} \tag{4.38}$$

The minimum selling price of the call option

$$\overline{V_0} = S(0)W_0(0)$$
$$\overline{W_{t_j}}(k) = B(t_j, t_{j+1})\left[\overline{P}(t_j)\overline{W}_{t_{j+1}}(k+1) + (1-\overline{P}(t_j))\overline{W}_{t_{j+1}}(k-1)\right] \quad (4.39)$$

where $B(t_j, t_{j+1})$ is the discount factor from time t_j to time t_{j+1}.

Similarly, we can construct the tree for the look-back put option with float payoff as well. By definition, the payoff of this kind of option is $\max_{0\leq i\leq m} S(t_i) - S(T)$, where $\max_{0\leq i\leq m} S(t_i)$ is the maximum stock price during the whole option life period. Then using a function to represent this value at time t_j is $\overline{M}(t_j) = \max_{0\leq i\leq m} S(t_i) = S(t_j)u^{-k}$. The option value $V(S(t_j), \overline{M}(t_j), t_j)$ depends on three factors, stock price, maximum stock price and the time to maturity. Define the power of upward movement factor u:

$$k = \ln[S(t_j)/\overline{M}(t_j)]/\ln(u) \quad (4.40)$$

As we known, $S(t_j)$ is always less than or equal to $\overline{M}(t_j)$, then k is a negative integer belongs to the set of values $\{0, \ldots, -j\}$. Then we can rewrite the option value as:

$$V(S(t_j), \overline{M}(t_j), t_j) = \max_{0\leq i\leq m} S(t_i) - S(t_j) = (u^{-k} - 1)S(t_j) = S(t_j)G_{t_j}(k) \quad (4.41)$$

Define $G_{t_j}(k) = u^{-k} - 1$, then we construct a new binomial tree of $G(k, t_j)$.

In the tree, Figure 4.14, when k is negative and the stock price goes down at t_j, for the next time step the power of u is $k-1$. Or if the stock price goes up at time t_j, the power of u is $k+1$ at time t_{j+1}. When $k = 0$, the downward movement will change the k to $k-1$, but upward movement won't change the power of u. From this monotone tree, we can use the NPI method to calculate the maximum buying price and the minimum selling price of the option. Here n is the number of the historical stock price, among them s stock prices go down. Then for each downward

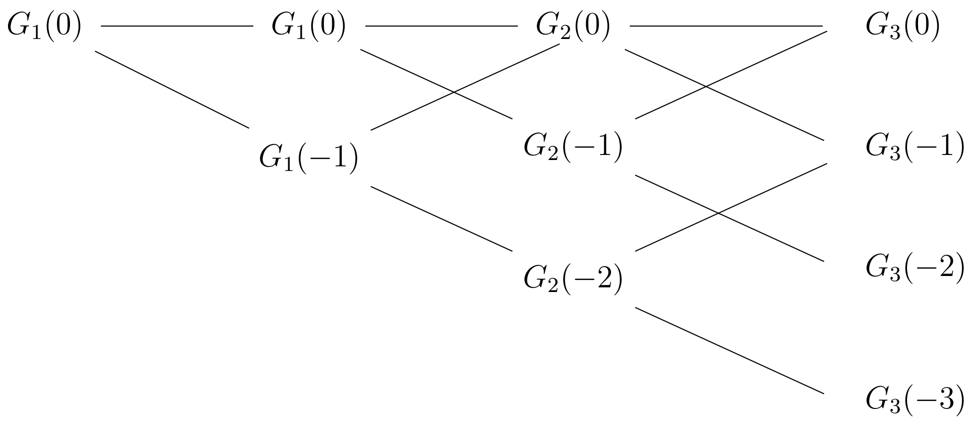


Figure 4.14: The lookback put option with float strike price

path, we have the upper and lower probabilities as:

$$\overline{P}(t_j) = \frac{s-k+1}{n+t_j+1} \tag{4.42}$$

$$\underline{P}(t_j) = \frac{s-k}{n+t_j+1} \tag{4.43}$$

The backward method for each node can be formulated as :

The maximum buying price of the put option

$$\begin{aligned} &\underline{V_0} = S(0)G(0,0) \\ &\underline{G}_{t_j}(k) = B(t_j, t_{j+1})\left[(1-\underline{P}(t_j))\underline{G}_{t_{j+1}}(k+1) + \underline{P}(t_j)\underline{G}_{t_{j+1}}(k-1)\right] \end{aligned} \tag{4.44}$$

The minimum selling price of the put option

$$\begin{aligned} &\overline{V_0} = S(0)G(0,0) \\ &\overline{G}_{t_j}(k) = B(t_j, t_{j+1})\left[(1-\overline{P}(t_j))\overline{G}_{t_{j+1}}(k+1) + \overline{P}(t_j)\overline{G}_{t_{j+1}}(k-1)\right] \end{aligned} \tag{4.45}$$

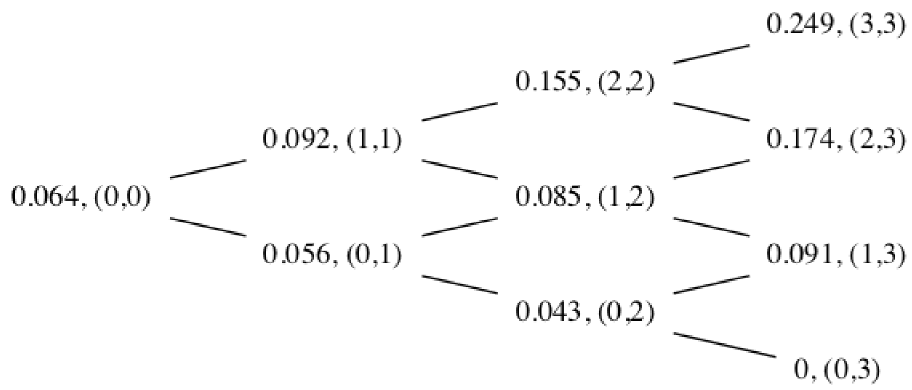


Figure 4.15: The binomial tree of a look-back call option with the float strike price

Example 4.7

In this example, we use R program in Appendix B.6 to predict the value of a look-back call option with the float strike price $K = \min_{0 \leq i \leq m} S(t_i)$ derived from stock with an initial stock price $S_0 = 20$. Following the mathematical description for the look-back call option, we first set up a binomial tree of $W_{t_j}(k)$ as Figure 4.15. In the tree, there are three values at each node. Two values in the parenthesis are the time steps t and the power of the upward movement factor k. The value outside the parenthesis is the value of $W_{t_j}(k)$ at each node. After this discounted backward evaluation method, we get the value $W_0(0) = 0.064$, then we can calculate the option price, which is $V(0) = S_0 W_0(0) = 20 \times 0.064 = 1.28$.

4.5 Concluding remarks

The NPI method can be used in the evaluation of the exotic options, for the type of option with a monotonic binomial tree the probability assignment of each path is quite applicable and understandable, like the digital option and the barrier option. For the options with the non-monotonic binomial tree, we can manipulate the payoff definition and construct a monotonic binomial tree and use the NPI method lower and upper probabilities to calculate the maximum buying option price and the minimum selling option price as shown for the look-back options with the float strike price. Or we can also give the original tree a different probability assignment in order to calculate the upper and lower expectations, which can be challenging and appealing. Although we do not discuss the performance of the NPI option pricing method for these three exotic options, we believe that the conclusion of its performance is the same as the NPI method for vanilla options, which is the NPI method performs better than the CRR model when the CRR model uses the substantially wrong assumptions. There are still a lot of research topics about the exotic option on the basis of the NPI method. One of them is to explore the NPI method for more exotic options, like the Asian option [76], and the study of the NPI performance is also of interest.

Chapter 5

Conclusion

In this thesis, we have applied Nonparametric Predictive Inference (NPI) to a new area, the option pricing. To start the NPI method for option pricing at an early stage, we developed the NPI method only focusing on the binomial tree model. Using NPI for Bernoulli data, as reviewed in Chapter 1, we set up the pricing model based on the NPI method not only for vanilla options but also some exotic options.

We first priced the European option by applying the NPI method to the model. Instead of getting a precise expected option price, we got an interval of European option prices, with upper and lower boundaries, called the minimum selling price and the maximum buying price, respectively. The boundary prices are named according to the trade preferences, and all the values in the interval are reasonable for this option from the model prediction. Then any value greater than the upper bound is appealing to sell, and any value less than the lower bound is appealing to buy. Then we checked into a famous property in the classic theory of the European option: the put-call parity. In the classic theory, this parity is only valid when there is no arbitrage opportunity in the market. Although our prediction is imprecise indicating arbitrage opportunity, the boundary prices still follow the put-call parity. From the formula we know that the maximum buying price of a call option has an equilibrium relationship with the minimum selling price of a put option, also applying to the maximum buying price of a put option and the minimum selling price of a call option. These are following the explanation of the classic theory, that holding a call option and selling a put option is equivalent to holding a forward contract with the same strike price and maturity based on the same underlying asset. Finally, the NPI method performance has been assessed by doing an analytical study of trade between the NPI method and the CRR model in two extreme scenarios. In Scenario 1, the CRR model is using the same information as the real market, while in Scenario 2, the CRR model uses the wrong assumption. We did this study in two ways both with and without the discount procedure. Excluding the discount procedure shows the difference causing by the price method differences, which the NPI prediction is never better than the CRR prediction in Scenario 1, but better than the CRR perdition in Scenario 2. However, the loss of the NPI investor in Scenario 1 can be reduced by enlarging the historical data. In Scenario 2, when the difference between the real market probability and the risk-neutral probability gets larger, the profit earned by the NPI investor raises as well. By adding the discount procedure, we found out that an appropriate discount rate can reduce the prediction error of the NPI method in Scenario 1 and it does not change the result of NPI method performance in Scenario 2.

In Chapter 3, we set up the NPI pricing method for the American option. Different from the European option, because of the early exercise feature, the NPI method for American option pricing does not have a closed formula. Instead, we offered a backward strategy for prediction and described it mathematically. After setting up the pricing model, we presented that in our method the American option without dividends is possible to be early exercise. We also gave the holding condition for both put and call options. Then based on all the knowledge, we studied the stopping time and the NPI method performance by simulation. Due to the different stopping time between our method and the CRR model, the option payoff, and the option price varies that influent the P&L of the NPI method in the performance study. We picked the example of trade between the NPI method and the CRR model with entirely different stopping time to see the extreme case of the NPI method. Similar to the European option, we also did this study in two extreme scenarios. It turns out that the NPI person hardly gets profit from this trade in Scenario 1 but performs very well in Scenario 2. This excellent performance is reflected from two aspects, first the NPI prediction result guide the investor who uses the NPI method only to take part in the smart trading position that has more chance to earn some money, and in these wise trading position the P&L of the NPI method is better than the ones in the first scenario. The influence of the number of historical data was also developed in the examples, showing that a sufficient amount of historical data supports the NPI investor trades more wisely and earns more money. From the results, it is also evident that a larger difference between the real market probability and the risk-neutral probability helps the NPI investor beat the CRR investor in the trade.

Last but not least, we also presented the NPI method for some exotic options, i.e., the digital option, the barrier option, and the look-back option. For the digital option and the barrier option, the NPI application is straightforward, for the payoffs of these options are monotonic with the underlying asset price binomial tree. Therefore, according to the option definition, the minimum selling price and the maximum buying price is achieved by assigning the upper or the lower probabilities to each movement path in the binomial tree. When it comes to the option having non-monotonic payoffs, there are two ways we could deal with; one is manipulating the binomial tree making monotonic payoffs, like what we did for the look-back option with the float strike price. Or we assign the new imprecise probability to the underlying asset price binomial tree, which is more challenging for future study because we need to check its consistency and exchangeability.

A further topic of interest for further study in our method is whether or not all historical data should be taken into account. It is better to do so if one can safely assume that the future observations will be exchangeable with all the past data. However, if one believes that there has been a considerable change in the data at some point in the past, it may be appropriate to restrict the historical data to observations after such a change. We only consider the basic binomial tree model as a simple first step of this research. This method is only used for ideal model situations, but it underpins a range of more realistic models for which we aim to investigate in the future, for instance, the trinomial model. We study the NPI method only by comparing its performance in regard to another trader who would use the CRR model, either with perfect or imperfect knowledge. Real world scenarios with multiple traders are also interesting for further research. It is also interesting to investigate the application of the NPI method for option pricing in real markets where it may also be possible to improve the method by creating hybrid strategies based on multiple pricing methods. Also, NPI for real-valued data can be developed for option pricing in real-time with continuous-valued increment. Then it is interesting to study aspects of the investigation of the discrete tree model to the continuous model.

Appendix A

Financial Terminology

The financial definition are from the textbook, Options, Futures, and Other Derivatives, written by Hull [45]

- **All-or-Nothing Option** An option has the a predetermined value as it payoff if the option is exercised.

- **American Option** A vanilla option which is able to be exercised before its maturity.

- **Arbitrage** A trading strategy that an investor could take when financial products are mispriced, and a certain profit can be gain through opposite trading action on the same product or on more products but should have the same value.

- **Asset-or-Nothing Option** An option has the value equal to the asset price as its payoff if the option is exercised.

- **At the Money option** An option in which the strike price equals the underlying asset price.

- **Backward Strategy** A procedure for working from the end of a tree to its beginning in order to value an option.

- **Barrier Option** An option whose payoff depends on if the path of the underlying asset has reached a barrier.
- **Call Option** An option that option buyer agree to buy the underlying asset at a certain price by a settled date, and its payoff is $[S-K]^+$.
- **Derivative** An financial product whose prices depends on the price of another asset.
- **Digital Option** An option with a discontinuous payoff; for example, a all-or-nothing option or a asset-or-nothing option.
- **Early Exercise** The exercise time is before the maturity date.
- **European Option** An option that can only be exercised at its maturity.
- **Exotic Option** A non-vanilla option.
- **Forward** A financial instrument that obligates the holder to buy or sell an asset for a predermined price at a predetermined future time.
- **Hedge** A trading action that is designed to eliminate or reduce risk.
- **Instant Value** Either a call option value where the asset price minus the strike price or a put option value where the strike price minus the asset price.
- **In the Money Option** The call option where the asset price is greater than the strike price, or the put option where the asset price is less than the strike price.
- **Long Position** A position an investor can be in through purchasing the asset.
- **Look-back Option** An option has the payoff that is depend on the maximum or minimum of the asset price achieved during a specific period of time.

- **Option** One of the financial derivatives designed to let the buyer has the right to buy or sell the underlying asset at a certain price by the certain date.

- **Out of the Money Option** The call option where the asset price is less than the strike price, or the put option where the asset price is greater than the strike price.

- **Payoff** The value gained by the buyer of an option or other derivatives at the end of its life.

- **Put Option** An option the holder agree to sell the underlying asset at a certain price by a certain date, and its payoff is $[K - S]^+$.

- **Risk-free Rate** The interest rate earned from the asset without any risk.

- **Risk Neutral Valuation** The valuation of an option or other products in the derivative market under the assumption of risk neutral market.

- **Risk Neutral Market** A market that investors are assumed to require no extra return on average for bearing risks.

- **Short Position** A position that the holder decides to be in by selling the asset.

- **Spot Price** The price for immediate delivery.

- **Strike Price** The price at which the underlying asset may be bought or sold at the exercise time.

- **Time Value** The value of an option arising from the time left to maturity.

Appendix B

R code

B.1 American option

```r
##This file is to generate the binomial tree of the American
   option.
##Use function treedot, where Asset is the initial stock price,
##u is the up factor, d is the down factor,
##IntRate is the discount rate,
##Strike is the strike price, NoSteps is the future time steps,
##n is the historical data, s is the successful historical data.
##Then type in the storage path in the terminal with graphviz
   installed.


#instant payoff of the American option
payoff.vanilla.call <- function(Asset, Strike) {
  return(max(0, Asset - Strike))
}

payoff.vanilla.put <- function(Asset, Strike) {
  return(max(0, Strike - Asset))
}
```

```
# Generate a binomial lattice for American option.
# This function is modified based on the NPI method according to
    the function  written by Ollie's Universe.
# web:http://www.frolovs.me/
genlattice.american <-
  function(Asset, u, d, IntRate, Strike, NoSteps, Payoff, n, s,
      OptionType) {
    OptionPosition <-
      readline(prompt = "Enter option investor's position: buy
         or sell:")
    # The number of tree nodes to process.
    count <- sum(1:(NoSteps + 1))

    # This data frame will store asset and option prices.
    # The mapping from tree node (i,j) to linear index
    # inside the data frame will have to be computed.
    # The early exercise flag is also stored.
    X <- data.frame(matrix(NA, nrow = count, ncol = 3))
    names(X) <- c("asset", "option", "exercise")

    #Option price discount factor.
    #Here the IntRate is the expected stock return in the thesis
    DiscountFactor <- 1 / (1 + IntRate)

    #p is the NPI boundary probability
    p <- 0

    # Compute the asset and option prices, starting from the
       last node of the tree, which is its bottom right corner
       when viewed as a graph.
    # Work up and backwards.
    for (i in NoSteps:0) {
      for (j in i:0) {
```

```
AssetCurrent <- Asset * u ^ (i - j) * d ^ j
X$asset[count] <- AssetCurrent

#Compute the payoff directly for the last step's nodes,
   otherwise use a formula.
if (i == NoSteps) {
  X$option[count]   <- Payoff(AssetCurrent, Strike)
  X$exercise[count] <- FALSE
} else {
  #The up and down jump factors
  up   <- X$option[sum(1:(i + 1), j, 1)]
  down <- X$option[sum(1:(i + 1), j + 1, 1)]
  if ((OptionType == "Call" &
       OptionPosition == "buy") |
      (OptionType == "Put" & OptionPosition == "sell"))
         {
    p <- (s + i - j) / (n + i + 1)# the lower
       probability
  } else if ((OptionType == "Call" &
              OptionPosition == "sell") |
             (OptionType == "Put" &
              OptionPosition == "buy")) {
    p <- (s + i - j + 1) / (n + i + 1)# the upper
       probability
  }

  #Possible option values when discounted or when early
     exercise is applied.
  V <- DiscountFactor * (p * up + (1 - p) * down)
  V.early <- Payoff(AssetCurrent, Strike)

  # The greatest of two possible values is stored
  X$option[count] <- max(V, V.early)
```

```
          #Should the option be exercised early?
          X$exercise[count]  <- (V.early > V)
        }

        count <- count - 1
      }
    }

    return(X)
  }

#Genlattice family functions:different lattice by the option
   type: call or put
genlattice.vanilla.american.call <-
  function(Asset, u, d, IntRate, Strike, NoSteps, n, s,
     OptionType = OptionType) {
    return (
      genlattice.american( Asset = Asset, u = u, d = d, IntRate
         = IntRate, Strike = Strike, NoSteps = NoSteps, Payoff =
          payoff.vanilla.call, n = n, s = s, OptionType =
         OptionType )
    )
  }

genlattice.vanilla.american.put <-
  function(Asset, u, d, IntRate, Strike, NoSteps, n, s,
     OptionType = OptionType) {
    return (
      genlattice.american( Asset = Asset, u = u, d = d, IntRate
         = IntRate, Strike = Strike, NoSteps = NoSteps, Payoff =
          payoff.vanilla.put, n = n, s = s, OptionType =
         OptionType)
```

```
    )
  }

#Generates a graph specification that can be fed into graphviz.
#Input: the binomial lattice produced by one of genlattice
   family functions.
#This function was borrowed Rory Winston:
#http://www.theresearchkitchen.com/archives/738
dotlattice <- function(S, digits = 2) {
  shape <- "plaintext"

  cat("digraph G {", "\n", sep = "")
  cat("node[shape=", shape, "];", "\n", sep = "")
  cat("rankdir=LR;", "\n")

  cat("edge[arrowhead=none];", "\n")

  # Create a dot node for each element in the lattice
  for (i in 1:nrow(S)) {
    x <- round(S$asset[i], digits = digits)
    y <- round(S$option[i], digits = digits)

    # Detect the American tree and draw accordingly
    early.exercise <- ""
    if (("exercise" %in% colnames(S)) && S$exercise[i]) {
      early.exercise <- "shape=oval,"
    }

    cat("node", i, "[", early.exercise, "label=\"", x, ", ", "("
       , y, ")",  "\"];", "\n", sep = "")
  }

  # The number of levels in a binomial lattice of length L is
```

```
    `$\frac{\sqrt{8N+1}-1}{2}$`
  L <- ((sqrt(8 * nrow(S) + 1) - 1) / 2 - 1)

  k <- 1
  for (i in 1:L) {
    tabs <- rep("\t", i - 1)
    j <- i
    while (j > 0) {
      cat("node", k, "->", "node", (k + i), ";\n", sep = "")
      cat("node", k, "->", "node", (k + i + 1), ";\n", sep = "")
      k <- k + 1
      j <- j - 1
    }
  }

  cat("}", sep = "")
}

#Plot the binomial tree of the American option in file lattice.
   dot.
treedot <- function(Asset, u, d, IntRate, Strike, NoSteps, n, s)
    {
  OptionType <- readline(prompt = "Enter option type: Call or
     Put:")
  if (OptionType == "Call") {
    x <- genlattice.vanilla.american.call(  Asset = Asset, u = u
       , d = d, IntRate = IntRate, Strike = Strike, NoSteps =
       NoSteps, n = n, s = s, OptionType = OptionType)
  } else if (OptionType == "Put") {
    x <- genlattice.vanilla.american.put( Asset = Asset, u = u,
       d = d, IntRate = IntRate, Strike = Strike, NoSteps =
       NoSteps, n = n, s = s, OptionType = OptionType)
  }
```

```
  y <- capture.output(dotlattice(x, digits = 3))

  cat(y, file = "lattice.dot")

}

#Terminal: dot -Tpng -o binomial-tree-american-call-1.png /...
   the path where you save lattice.dot
```

B.2 American all-or-nothing digital option

```
##This file is to generate the binomial tree of the American all
   -or-nothing digital option.
##Use function treedot, where Asset is the initial stock price,
##u is the up factor, d is the down factor,
##IntRate is the discount rate,
##Strike is the strike price, NoSteps is the future time steps,
##n is the historical data, s is the successful historical data.
##PX is the constant payoff
##Then type in the storage path in the terminal with graphviz
   installed.

#The payoff for call and put options.
payoff.vanilla.call <- function(Asset, Strike, PX) {
  if (Asset - Strike > 0) {
    pay <- PX
  } else{
    pay <- 0
  }
  return(pay)
}
```

```
payoff.vanilla.put <- function(Asset, Strike, PX) {
  if (Strike - Asset > 0) {
    pay <- PX
  } else {
    pay <- 0
  }
  return(pay)
}

# Generate a binomial lattice for the digital option.
genlattice.american <- function(Asset, u, d,IntRate, Strike,
   NoSteps, Payoff, n, s, PX, OptionType) {
    OptionPosition<-readline(prompt = "Enter option investor's
       position: buy or sell:")
    # The number of tree nodes to process.
    count <- sum(1:(NoSteps + 1))

    # This data frame will store asset and option prices.
    # The mapping from tree node (i,j) to linear index
    # inside the data frame will have to be computed.
    # The early exercise flag is also stored.
    X <- data.frame(matrix(NA, nrow = count, ncol = 3))
    names(X) <- c("asset", "option", "exercise")

    # Option price discount factor.
    #Here the IntRate is the expected return for only onestep
    DiscountFactor <- 1 / (1 + IntRate)

    #The NPI probability.
    p <- 0

    # Compute the asset and option prices, starting from the
       last node of the tree, which is its bottom right corner
```

```
      when viewed as a graph.
  # Work up and backwards.
  for (i in NoSteps:0) {
    for (j in i:0) {
      AssetCurrent <- Asset * u ^ (i - j) * d ^ j
      X$asset[count] <- AssetCurrent

      # Compute the payoff directly for the last step's nodes,
          otherwise use a formula.
      if (i == NoSteps) {
        X$option[count]   <- Payoff(AssetCurrent, Strike, PX)
        X$exercise[count] <- FALSE
      } else {
        #The up and down jump factors
        up   <- X$option[sum(1:(i + 1), j, 1)]
        down <- X$option[sum(1:(i + 1), j + 1, 1)]
        if ((OptionType == "Call" &
             OptionPosition == "buy") |
            (OptionType == "Put" & OptionPosition == "sell"))
               {
          p <- (s + i - j) / (n + i + 1)# the lower
             probability
        } else if ((OptionType == "Call" &
                    OptionPosition == "sell") |
                   (OptionType == "Put" &
                    OptionPosition == "buy")) {
          p <- (s + i - j + 1) / (n + i + 1)# the upper
             probability
        }

        # Possible option values when discounted or when early
            exercise is applied
        V <- DiscountFactor * (p * up + (1 - p) * down)
```

```
          V.early <- Payoff(AssetCurrent, Strike, PX)

          # The greatest of two possible values is stored
          X$option[count] <- max(V, V.early)

          # Should the option be exercised early?
          X$exercise[count]  <- (V.early > V)
        }

        count <- count - 1
      }
    }

    return(X)
  }

#Genlattice family functions:
#different lattice by the option type:call or put
genlattice.vanilla.american.call <- function(Asset, u, d,
   IntRate,  Strike,  NoSteps, n, s, PX, OptionType=OptionType)
   {
  return (genlattice.american(Asset = Asset, u = u, d = d,
     IntRate = IntRate, Strike = Strike,  NoSteps = NoSteps,
     Payoff = payoff.vanilla.call, n = n, s = s, PX = PX,
     OptionType = OptionType))
}

genlattice.vanilla.american.put <- function(Asset, u, d, IntRate
   , Strike, NoSteps, n, s, PX, OptionType=OptionType) {
  return (genlattice.american(Asset = Asset, u = u, d = d,
     IntRate = IntRate, Strike = Strike,  NoSteps = NoSteps,
     Payoff = payoff.vanilla.put, n = n, s = s, PX = PX,
     OptionType = OptionType))
```

```
}

# Generates a graph specification that can be fed into graphviz.
# Input: the binomial lattice produced
#by one of genlattice family functions.
dotlattice <- function(S, digits = 2) {
  shape <- "plaintext"

  cat("digraph G {", "\n", sep = "")
  cat("node[shape=", shape, "];", "\n", sep = "")
  cat("rankdir=LR;", "\n")

  cat("edge[arrowhead=none];", "\n")

  # Create a dot node for each element in the lattice
  for (i in 1:nrow(S)) {
    x <- round(S$asset[i], digits = digits)
    y <- round(S$option[i], digits = digits)

    # Detect the American tree and draw accordingly
    early.exercise <- ""
    if (("exercise" %in% colnames(S)) && S$exercise[i]) {
      early.exercise <- "shape=oval,"
    }

   cat("node", i, "[", early.exercise, "label=\"", x, ", ","(",
      y,")", "\"];", "\n", sep="")
  }

  # The number of levels in a binomial lattice of length L is
     `$\frac{\sqrt{8N+1}-1}{2}$`
  L <- ((sqrt(8 * nrow(S) + 1) - 1) / 2 - 1)
```

```
  k <- 1
  for (i in 1:L) {
    tabs <- rep("\t", i - 1)
    j <- i
    while (j > 0) {
      cat("node", k, "->", "node", (k + i), ";\n", sep = "")
      cat("node", k, "->", "node", (k + i + 1), ";\n", sep = "")
      k <- k + 1
      j <- j - 1
    }
  }

  cat("}", sep = "")
}

#Plot the binomial tree of the Ameican all-or-nothing digital
   option in file lattice.dot.
treedot<-function(Asset, u, d, IntRate, Strike, NoSteps, n, s,
   PX) {
  OptionType <- readline(prompt = "Enter option type: Call or
     Put:")
  if (OptionType == "Call") {
    x <- genlattice.vanilla.american.call(Asset = Asset, u = u,
       d = d, IntRate = IntRate,  Strike = Strike,  NoSteps =
       NoSteps, n = n, s = s, PX = PX, OptionType = OptionType)
  } else if (OptionType == "Put") {
    x <- genlattice.vanilla.american.put(Asset = Asset, u = u, d
        = d, IntRate = IntRate,  Strike = Strike,  NoSteps =
       NoSteps, n = n, s = s, PX = PX, OptionType = OptionType)
  }
  y <- capture.output(dotlattice(x, digits = 3))
  cat(y, file = "lattice.dot")
}
```

```
# Terminal: dot -Tpng -o binomial-tree-american-call-1.png /...
   the path where you save lattice.dot
```

B.3 American asset-or-nothing digital option

```
##This file is to generate the binomial tree of the American
   asset-or-nothing digital option.
##Use function treedot, where Asset is the initial stock price,
##u is the up factor, d is the down factor,
##IntRate is the discount rate,
##Strike is the strike price, NoSteps is the future time steps,
##n is the historical data, s is the successful historical data.
##Then type in the storage path in the terminal with graphviz
   installed.

#The payoff for call and put options.
payoff.vanilla.call <- function(Asset, Strike) {
  if (Asset - Strike > 0) {
    pay <- Asset
  } else{
    pay <- 0
  }
  return(pay)
}

payoff.vanilla.put <- function(Asset, Strike) {
  if (Strike - Asset > 0) {
    pay <- Asset
  } else {
    pay <- 0
  }
  return(pay)
```

```
}


# Generate a binomial lattice for the American asset-or-nothing
   option.
genlattice.american <-
  function(Asset, u, d, IntRate, Strike, NoSteps, Payoff, n, s,
     OptionType) {
    OptionPosition <-
      readline(prompt = "Enter option investor's position: buy
         or sell:")
    # The number of tree nodes to process.
    count <- sum(1:(NoSteps + 1))

    # This data frame will store asset and option prices.
    # The mapping from tree node (i,j) to linear index
    # inside the data frame will have to be computed.
    # The early exercise flag is also stored.
    X <- data.frame(matrix(NA, nrow = count, ncol = 3))
    names(X) <- c("asset", "option", "exercise")



    # Option price discount factor.
    #Here the IntRate is the expected return for only onestep
    DiscountFactor <- 1 / (1 + IntRate)

    #The NPI probability.
    p <- 0

    # Compute the asset and option prices, starting from the
       last node of the tree, which is its bottom right corner
       when viewed as a graph.
    # Work up and backwards.
    for (i in NoSteps:0) {
```

```
for (j in i:0) {
  AssetCurrent <- Asset * u ^ (i - j) * d ^ j
  X$asset[count] <- AssetCurrent

  # Compute the payoff directly for the last step's nodes,
     otherwise use a formula.
  if (i == NoSteps) {
    X$option[count]   <- Payoff(AssetCurrent, Strike)
    X$exercise[count] <- FALSE
  } else {
    #The up and down jump factors
    up   <- X$option[sum(1:(i + 1), j, 1)]
    down <- X$option[sum(1:(i + 1), j + 1, 1)]
    if ((OptionType == "Call" &
         OptionPosition == "buy") |
        (OptionType == "Put" & OptionPosition == "sell"))
           {
      p <- (s + i - j) / (n + i + 1) # the lower
         probability
    } else if ((OptionType == "Call" &
                OptionPosition == "sell") |
               (OptionType == "Put" &
                OptionPosition == "buy")) {
      p <- (s + i - j + 1) / (n + i + 1) # the upper
         probability
    }

    # Possible option values when discounted or when early
       exercise is applied
    V <- DiscountFactor * (p * up + (1 - p) * down)
    V.early <- Payoff(AssetCurrent, Strike)

    # The greatest of two possible values is stored
```

```
          X$option[count] <- max(V, V.early)

          # Should the option be exercised early?
          X$exercise[count]  <- (V.early > V)
        }

        count <- count - 1
      }
    }

    return(X)
  }
#Genlattice family functions:different lattice by the option
   type:
#call or put
genlattice.vanilla.american.call <-
  function(Asset, u, d, IntRate, Strike, NoSteps, n, s,
     OptionType = OptionType) {
    return (
      genlattice.american( Asset = Asset, u = u, d = d, IntRate
         = IntRate, Strike = Strike,  NoSteps = NoSteps, Payoff
         = payoff.vanilla.call, n = n, s = s, OptionType =
         OptionType)
    )
  }

genlattice.vanilla.american.put <-
  function(Asset, u, d, IntRate, Strike, NoSteps, n, s,
     OptionType = OptionType) {
    return (
      genlattice.american(Asset = Asset, u = u, d = d, IntRate =
          IntRate, Strike = Strike, NoSteps = NoSteps, Payoff =
         payoff.vanilla.put, n = n, s = s, OptionType =
```

```
        OptionType)
    )
  }

# Generates a graph specification that can be fed into graphviz.
# Input: the binomial lattice produced by
# one of genlattice family functions.
dotlattice <- function(S, digits = 2) {
  shape <- "plaintext"

  cat("digraph G {", "\n", sep = "")
  cat("node[shape=", shape, "];", "\n", sep = "")
  cat("rankdir=LR;", "\n")

  cat("edge[arrowhead=none];", "\n")

  # Create a dot node for each element in the lattice
  for (i in 1:nrow(S)) {
    x <- round(S$asset[i], digits = digits)
    y <- round(S$option[i], digits = digits)

    # Detect the American tree and draw accordingly
    early.exercise <- ""
    if (("exercise" %in% colnames(S)) && S$exercise[i]) {
      early.exercise <- "shape=oval,"
    }

    cat("node", i, "[", early.exercise, "label=\"", x, ", ", "("
      , y, ")", "\"];", "\n", sep = "")
  }

  # The number of levels in a binomial lattice of length L is
     `$\frac{\sqrt{8N+1}-1}{2}$`
```

```
  L <- ((sqrt(8 * nrow(S) + 1) - 1) / 2 - 1)

  k <- 1
  for (i in 1:L) {
    tabs <- rep("\t", i - 1)
    j <- i
    while (j > 0) {
      cat("node", k, "->", "node", (k + i), ";\n", sep = "")
      cat("node", k, "->", "node", (k + i + 1), ";\n", sep = "")
      k <- k + 1
      j <- j - 1
    }
  }

  cat("}", sep = "")
}

# Plot the binomial tree of the Ameican all-or-nothing digital
   option in file lattice.dot.
treedot <- function(Asset, u, d, IntRate, Strike, NoSteps, n, s)
    {
  OptionType <- readline(prompt = "Enter option type: Call or
     Put:")
  if (OptionType == "Call") {
    x <-
      genlattice.vanilla.american.call(Asset = Asset, u = u, d =
          d, IntRate = IntRate, Strike = Strike,
        NoSteps = NoSteps, n = n, s = s, OptionType = OptionType
      )
  } else if (OptionType == "Put") {
    x <-
      genlattice.vanilla.american.put(Asset = Asset, u = u, d =
         d, IntRate = IntRate, Strike = Strike, NoSteps =
```

```
            NoSteps, n = n, s = s, OptionType = OptionType)
  }
  y <- capture.output(dotlattice(x, digits = 3))
  cat(y, file = "lattice.dot")
}

# Terminal: dot -Tpng -o binomial-tree-american-call-1.png /...
    the path where you save lattice.dot
```

B.4 Knock up-and-out barrier option

```
##This is the R code for knock up and out option
##Use function treedot, where Asset is the initial stock price,
##u is the up factor, d is the down factor,
##IntRate is the discount rate,
##Strike is the strike price, NoSteps is the future time steps,
##n is the historical data, s is the successful historical data.
##B is the barrier value.
##Then type in the storage path in the terminal with graphviz
    installed.

# The payoff for call and put options.
payoff.vanilla.call <- function(Asset, Strike) {
  pay <- max(Asset - Strike, 0)
  return(pay)
}

payoff.vanilla.put <- function(Asset, Strike) {
  pay <- max(Strike - Asset, 0)
  return(pay)
}
```

```
# Generate a binomial lattice for the knock up and out option.
genlattice.american <-
  function(Asset, u, d, IntRate, Strike, NoSteps, Payoff, n, s,
     OptionType, B) {
    OptionPosition <-
      readline(prompt = "Enter option investor's position: buy
         or sell:")
    # The number of tree nodes to process.
    count <- sum(1:(NoSteps + 1))

    # This data frame will store asset and option prices.
    # The mapping from tree node (i,j) to linear index
    # inside the data frame will have to be computed.
    # The early exercise flag is also stored.
    X <- data.frame(matrix(NA, nrow = count, ncol = 3))
    names(X) <- c("asset", "option", "exercise")



    # Option price discount factor.
    # Here the IntRate is the expected return for only onestep.
    DiscountFactor <- 1 / (1 + IntRate)

    #The NPI probability.
    p <- 0

    # Compute the asset and option prices, starting from the
       last node of the tree, which is its bottom right corner
       when viewed as a graph.
    # Work up and backwards.
    for (i in NoSteps:0) {
      for (j in i:0) {
        AssetCurrent <- Asset * u ^ (i - j) * d ^ j
        X$asset[count] <- AssetCurrent
```

```
    # Compute the payoff directly for the last step's nodes,
        otherwise use a formula.
    if (i == NoSteps) {
      #Check the last maturity stock price reaching the
         barrier or not.
      if (X$asset[count] < B) {
        X$option[count]   <- Payoff(AssetCurrent, Strike)
        X$exercise[count] <- "FALSE"
      } else{
        X$option[count] <- 0
        if (Payoff(AssetCurrent, Strike) != 0) {
          X$exercise[count] <- "TRUE"
        } else{
          X$exercise[count] <- "FALSE"
        }

      }
    } else {
      # The up and down jump factors
      up   <- X$option[sum(1:(i + 1), j, 1)]
      down <- X$option[sum(1:(i + 1), j + 1, 1)]
      if (X$asset[count] < B) {
        if ((OptionType == "Call" &
             OptionPosition == "buy") |
            (OptionType == "Put" & OptionPosition == "sell")
               ) {
          p <- (s + i - j) / (n + i + 1)
        } else if ((OptionType == "Call" &
                   OptionPosition == "sell") |
                  (OptionType == "Put" & OptionPosition ==
                     "buy")) {
          p <- (s + i - j + 1) / (n + i + 1)
```

```
            }
            X$option[count] <- DiscountFactor * (p * up + (1 - p
               ) * down)
            X$exercise[count] <- "FALSE"
          } else{
            X$option[count] <- 0
            #check the stock price reaching the barrier or not
            if (up != 0 || down != 0) {
              X$exercise[count] <- "TRUE"
            }
            else {
              X$exercise[count] <- "FALSE"
            }
          }
        }
        count <- count - 1
      }
    }

    return(X)
  }

#Genlattice family functions:
#different lattice by the option type:call or put
genlattice.vanilla.american.call <-
  function(Asset, u, d, IntRate, Strike, NoSteps, n, s,
     OptionType = OptionType, B) {
    return (
      genlattice.american(Asset = Asset, u = u, d = d, IntRate =
          IntRate, Strike = Strike, NoSteps = NoSteps, Payoff =
         payoff.vanilla.call, n = n, s = s, OptionType =
         OptionType, B = B)
    )
```

```
  }

genlattice.vanilla.american.put <-
  function(Asset, u, d, IntRate, Strike, NoSteps, n, s,
     OptionType = OptionType, B) {
    return (
      genlattice.american(Asset = Asset, u = u, d = d, IntRate =
          IntRate, Strike = Strike, NoSteps = NoSteps, Payoff =
         payoff.vanilla.put, n = n, s = s, OptionType =
         OptionType, B = B)
    )
  }
# Generates a graph specification that can be fed into graphviz.
# Input: the binomial lattice produced
# by one of genlattice family functions.
dotlattice <- function(S, digits = 2) {
  shape <- "plaintext"

  cat("digraph G {", "\n", sep = "")
  cat("node[shape=", shape, "];", "\n", sep = "")
  cat("rankdir=LR;", "\n")

  cat("edge[arrowhead=none];", "\n")

  # Create a dot node for each element in the lattice
  for (i in 1:nrow(S)) {
    x <- round(S$asset[i], digits = digits)
    y <- round(S$option[i], digits = digits)

    # Detect the American tree and draw accordingly
    early.exercise <- ""
    if (S$exercise[i] == "TRUE") {
      early.exercise <- "shape=box,"
```

```
    }
    if (i == 1) {
      cat("node", i, "[", early.exercise, "label=\"", x, ", ", "
         (", y, ")", "\"];", "\n", sep = "")
    } else{
      cat("node", i, "[", early.exercise, "label=\"", x, "\"];",
          "\n", sep = "")
    }
  }

  # The number of levels in a binomial lattice of length L is
     `$\frac{\sqrt{8N+1}-1}{2}$`
  L <- ((sqrt(8 * nrow(S) + 1) - 1) / 2 - 1)

  k <- 1
  for (i in 1:L) {
    tabs <- rep("\t", i - 1)
    j <- i
    while (j > 0) {
      cat("node", k, "->", "node", (k + i), ";\n", sep = "")
      cat("node", k, "->", "node", (k + i + 1), ";\n", sep = "")
      k <- k + 1
      j <- j - 1
    }
  }

  cat("}", sep = "")
}
#Plot the binomial tree of the Ameican all-or-nothing digital
   option in file lattice.dot.
treedot <- function(Asset, u, d, IntRate, Strike, NoSteps, n, s,
    B) {
  OptionType <- readline(prompt = "Enter option type: Call or
```

```
     Put:")
  if (OptionType == "Call") {
    x <-
      genlattice.vanilla.american.call(Asset = Asset, u = u, d =
          d, IntRate = IntRate, Strike = Strike, n = n, s = s,
         OptionType = OptionType, B = B)
  } else if (OptionType == "Put") {
    x <-
      genlattice.vanilla.american.put(Asset = Asset, u = u, d =
         d, IntRate = IntRate, Strike = Strike, NoSteps =
         NoSteps, n = n, s = s, OptionType = OptionType, B = B)
  }
  y <- capture.output(dotlattice(x, digits = 3))
  cat(y, file = "lattice.dot")
}
# Terminal: dot -Tpng -o binomial-tree-american-call-1.png /...
   the path where you save lattice.dot
```

B.5 Knock up-and-in barrier option

```
##This file is to generate the binomial tree of the knock up and
    in option.
##Use function treedot, where Asset is the initial stock price,
##u is the up factor, d is the down factor,
##IntRate is the discount rate,
##Strike is the strike price, NoSteps is the future time steps,
##n is the historical data, s is the successful historical data.
##B is the barrier.
##Then type in the storage path in the terminal with graphviz
   installed.

# The payoff for call and put knock up and in options.
```

```
payoff.vanilla.call <- function(Asset, Strike) {
  pay <- max(Asset - Strike, 0)
  return(pay)
}

payoff.vanilla.put <- function(Asset, Strike) {
  pay <- max(Strike - Asset, 0)
  return(pay)
}

# European option option pricing for the first node reaching the
    barrier.
EUoption <-
  function(S0, K, u, d, n, s, m, OptionType, IntRate,
     OptionPosition) {
    if (OptionType == "Call") {
      kstar <- (log(K) - log(S0) - m * log(d)) / (log(u) - log(d
         ))
      kstarPlus <- as.integer(kstar) + 1
      k <- c(kstarPlus:m)
      if (OptionPosition == "buy") {
        option.value <-
          choose(n + m, m) ^ (-1) * sum((u ^ k * d ^ (m - k) * S
             0 - K) * choose(s + k - 1, k) *choose(n - s + m - k
             , m - k)) * (1 + IntRate) ^ (-m)
      } else if (OptionPosition == "sell") {
        option.value <-
          choose(n + m, m) ^ (-1) * sum((u ^ k * d ^ (m - k) * S
             0 - K) * choose(s +k, k) *choose(n - s + m - k - 1,
              m - k)) * (1 + IntRate) ^ (-m)
      }
    } else if (OptionType == "Put") {
      kstar <- (log(K) - log(S0) - m * log(d)) / (log(u) - log(d
```

```
        ))
      kstarMinus <- as.integer(kstar)
      k <-  c(0:kstarMinus)
      if (OptionPosition == "buy") {
        option.value <-
          choose(n + m, m) ^ (-1) * sum((K - u ^ k * d ^ (m - k)
              * S0) * choose(s + k, k) *choose(n - s + m - k - 1
              , m - k)) * (1 + IntRate) ^ (-m)
      } else if (OptionPosition == "sell") {
        option.value <-
          choose(n + m, m) ^ (-1) * sum((K - u ^ k * d ^ (m - k)
              * S0) * choose(s +k - 1, k) *choose(n - s + m - k,
              m - k)) * (1 + IntRate) ^ (-m)
      }
    }
    return(option.value)
  }

# Generate a binomial lattice for knock up and in option.
genlattice.american <-
  function(Asset, u, d, IntRate, Strike, NoSteps, Payoff, n, s,
     OptionType, B) {
    OptionPosition <- readline(prompt = "Enter option investor's
        position: buy or sell:")
    # The number of tree nodes to process.
    count <- sum(1:(NoSteps + 1))

    # This data frame will store asset and option prices.
    # The mapping from tree node (i,j) to linear index
    # inside the data frame will have to be computed.
    # "option" stores the first node reaching the barrier
    X <- data.frame(matrix(NA, nrow = count, ncol = 3))
    names(X) <- c("asset", "option", "price")
```

```
# Option price discount factor. Here the IntRate is the
   expected return for only onestep
DiscountFactor <- 1 / (1 + IntRate)

#The NPI probability.
p <- 0
#Set up stock price tree
for (i in NoSteps:0) {
  for (j in i:0) {
    AssetCurrent <- Asset * u ^ (i - j) * d ^ j
    X$asset[count] <- AssetCurrent
    count <- count - 1
  }
}

#Decide the which node is the first node reaching the
   barrier.
l <- 0
for (i in 0:NoSteps) {
  for (j in l:i) {
    g <- (1 + i) * i / 2 + j + 1
    if (X$asset[g] >= B) {
      l = j + 1
      if (i == NoSteps) {
        if (OptionType == "Call") {
          X$option[g] <- payoff.vanilla.call(X$asset[g],
             Strike)
        } else if (OptionType == "Put") {
          X$option[g] <- payoff.vanilla.put(X$asset[g],
             Strike)
        }
      } else{
```

```
          X$option[g] <- EUoption( X$asset[g], Strike, u, d,
             n + i, s + i - j, NoSteps - i, OptionType,
             IntRate, OptionPosition)
        }
      } else{
        X$option[g] <- 0
      }
    }
    if (l == i + 1 && l != 0) {
      break
    }
  }
X$option[is.na(X$option)] <- 0
count <- sum(1:(NoSteps + 1))

# Compute the asset and option prices, starting from the
   last node of the tree, which is its bottom right corner
   when viewed as a graph.
# Work up and backwards.
for (i in NoSteps:0) {
  for (j in i:0) {
    if (i == NoSteps) {
      X$price[count]   <- X$option[count]
    } else {
      up   <- X$price[sum(1:(i + 1), j, 1)]
      down <- X$price[sum(1:(i + 1), j + 1, 1)]
      if ((OptionType == "Call" &
           OptionPosition == "buy") |
          (OptionType == "Put" & OptionPosition == "sell"))
             {
        p <- (s + i - j) / (n + i + 1)
      } else if ((OptionType == "Call" &
                  OptionPosition == "sell") |
```

```
                     (OptionType == "Put" & OptionPosition == "
                        buy")) {
            p <- (s + i - j + 1) / (n + i + 1)
          }
          #Justify the node is to be calculated as european
             option or not
          if (X$option[count] != 0) {
            X$price[count] <- X$option[count]
          } else{
            X$price[count] <- DiscountFactor * (p * up + (1 - p)
                * down)
          }
        }

        count <- count - 1
      }
    }

    return(X)
  }
#Genlattice family functions:
#different lattice by the option type:call or put
genlattice.vanilla.american.call <- function(Asset, u, d,
   IntRate, Strike, NoSteps, n, s, OptionType = OptionType, B) {
    return (
      genlattice.american( Asset = Asset, u = u, d = d, IntRate
         = IntRate, Strike = Strike, NoSteps = NoSteps, Payoff =
          payoff.vanilla.call, n = n, s = s, OptionType =
         OptionType, B = B)
    )
  }

genlattice.vanilla.american.put <-
```

```
  function(Asset, u, d, IntRate, Strike, NoSteps, n, s,
     OptionType = OptionType, B) {
    return (
      genlattice.american( Asset = Asset,  u = u, d = d,
         IntRate = IntRate,  Strike = Strike,  NoSteps = NoSteps
         , Payoff = payoff.vanilla.put, n = n, s = s, OptionType
          = OptionType, B = B)
    )
  }

# Generates a graph specification that can be fed into graphviz.
# Input: the binomial lattice produced
#by one of genlattice family functions.
dotlattice <- function(S, digits = 2) {
  shape <- "plaintext"

  cat("digraph G {", "\n", sep = "")
  cat("node[shape=", shape, "];", "\n", sep = "")
  cat("rankdir=LR;", "\n")

  cat("edge[arrowhead=none];", "\n")

  # Create a dot node for each element in the lattice
  for (i in 1:nrow(S)) {
    x <- round(S$asset[i], digits = digits)
    y <- round(S$price[i], digits = digits)

    # Detect the American tree and draw accordingly
    early.exercise <- ""
    if (S$option[i] != 0) {
      early.exercise <- "shape=box,"
    }
    if (i == 1) {
```

```
      cat("node", i, "[", early.exercise, "label=\"", x, ", ", "
         (", y, ")", "\"];", "\n", sep = "")
    } else{
      cat("node", i, "[", early.exercise, "label=\"", x, "\"];",
          "\n", sep = "")
    }
  }

  # The number of levels in a binomial lattice of length L is
     `$\frac{\sqrt{8N+1}-1}{2}$`
  L <- ((sqrt(8 * nrow(S) + 1) - 1) / 2 - 1)

  k <- 1
  for (i in 1:L) {
    tabs <- rep("\t", i - 1)
    j <- i
    while (j > 0) {
      cat("node", k, "->", "node", (k + i), ";\n", sep = "")
      cat("node", k, "->", "node", (k + i + 1), ";\n", sep = "")
      k <- k + 1
      j <- j - 1
    }
  }

  cat("}", sep = "")
}

#Plot the binomial tree of the knock up and in option in file
   lattice.dot.
treedot <- function(Asset, u, d, IntRate, Strike, NoSteps, n, s,
    B) {
  OptionType <- readline(prompt = "Enter option type: Call or
     Put:")
```

```
  if (OptionType == "Call") {
    x <- genlattice.vanilla.american.call( Asset = Asset, u = u,
        d = d, IntRate = IntRate, Strike = Strike, NoSteps =
       NoSteps, n = n, s = s, OptionType = OptionType, B = B)
  } else if (OptionType == "Put") {
    x <- genlattice.vanilla.american.put(Asset = Asset, u = u, d
        = d, IntRate = IntRate, Strike = Strike, NoSteps =
       NoSteps, n = n, s = s, OptionType = OptionType, B = B)
  }
  y <- capture.output(dotlattice(x, digits = 3))
  cat(y, file = "lattice.dot")
}
# Terminal: dot -Tpng -o binomial-tree-american-call-1.png /...
   the path where you save lattice.dot
```

B.6 Look-back call option

```
##This file is to generate the binomial tree of the look-back
   call option
##with float strike price.
##Use function treedot, where Asset is the initial stock price,
##u is the up factor,
##IntRate is the discount rate,
##NoSteps is the future time steps,
##n is the historical data, s is the successful historical data.
##Then type in the storage path in the terminal with graphviz
   installed.

#The value of $W_{T_j}(k)$
payoff.vanilla.call <- function(u, k) {
  pay <- 1 - u ^ (-k)
  return(pay)
```

```
}

# Generate a binomial lattice for the look-back call option.
genlattice.american <- function(u, IntRate, NoSteps, n, s) {
  OptionPosition <- readline(prompt = "Enter option investor's
     position: buy or sell:")
  # The number of tree nodes to process.
  count <- sum(1:(NoSteps + 1))

  # This data frame will store $W_{T_j}(k)$ as "asset"
  # and $W_{t_j}(k)$ as "option".
  # The mapping from tree node (i,j) to linear index
  # inside the data frame will have to be computed.
  # The time and $k$ are also stored.
  X <- data.frame(matrix(NA, nrow = count, ncol = 4))
  names(X) <- c("asset", "option", "time", "kvalue")
  # Option price discount factor.
  # Here the IntRate is the expected return for only onestep
  DiscountFactor <- 1 / (1 + IntRate)
  #The NPI probability.
  p <- 0
  # Compute the $W$ founction binomial tree, starting from the
     last node of the tree, which is its bottom right corner
     when viewed as a graph.
  # Work up and backwards.
  for (i in NoSteps:0) {
    for (j in i:0) {
      X$asset[count] <- 1 - u ^ (j - i)
      if (i == NoSteps) {
        X$option[count] <- X$asset[count]
      } else{
        if (j == i) {
          down <- X$option[sum(1:(i + 2))]
```

```
        } else{
          down <- X$option[sum(1:(i + 1), j + 1, 2)]
        }
        if (OptionPosition == "buy") {
          p <- (s + i - j) / (n + i + 1)
        } else if (OptionPosition == "sell") {
          p <- (s + i - j + 1) / (n + i + 1)
        }
        up <- X$option[sum(1:(1 + i), j, 1)]
        X$option[count] <- DiscountFactor * (p * up + (1 - p) *
           down)
      }

      X$time[count] <- i
      X$kvalue[count] <- i - j
      count <- count - 1
    }
  }
  return(X)
}

# Generates a graph specification that can be fed into graphviz.
dotlattice <- function(S, digits = 2) {
  shape <- "plaintext"

  cat("digraph G {", "\n", sep = "")
  cat("node[shape=", shape, "];", "\n", sep = "")
  cat("rankdir=LR;", "\n")

  cat("edge[arrowhead=none];", "\n")

  # Create a dot node for each element in the lattice
  for (i in 1:nrow(S)) {
```

```
    x <- round(S$option[i], digits = digits)
    y <- round(S$time[i], digits = digits)
    z <- round(S$kvalue[i], digits = digits)
    cat("node", i, "[", "label=\"", x, ", ", "(", z, ",", y, ")"
       , "\"];", "\n", sep = "")
  }

  # The number of levels in a binomial lattice of length L is
     `$\frac{\sqrt{8N+1}-1}{2}$`
  L <- ((sqrt(8 * nrow(S) + 1) - 1) / 2 - 1)

  k <- 1
  for (i in 1:L) {
    tabs <- rep("\t", i - 1)
    j <- i
    while (j > 0) {
      cat("node", k, "->", "node", (k + i), ";\n", sep = "")
      cat("node", k, "->", "node", (k + i + 1), ";\n", sep = "")
      k <- k + 1
      j <- j - 1
    }
  }

  cat("}", sep = "")
}

#Plot the binomial tree of thelook-back call option in file
   lattice.dot.
treedot <- function(u, IntRate, NoSteps, n, s) {
  x <- genlattice.american(u, IntRate, NoSteps, n, s)
  y <- capture.output(dotlattice(x, digits = 3))
  cat(y, file = "lattice.dot")
}
```

```
# Terminal: dot -Tpng -o binomial-tree-american-call-1.png /...
   the path where you save lattice.dot
```

Bibliography

[1] A.M. Aboalkhair. *Nonparametric Predictive Inference for System Reliability.* PhD thesis, University of Durham, 2012.

[2] K.I. Amin. Jump diffusion option valuation in discrete time. *The Journal of Finance*, 48:1833–1863, 1993.

[3] E. Appolloni, M. Gaudenzi, and A. Zanette. The binomial interpolated lattice method for step double barrier options. *International Journal of Theoretical and Applied Finance*, 17:1450035, 2014.

[4] T. Arnold and T.F. Crack. Option pricing in the real world: a generalized binomial model with applications to real options, 27 Dec 2000. http://hdl.handle.net/10523/1548.

[5] T. Augustin and F.P.A. Coolen. Nonparametric predictive inference and interval probability. *Journal of Statistical Planning and Inference*, 124:251–272, 2004.

[6] T. Augustin, F.P.A. Coolen, G. de Cooman, and M.C.M Troffaes. *Introduction to Imprecise Probability.* Wiley, 2014.

[7] S. Babbs. Binomial valuation of lookback options. *Journal of Economic Dynamics and Control*, 24:1499–1525, 2000.

[8] R.M. Baker, T.A. Coolen-Maturi, and F.P.A. Coolen. Nonparametric predictive inference for stock returns. *Journal of Applied Statistics*, 44(8):1333–1349, 2017.

[9] L. Bauwens and M. Lubrano. Bayesian option pricing using asymmetric garch models. *Journal of Empirical Finance*, 9:321–342, 2002.

[10] D.T. Beerden and R.H. Litzenberger. Prices of state-contingent claims implicit in option prices. *The Journal of Business*, 54:621–651, 1978.

[11] D. Berleant, L. Andrieu, J.P. Argaud, F. Barjon, M.P. Cheong, M. Dancre, G. Sheble, and C.C. Teoh. Portfolio management under epistemic uncertainty using stochastic dominance and information-gap theory. *International Journal of Approximate Reasoning*, 49:101–116, 2008.

[12] F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81:637–654, 1973.

[13] Z. Bodie, A. Kane, and A.J. Marcus. *Investments and Portfolio Management*. McGraw-Hill/Irwin, global edition edition, 2011.

[14] G. Boole. *An investigation of the laws of thought on which are founded the mathematical theories of logic and probabilities*. Cambridge University Press, 1854.

[15] P.P. Boyle. A lattice framework for option pricing with two state variables. *Journal of Financial and Quantitative Analysis*, 23:1–12, 1988.

[16] P.P. Boyle and A.L. Ananthanarayanan. The impact of variance estimation in option valuation models. *Journal of Financial Economics*, 5:375–387, 1997.

[17] P.P. Boyle and S.H. Lau. Bumping up against the barrier with the binomial method. *The Journal of Derivatives*, 1:6–14, 1994.

[18] M. Capiński. *Discrete Models of Financial Markets*. Cambridge : Cambridge University Press, 2012.

[19] J. Chen. *Contributions to Nonparametric Predictive Inference for Bernoulli data with application in finance (submit for examination).* PhD thesis, University of Durham, 2019.

[20] T.H.F. Cheuk and T.C.F. Vorst. Currency lookback options and observation frequency: a binomial approach. *Journal of International Money and Finance*, 16:173–187, 1997.

[21] F.P.A. Coolen. Low structure imprecise predictive inference for bayes' problem. *Statistics and Probability Letters*, 36:349–357, 1998.

[22] F.P.A. Coolen. On nonparametric predictive inference and objective bayesianism. *Journal of Logic, Language and Information*, 15:21–47, 2006.

[23] F.P.A. Coolen. *Nonparametric Predictive Inference in : International encyclopedia of statistical science.* M.Lovric,ed., Springer, Berlin, 2011.

[24] F.P.A. Coolen, T. Coolen-Maturi, and A.H. Al-nefaiee. Nonparametric predictive inference for system reliability using the survival signature. *Journal of Risk and Reliability*, 42:437–448, 2014.

[25] F.P.A. Coolen, P. Coolen-Schrijner, T. Coolen-Maturi, and F.F. Elkhafifi. Nonparametric predictive inference for ordinal data. *Communications in Statistics - Theory and Methods*, 42:3478–3496, 2013.

[26] T. Coolen-Maturi and F.P.A. Coolen. Nonparametric predictive inference for the validation of credit rating systems. *Journal of the Royal Statistical Society: Series A, to appear*, 2018.

[27] P. Coolen-Schrijner and F.P.A. Coolen. Nonparametric adaptive age replacement with a one-cycle criterion. *Reliability Engineering and System Safety*, 92:74–84, 2007.

[28] P. Coolen-Schrijner, S.C. Shaw, and F.P.A. Coolen. Opportunity-based age replacement with a one-cycle criterion. *Journal of the Operational Research Society*, 60:1428–1438, 2009.

[29] J. Cox and M. Rubinstein. *Options Markets.* Prentice–Hall, New Jersey, 1985.

[30] J.C. Cox, S.A. Ross, and M. Rubinstein. Option pricing: A simplified approach. *Journal of Financial Economics*, 7:229–263, 1979.

[31] B. De Finetti. *Theory of Probability.* Wiley, London, 1974.

[32] J. de la Vega. *Confusion de Confusions.* reprinted in M.Fridson(ed.)(1996), 1688.

[33] I. de Pinto. *An essay on Circulation of currency and credit in four parts and a letter on the jealousy of commerce. translated with annotions by S. Baggs(1774), London.* reprinted by Gregg International Publishers (1969), 1771.

[34] H. Emery. *Speculation on the stock and produce exchanges of the United States.* reprinted by AMS Press, 1896.

[35] M. Gaudenzi and M.A. Lepellere. Pricing and hedging american barrier options by a modified binomial method. *International Journal of Theoretical and Applied Finance*, 9:533–553, 2006.

[36] O. Gelderblom and J. Jonker. *Amsterdam as the Cradle of Modern Futures and Options Trading, 1550-1630, chapter 11 in W. Goetsmann and K. Rouwenhorst (eds.).* The Origins of Value, 2005.

[37] A.V. Gerbessiotis. Architecture independent parallel binomial tree option price valuations. *Journal Parallel Computing*, 30:301–316, 2004.

[38] R. Geske and R. Roll. On valuing american call options with the black-scholes european formula. *International Journal of Approximate Reasoning*, XXXIX:443–455, 1984.

[39] M. Goldman, H. Sosin, and M. Gatto. Path dependent options: Buy at the low, sell at the high. *Journal of Finance*, 34:1111–1128, 1979.

[40] T. He, F.P.A. Coolen, and T. Coolen-Maturi. Nonparametric predictive inference for european option pricing based on the binomial tree model. *Journal of the Operational Research Society*, 2019. https://doi.org/10.1080/01605682.2018.1495997.

[41] B.M. Hill. Posterior distribution of percentiles: Bayes' theorem for sampling from a population. *Journal of the American Statistical Association*, 63:677–691, 1968.

[42] B.M. Hill. De Finetti's theorem, induction, and A(n) or Bayesian nonparametric predictive inference (with discussion). In J.M. Bernardo, et al. (eds.). *Bayesian Statistics*, 3:Bayesian Statistics 3, 211–241, 1988.

[43] X. Hu and J. Cao. Randomized binomial tree and pricing of american-style options. *Mathematical Problems in Engineering*, 2014:1563–5147, 2014.

[44] J Hull and A. White. Efficient procedures for valuing european and american path-dependent options. *Journal of Derivatives*, 1(1):21–31, 1993.

[45] J.C. Hull. *Options, futures and other derivatives*. Pearson Education, 2009.

[46] J.C. Jackwerth and M. Rubinstein. Recovering probability distributions from option prices. *The Journal of Finance*, 51:1611–1631, 1996.

[47] E. Jacquier and R. Jarrow. Bayesian analysis of contingent claim model error. *Journal of Econometrics*, 94:145–180, 2000.

[48] J.Y. Jaffray and M. Jeleva. Information processing under imprecise risk with an insurance demand illustration. *International Journal of Approximate Reasoning*, 49:117–129, 2008.

[49] R. Jarrow and A. Rudd. *Option pricing.* Down Jones-Irwin Publishing, Homewood, IL, 1983.

[50] D. Ji and B. Brorsen. A recombining lattice option pricing model that relaxes the assumption of lognormality. *Review of Derivatives Research*, 14:349–367, 2011.

[51] K.I. Kim, H.S. Park, and X. Qian. A mathematical modeling for the lookback option with jump–diffusion using binomial tree method. *Journal of Computational and Applied Mathematics*, 235:5140–5154, 2011.

[52] K.I. Kim and X. Qian. Convergence of the binomial tree method for asian options in jump-diffusion models. *Journal of Mathematical Analysis and Applications*, 330:10–23, 2007.

[53] Y.S. Kima, S. Stoyanov, S. Rachev, and F. Fabozzi. Multi-purpose binomial model: Fitting all moments to the underlying geometric brownian motion. *Economics Letters*, 145:225–229, 2016.

[54] R.C. Klemkosky and B.G. Resnick. Put-call parity and market efficiency. *The Journal of Finance*, 34:1141–1155, 1979.

[55] G.M. Martin, C.S. Forbes, and Martin V.L. Implicit bayesian inference using option prices. *Journal of Time Series Analysis*, 26:437–462, 2005.

[56] T.A. Maturi. *Nonparametric Predictive Inference for multiple comparisons.* PhD thesis, University of Durham, 2010.

[57] R.C. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4:141–183, 1973.

[58] S. Muzzioli and H. Reynaerts. American option pricing with imprecise risk-neutral probabilities. *International Journal of Approximate Reasoning*, 49:140–147, 2008.

[59] S.C. Myers. Finance theory and financial strategy. *Interfaces*, 14:126–137, 1984.

[60] I. Newton. *Methodus fluxionum et serierum infinitarum.* 1664 - 1671.

[61] S. Palan. Digital options and efficiency in experimental asset markets. *Journal of Economic Behavior and Organization*, 75:506–522, 2010.

[62] H.S. Park. Analytical binomial lookback options with double-exponential jumps. *Journal of the Korean Statistical Society*, 38:397–404, 2009.

[63] G. Poitras. The early history of option contracts. in: Hafner w., zimmermann h. (eds) vinzenz bronzin's option pricing models. 2009.

[64] N.G. Polson and J.R. Stroud. Bayesian inference for derivative prices. *Bayesian Statistics*, 7:641–650, 2002.

[65] X. Qian, C. Xu, L. Jiang, and B. Bian. Convergence of the binomial tree method for american options in jump-diffusion models. *SIAM Journal on Numerical Analysis*, 42:1899–1913, 2005.

[66] J. Raphson. *Analysis aequationum universalis.* London, 1690.

[67] M. Reimer and K. Sandmann. A discrete time approach for european and american barrier options. *Discussion Paper, Murray State University*, SFB 303(No. B-272), 1996.

[68] M. Rubinstein. Exotic Options. `http://www.haas.berkeley.edu/groups/finance/WP/rpf220.pdf`, 1990.

[69] M. Rubinstein and E. Reiner. Breaking down the barriers. *RISK*, 4:28–35, 1991.

[70] A. Thavaneswaran, S.S. Appadoo, and J. Frank. Binary option pricing using fuzzy numbers. *Applied Mathematics Letters*, 26:65–72, 2013.

[71] Y. Tian. A modified lattice approach to option pricing. *Journal of Futures Markets*, 13:563–577, 1993.

[72] Y. Tian. A flexible binomial option pricing model. *Journal of Futures Markets*, 19:817–843, 1999.

[73] P. Walley. *Statistical reasoning with imprecise probabilities.* Chapman and Hall, London, 1991.

[74] K. Weichselberger. The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning*, 24(2-3):149–170, 2000.

[75] K. Weichselberger. *Elementare Grundbegriffe einer Allgemeineren Wahrscheinlichkeitsrechnung I. Intervallwahrscheinlichkeit als Umfassendes Konzept.* Physika, Heidelberg, 2001.

[76] F. William and D. Turner. The evolution of a market. *In Managing Energy Price Risk*, London: Risk Books, 1999.

[77] Z. Zdenek. Generalised soft binomial american real option pricing model (fuzzy–stochastic approach). *European Journal of Operational Research*, 207:1096–1103, 2010.