A PANDAS-PYTHON

Special Report of Stock-Option Pricing

on U.S. DRUG Companies relative to S&P500 (SPY)

Columbia University\*

Dec 11th, 2019

Piruz Alemi \* Edward Joseph \* Martha Karam

\*Submitted to Ilya Altshteyn, Terry Huang, Ju Chen Ta, Jason Ta

Note: *Sections on modification & enhancement of R. Cox, Rubenstein Binomial Model is proprietary information of its authors and cannot be shared in public. These sections will be uploaded to a private repository rather than public.*

A useful and very popular technique for pricing a stock option involves constructing a binomial tree. This is a diagram that represents different possible paths that might be followed by the stock price over the life of the option.

In this report and its implementation in a PANDAS-PYTHON program, the authors pioneer a python program that looks at the binomial trees and their relationship to important principals in risk-neutral valuation.

The approach we adopt here is based on Cox, Ross and Rubinstein in 1979 – the basics of which has not changed since 1979. Once we operationalize the Option Pricing model, we then enhanced the original model, by changing certain parameters, provided in a black box – as it remains a proprietary information of those who help us enhance the model.

**The Core Code:[[1]](#footnote-1)**

Step0: Go to: <https://finance.yahoo.com/quote/SPY?p=SPY>

      if you go to finance.yahoo.com, then on the top right corner see: Quote Look Up, enter SPY.

1. Build a tree of probable levels of stocks going up or down or not moving at all. This gives a tree on its first step consisting of stock prices starting today at stock prices on say $314.04 for SPY, and going up in LEVELS by .01, .02, .03, or down by -.01, -.02, -.03 or same as before at zero.

     This is the tree of hypothetical stocks. Plot this chart #1

1. We take say n days of trading data, in our case n=7 or more accurately starting on

Monday and ending on Friday, which gives us 5 days of data, for which we would say

buy a Call Option.

The probabilities of stock-level-tree in Step1 are now calculated: This is the probability of reaching a node from any directions. In a two step model (Monday, Tuesday, Wed) these would be 0.25, 0.5, and 0.25, for the highest (up, up), mid-point (up-down, down-up) and lowest node (down, down).

 Say we calculate all the probabilities of reaching the nodes resting on Friday (after 4 periods). We derive for Friday all the probabilities for all the Friday-nodes on the binomial tree.

Plot Chart #2.

1. For a given Strike-Price (strike less of hypothetical stock prices on the tree) an American

Option Contract for Friday is derived, The option price (profit), according to the tree in Steps 1+2 is (StrikePrice – StockPrice) for each of the nodes present.

1. The weighted sum of the Option profits of the last period (Froiday), where the weights are probabilities of each node must be the fair-value of the Call option.

If StockPrice > StrikePrice, then Option = 0 i.e. the Call-Option-Value is Max (Strike-StockPrice,0) for the call option.

1. We take VIX as input to the equations in Hull (Chapter 10. Section 10.7 & 10.8), matching volatility with u and d. See equations 10.11, 10.12

A screenshot of a cell phone

Description automatically generated

1. Thus we understand the u/d (Up/Down) ratios and probability of up (p) and down (1-p) are contingent on yield, sigma(volatility) and time. We derive all these values for all the nodes of the binomial trees constructed along the basis of R.Cox, Rubenstein Model of 1979. Part of the Node(Object), is defined in a Class as

class Node():

# ------------------------------------------------------------------------------------------

# Node simulates a tree of Nodes. On each path of possible Trajectory of Stock prices we build

# A node [A rocket], whose attributes are subject to change along that trajectory.

# Each path has its own Node [Rocket]. A fleet of these Nodes [Rockets] are initiated at the start

# that can move up or down in accordance with the permutation of up/down paths.

# How the Node moves also affects its internal(self) values....

# https://www.macrotrends.net/2577/sp-500-pe-ratio-price-to-earnings-chart

# ------------------------------------------------------------------------------------------

def \_\_init\_\_(self, x=0,y=1,level=0):

# Each node has an (x,y) position and contains a stock Value.

self.x = x

self.y = y

self.level=0

self.stockValue=312

self.strikePrice=314

def setup\_parameters(self,rfr=0.03,vix\_sigma=0.12,dt=.0197):

self.sigma=vix\_sigma

self.r=rfr

self.dt=dt

# See equation 10.11 in Hull. The stock price either moves up by u or down by d proportions

# where u/d = 1. This proprtionality is made an Exp function of volatility of stock sigma & time.

self.u = math.exp(self.sigma \* math.sqrt(self.dt))

#self.u = 0.5

self.d = 1./self.u

# See equation 10.9 in Hull. To match the expected return on Stock u, with the trees parameters...

# Probability of Node going up = qu and going down = qd

self.qu = ((math.exp((self.r \* self.dt))-self.d))/(self.u-self.d)

self.qd = 1-self.qu

def move\_up(self, y\_up=1.1, level=+1, rfr=0.03, vix\_sigma=0.12, dt=.0197):

# ------------------------------------------------------------------------------------

# 1. Increment the level-position of the Node, each time you move-up

# 2. lower the probability of going up by qu \* qu

# 3. Move the node to the right, as to capture the trajectory of the node on a given path

# -------------------------------------------------------------------------------------

self.sigma = vix\_sigma

self.r = rfr

self.dt = dt

self.u = math.exp(self.sigma \* math.sqrt(self.dt))

self.d = 1./self.u

self.qu = ((math.exp((self.r \* self.dt))-self.d))/(self.u-self.d)

self.qu = self.qu \* self.qu

#-----------------------------------------------

self.y = y\_up \* self.y

self.level += level

self.factor=y\_up

#self.stockValue = self.factor \* self.stockValue

self.stockValue = self.u \* self.stockValue

#self.qu=self.qu \* self.qu

self.optionValue=max(self.stockValue - self.strikePrice,0)

self.x += 1

def move\_down(self, y\_down=0.90,level=-1,rfr=0.03,vix\_sigma=0.12,dt=.0197):

# Increment the y-position of the Node.

self.sigma=vix\_sigma

self.r=rfr

self.dt=dt

self.u = math.exp(self.sigma \* math.sqrt(self.dt))

self.d = 1./self.u

self.qu = ((math.exp((self.r \* self.dt))-self.d))/(self.u-self.d)

self.qd = 1-self.qu

self.qd = self.qd \* self.qd

#-------------------------------------------------

self.level += level

self.y = y\_down \* self.y

self.factor=y\_down

#self.stockValue = self.factor \* self.stockValue

self.stockValue = self.d \* self.stockValue

self.optionValue=max(self.stockValue - self.strikePrice,0)

#self.qd=self.qd \* self.qd

self.x += 1

def move\_right(self):

# Increment the x-position of the Node.

self.x += 1

def move\_left(self):

# Decrease the x-position of the Node.

self.x += -1

def move\_node(self,time\_position=0,ud\_pos=0):

self.x += time\_position

self.y += ud\_pos

def outputDetail(self):

return [self.x,self.level,self.stockValue,self.optionValue,self.u,self.qd]

1. when binomial trees are used in practice, the life of the option is typically divided into 30 or more time steps of length σt. In each time step there is a binomial stock price movement. With 30 time steps, this means that 31 terminal stock prices and 230 or about 1 billion possible stocks price paths to be considered

For a given contract, say a call option to be exercised on Dec 16th, 2019, from Finance.Yahoo.com Call option strike prices are given. Let us assume we select for Call Option with a Strike Price of $314, today on Dec 10th, 2019. days later, if current Stock = 313.58?

A screenshot of a cell phone

Description automatically generatedA close up of a logo

Description automatically generatedA screenshot of a cell phone

Description automatically generated

Note for the Python program we have the following:

Inputs:

1. API-Code to extract the data from Finance.Yahoo.com
2. Date the contract is exercised.
3. Strike-Price
4. Current Stock-Price
5. Options Bid Price (In this case, as we have a Call Option)
6. Expected Up/Down Levels (e.g., 1.1, 0, - 0.9)
7. Probabilities of Up/Down (e.g. 0.5).

Outputs:

1. Derived Option Price
2. Stock Price Variance
3. Stock Price Implied Variance
4. Tree Distribution of Stock Prices
5. Z-Values for Confidence levels
6. Plots of the distribution of Stock Prices on the Tree
7. Plot of Hedging share, for Rebalancing the portfolio

1. Check if the derived option is greater or less than the bid-price of option and by how much?

We arrived first at the following binomial tree of Stock-Options

A picture containing sky, text, indoor

Description automatically generated

Which given our inputs of Time, Strike Price, Possible Stock Prices, Risk Free Rate, Up/Down ratios

And derived probabilities gives us the following chart for Option Values. Interestingly, for the forward movement, the Option values are independent of probabilities, if we follow R. Cox, Rubenstein Model of Binomial Trees. However traversing the option tree backward, having derived the option value of day 5, we derive the value of option in day 4, as the weighted average of two nodes in day 5, weighted by the probabilities of up/down ratios. As it follows where rfr = Risk Free Rate, dt = Days as a portion of annual days traded, like 5/253. The up-ratio now has become a function of sigma (variations) and time, rather than the standard fixed portion of 1.1 up and 0.9 down. Thus we derived the probability of the node going up as qu and going down as qd:

self.sigma=vix\_sigma

self.r=rfr

self.dt=dt

self.u = math.exp(self.sigma \* math.sqrt(self.dt))

self.d = 1./self.u

self.qu = ((math.exp((self.r \* self.dt))-self.d))/(self.u-self.d)

self.qd = 1-self.qu

A picture containing text

Description automatically generated

**PART II**

**Part II – Option Pricing & Health Systems**

In part two, instead of SPY index we focus on a particular Corporations Stock and Options. In this case a nimber of drug companies as well as IHE. We are interested in the impact of a news (e.g. Drug Discovery) positive or negative on both the Stock and Options prices and weather the Drug-Stocks follows the SPY index

In particular we asked the following questions:

How tied is Each HealthCare Stock in Our Study to the S&P 500 Market Index?

1. Gilead is slightly negatively correlated (-.10) with the S&P 500. The R-Squared is .167. This shows a weak relationship between the Stock and the S&P 500.
2. Johnson & Johnson is slightly positively correlated (.12) with the S&P 500. The R-Squared is .108. This is a fairly weak relationship between the stock and the S&P 500.
3. Ely Lily is positively correlated (.22) with the S&P 500. The R-Squared is .131. This shows a weak  relationship between the stock and the market index.
4. Pfizer is moderately negatively correlated (-.49) with the S&P 500. The R-Squared is .32. This is a moderate negative relationship. However, this may be a coincidence, and we may not want to draw the conclusion of a negative relationship from it.
5. Stryker is highly positively correlated (.85) with the S&P 500. The R-Squared is .668. This shows a strong positive relationship between the stock and the S&P 500.

How Correlated is Each HealthCare Stock in Our Study To iShares U.S. Pharmaceuticals ETF(IHE)?

1. Gilead is positively correlated (.50) with the iShares ETF. The R-Squared is .883. This shows that there is a positive relationship between the stock and the etf.
2. Johnson & Johnson is strongly positively correlated (.78) with the iShares ETF. The R-Squared is .94. This shows that there is a strong positive relationship between the stock and the etf.
3. Ely Lily is strongly positively correlated (.72) with the iShares ETF. The R-Squared is .416. This shows that there is a strong positive relationship between the stock and the etf.
4. Pfizer is positively correlated (.66) with the iShares ETF. The R-Squared is .878. This shows that there is a positive relationship between the stock and the etf.
5. Stryker is slightly negatively correlated (.17) with the iShares ETF. The R-Squared is .557. This shows a weak relationship between the stock and the iShares ETF.

These correlations show that the 5 pharmaceuticals companies we looked at are not that strongly correlated with the S&P 500. However, Stryker, a medical devices company is strongly correlated with the S&P. This shows that pharma companies do not trade with the market and are not strongly influenced by the market, however this is not true for all healthcare companies as we can see that Stryker does trade with the market and may actually be influenced by it.

1. Clean the historical data from NaN values
2. Set all historical data to same frequency (daily).
3. Identify the spikes for a given spread of the historical data. For SPY use VIX.
4. Identify where the Variance is not constant, say low, medium, high variance periods
5. Identify where the Volume of Trade is not constant, say it is low, medium, high volume periods
6. Identify the time period of the news announced for positive / negative news.
7. For the given news indicator, identify if the stock price probability of Up or Down on Johnson & Johnson has changes from 50-50 to another ratio
8. Plot the results
9. Feed the results of part II back into Part I of the model.
10. Report the result of new discovery of the drug on the Binomial model and the actual Stock price.

**Part III Future Plan of Actions:**

We intend to enhance this project further as:

1. Automate the process
2. Read the data on-line, rather than off-line
3. Identify the inputs that can converge the Binomial Models results towards the Actual Stock Price. These input variables include:
   1. Tree Levels Up/Down Ratios
   2. The probability of up/down
   3. Implied volatility vs. Actual
   4. The shape of the “normal” distribution and whether the z-values on confidence levels need to be changed
   5. Whether the markets Call/Put Ratios could affect the convergence
   6. Whether external variables as opposed to Internal variables could have affected the convergence/divergence between the actual and the model
   7. Whether the SEM of the model through time, is random or is heteroskedastic?

**Sources:**

John C Cox, Stephen Ross, Mark. Rubenstein 1979. Option Pricing: A simplified Approach.

<http://static.stevereads.com/papers_to_read/option_pricing_a_simplified_approach.pdf>

John C. Hull 2003, Derivatives Options, Futures and Other Derivatives, Prentice Hall

<https://fac.ksu.edu.sa/sites/default/files/options_futures_and_other_derivatives_5th_ed.pdf>

Feller W. 1950 Probability Theory and its Applications Voil 1, and Vol 2. Wiley, New York.

<https://www.ime.usp.br/~fmachado/MAE5709/FellerV1.pdf>

1. I thank Victor Kitaigorodsky, PhD in math from Moscow University for his generous help, so I may understand how to write the Core Code section of the Binomial Model. Piruz Alemi. Nov 27th, 2019. [↑](#footnote-ref-1)