Ch 4. Binomial Tree Model

I. One-Period Binomial Tree

II. CRR Binomial Tree Model

III. Estimation and Calibration of μ and σ

IV. Dividends and Option Pricing

V. Introduction of Combinatorial Method

Appendix A. Binominal Tree Model for Jump-Diffusion Processes

- This chapter is devoted to introduce the binomial tree model, which is also known as a kind of lattice model. The lattice models, such as the binomial tree model introduced in this chapter or the finite difference method introduced in the next chapter, are popular numerical methods for option pricing, particularly for pricing American-style derivatives. They are also flexible since only nominal changes of the payoff function are needed for dealing with pricing complex, nonstandard options.
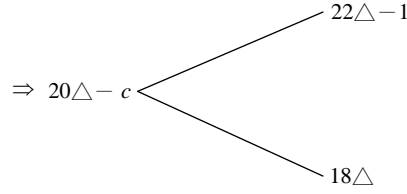
I. One-Period Binomial Tree

Figure 4-1
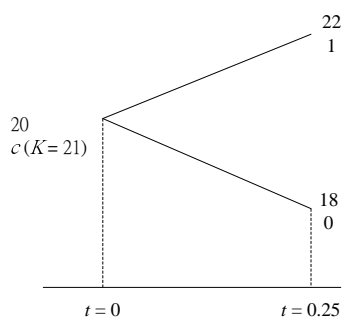
(i) Constructing a portfolio: long Δ shares and short 1 call

Figure 4-2

(ii) Deciding the value of Δ: if $22\Delta - 1 = 18\Delta \Rightarrow$ portfolio is riskless
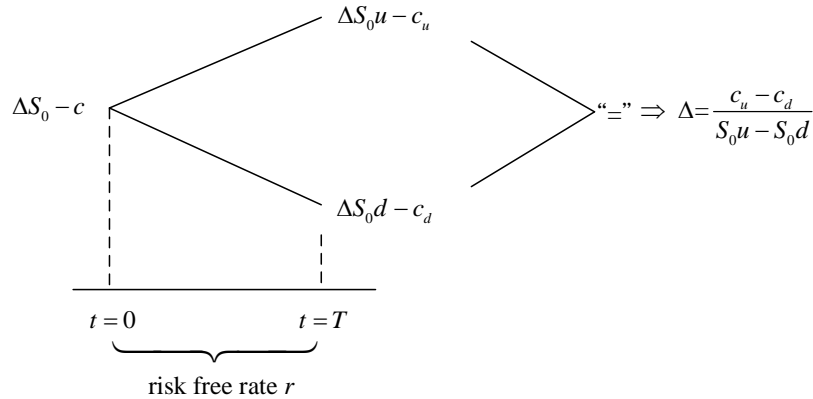
$$\Rightarrow 4\Delta = 1$$

$$\Delta = 0.25$$

(iii) Hence, at $t = 0.25$, the value of the portfolio is $22 \cdot 0.25 - 1 = 18 \cdot 0.25 = 4.5$

$$\Rightarrow 20\Delta - c = 4.5 \cdot e^{-r \cdot 0.25}, \text{ where } \Delta = 0.25 \text{ and } r = 0.12$$

$$\Rightarrow c = 0.633$$

* The emergence of the risk free interest rate r is due to the no-arbitrage argument.

Figure 4-3

$$\Delta = \frac{c_u - c_d}{S_0 u - S_0 d}$$

$$\Rightarrow \frac{c_u - c_d}{S_0 u - S_0 d}S_0 - c = \left(S_0 u \frac{c_u - c_d}{S_0 u - S_0 d} - c_u\right)e^{-rT}$$

$$\Rightarrow c = \frac{c_u - c_d}{u - d} - \left(\frac{c_u - c_d}{u - d}u - c_u\right)e^{-rT}$$

$$\Rightarrow c = e^{-rT}\left(\frac{c_u - c_d}{u - d}e^{rT} - \frac{(c_u - c_d)u}{u - d} + \frac{(u - d)}{u - d}c_u\right)$$

$$\Rightarrow c = e^{-rT}\left(\frac{(c_u - c_d)e^{rT} - c_u u + c_d u + u c_u - d c_u + c_d d - c_d d}{u - d}\right)$$

$$= e^{-rT}\left(\frac{(e^{rT} - d)c_u}{u - d} + \frac{(u - d)c_d}{u - d} - \frac{(e^{rT} - d)c_d}{u - d}\right)$$

$$= e^{-rT}\left(\frac{e^{rT} - d}{u - d} \cdot c_u + \left(1 - \frac{e^{rT} - d}{u - d}\right) \cdot c_d\right)$$

$\left(\frac{e^{rT} - d}{u - d} \text{ and } 1 - \frac{e^{rT} - d}{u - d} \text{ are like binomial probabilities, so they are denoted as } p \text{ and } 1 - p.\right)$

$$= e^{-rT}(p \cdot c_u + (1-p) \cdot c_d)$$

∗ For different options, the above equation remains valid, but different payoffs c_u and c_d should be considered.

- It is worth noting that p is not the probability for c_u (or for the upward movement of S). However, p can be regarded as the probability for c_u (or for the upward movement of S) in the risk neutral world. This is because in the real world, if the expected stock return is μ,

$$S_0 e^{\mu T} = S_0 u \cdot q + S_0 d \cdot (1-q) \Rightarrow q = \frac{e^{\mu T} - d}{u - d}.$$

Similarly, in the risk neutral world, since the expected returns for all securities are the same and equal to r,

$$S_0 e^{rT} = S_0 u \cdot p + S_0 d \cdot (1-p) \Rightarrow p = \frac{e^{rT} - d}{u - d}.$$

Therefore, p and $1-p$ are termed as risk neutral probabilities in the binomial tree framework.

- The option pricing equation $c = e^{-rT}(p \cdot c_u + (1-p) \cdot c_d)$ in the binomial tree model is consistent with the RNVR because both the expected growth rate of the underlying asset and the discount rate of the option payoff are the risk free rate.

- If the upward and dowanward probabilities in the real world are considered, it is difficult to identify a proper discount rate to discount the expected payoff of options, i.e.,

$$c = e^{-?T} \cdot (q \cdot c_u + (1-q) \cdot c_d).$$

  In practice, for securities with more risk or uncertainty, it should apply higher discount rates to future expected payoffs. Furthermore, it is well-known that options are riskier than their underlying assets due to the high-leverage characteristic for investing in options. So, it is not suited to use the expected return for the underlying asset, μ, to discount the expected payoff of the option.

  If we reconsider the above numerical example of the one-period binomial tree and assume $\mu = 16\%$, then we can derive the probability q in the real world to be 0.7041. Next by equalizing $e^{-?T} \cdot (q \cdot c_u + (1-q) \cdot c_d)$ to be 0.633, which is the true option value in the numerical example, we can derive the discount rate for the option to be 42.58%.

  In fact, the proper discount rates for expected payoffs of options depend not only on the expected returns (μ) and volatilities (σ) of underlying assets, but also on the different K and T of options. As a consequence, it is almost impossible to derive theoretical option prices directly in the real world.

II. CRR Binomial Tree Model

- Lognormal property

  If X is lognormally distributed, i.e., $\ln X$ follows a normal distribution with mean $= E[\ln X]$ and variance $= \text{var}(\ln X)$, then $E[X] = e^{E[\ln X] + \frac{1}{2} \cdot \text{var}(\ln X)}$, and $\text{var}(X) = e^{2 \cdot E[\ln X] + \text{var}(\ln X)} \cdot (e^{\text{var}(\ln X)} - 1)$.

  $\because \ln S_T \sim ND(\ln S_0 + (\mu - \frac{\sigma^2}{2})T, \sigma^2 T)$

  $\Rightarrow E[S_T] = S_0 e^{\mu T}$, and $\text{var}(S_T) = S_0^2 e^{2\mu T}(e^{\sigma^2 T} - 1)$

  $\left\|\begin{array}{l} \because \ln S_{t+\Delta t} \sim ND(\ln S_t + (\mu - \frac{\sigma^2}{2})\Delta t, \sigma^2 \Delta t) \\ \Rightarrow E[S_{t+\Delta t}] = S_t e^{\mu \Delta t} \\ \Rightarrow \text{var}(S_{t+\Delta t}) = S_t^2 e^{2\mu \Delta t}(e^{\sigma^2 \Delta t} - 1) \\ \qquad \approx S_t^2(1 + 2\mu\Delta t)(1 + \sigma^2 \Delta t - 1) \text{ (because } e^x \approx 1 + x \text{ when } x \to 0) \\ \qquad = S_t^2 \sigma^2 \Delta t + S_t^2(2\mu\Delta t)(\sigma^2 \Delta t) \\ \qquad \approx S_t^2 \sigma^2 \Delta t \text{ (because the term with } \Delta t^2 \text{ is relatively too small)} \\ \text{(In the next paragraph, I will use this property to derive the parameters } u, d, \text{ and } p \\ \text{in the CRR binomial tree framework.)} \end{array}\right.$

- Deriving u, d, and p in the CRR (Cox, Ross, and Rubinstein (1979)) binomial tree model, which is the most common and famous binomial tree model.

  Figure 4-4
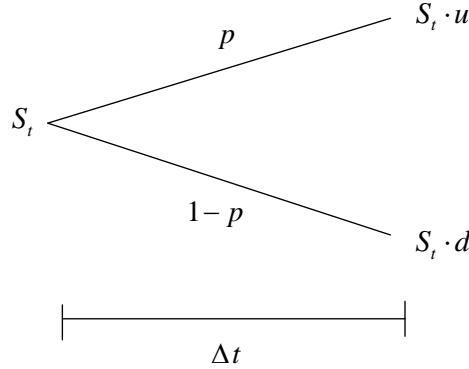
  (i) Matching mean: $\quad p \cdot S_t u + (1-p) \cdot S_t d = E[S_{t+\Delta t}] = S_t e^{r\Delta t}$

  $\qquad \Rightarrow p = \frac{e^{r\Delta t} - d}{u - d}$

(ii) Matching variance: $\text{var}(S_{t+\Delta t}) = E[S_{t+\Delta t}^2] - E[S_{t+\Delta t}]^2$

$$\Rightarrow \sigma^2 \Delta t = p \cdot u^2 + (1-p) \cdot d^2 - [p \cdot u + (1-p) \cdot d]^2 \text{ (both sides } \times (1/S_t)^2)$$

$$= p \cdot u^2 + (1-p) \cdot d^2 - p^2 \cdot u^2 - 2 \cdot p \cdot (1-p) \cdot u \cdot d - (1-p)^2 \cdot d^2$$

$$= u^2 \cdot (p - p^2) + [(1-p) - (1-p)^2] \cdot d^2 - 2 \cdot p \cdot (1-p) \cdot u \cdot d$$

$$= u^2 \cdot p \cdot (1-p) + (1-p) \cdot [1 - (1-p)] \cdot d^2 - 2 \cdot p \cdot (1-p) \cdot u \cdot d$$

$$= p \cdot (1-p) \cdot [u^2 - 2 \cdot u \cdot d + d^2] = p \cdot (1-p) \cdot (u-d)^2$$

$$\left\Vert \begin{array}{rl} p \cdot (1-p) &= p - p^2 = \frac{e^{r\Delta t} - d}{u - d} - \frac{e^{2r\Delta t} - 2 \cdot d \cdot e^{r\Delta t} + d^2}{(u-d)^2} \\[2ex] &= \frac{e^{r\Delta t} \cdot u - u \cdot d - e^{r\Delta t} \cdot d + d^2 - e^{2r\Delta t} + 2 \cdot d \cdot e^{r\Delta t} - d^2}{(u-d)^2} \\[2ex] &= \frac{e^{r\Delta t}(u - d + 2 \cdot d) - u \cdot d - e^{2r\Delta t}}{(u-d)^2} = \frac{e^{r\Delta t}(u+d) - u \cdot d - e^{2r\Delta t}}{(u-d)^2} \end{array} \right.$$

$$\Rightarrow \sigma^2 \Delta t = e^{r\Delta t}(u+d) - u \cdot d - e^{2r\Delta t}$$

Define $d = \frac{1}{u}$,

$$\Rightarrow \sigma^2 \Delta t = e^{r\Delta t}(u + \tfrac{1}{u}) - u \cdot \tfrac{1}{u} - e^{2r\Delta t}$$

$$\Rightarrow u + \tfrac{1}{u} = \frac{\sigma^2 \Delta t + 1 + e^{2r\Delta t}}{e^{r\Delta t}} = e^{-r\Delta t}\sigma^2 \Delta t + e^{-r\Delta t} + e^{r\Delta t}$$

$$\left\Vert \; e^{-r\Delta t} \approx (1 - r\Delta t),\; e^{r\Delta t} \approx (1 + r\Delta t), \text{ and } r \cdot \sigma^2 \cdot \Delta t^2 \to 0 \right.$$

$$\approx \sigma^2 \Delta t + 2$$

$$\Rightarrow u^2 - (\sigma^2 \Delta t + 2)u + 1 = 0$$

$$u = \frac{\sigma^2 \Delta t + 2 \pm \sqrt{(\sigma^2 \Delta t + 2)^2 - 4}}{2} = \frac{\sigma^2 \Delta t + 2 \pm \sqrt{\sigma^4 \Delta t^2 + 4\sigma^2 \Delta t + 4 - 4}}{2} \quad (\sigma^4 \Delta t^2 \to 0)$$

$$= \frac{\sigma^2 \Delta t}{2} + 1 \pm \sigma\sqrt{\Delta t}$$

$$\left\Vert \begin{array}{l} \text{Since } \sqrt{\Delta t} \text{ is far larger than } \Delta t \text{ for a small } \Delta t, \\ \text{and } \sigma^2 \text{ is relatively smaller than } \sigma \\ \Rightarrow \text{ we can ignore the first term } \frac{\sigma^2 \Delta t}{2} \end{array} \right.$$

$$\approx 1 \pm \sigma\sqrt{\Delta t}$$

$$\approx e^{+\sigma\sqrt{\Delta t}} \text{ (because } u > 1 \Rightarrow u \neq e^{-\sigma\sqrt{\Delta t}})$$
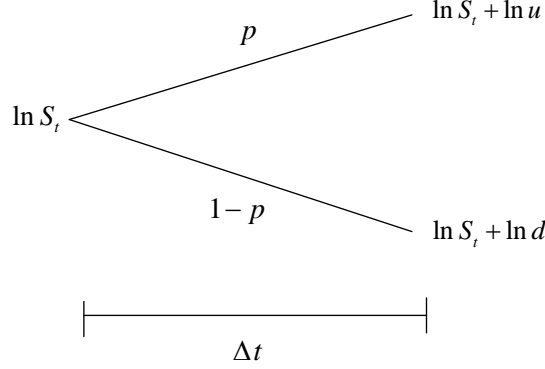
- Implementation of the CRR binomial tree model:

Figure 4-5

$S(n,0) = S_0 u^n d^0$

$S(2,0) = S_0 u^2$

$S(1,0) = S_0 u$

$\bullet \, S(i,j) = S_0 u^{i-j} d^j$

$S(n,j) = S_0 u^{n-j} d^j$

$S(2,1) = S_0$

$S(0,0) = S_0$

$S(1,1) = S_0 d$

$S(2,2) = S_0 d^2$

$S(n,n) = S_0 u^0 d^n$

1. $S(i,j) = S_0 u^{i-j} d^j$, for $0 \le j \le i \le n$ and $T/n \equiv \Delta t$
2. $c(n,j) = $ payoff for $S_0 u^{n-j} d^j$ at maturity
3. $c(i,j) = e^{-r\Delta t}[p \cdot c(i+1, j) + (1-p) \cdot c(i+1, j+1)]$ (backward induction)
4. For American options, $c(i,j) = \max(c(i,j),$ exercise value for $S_0 u^{i-j} d^j)$
5. $c(0,0)$ is the option price today
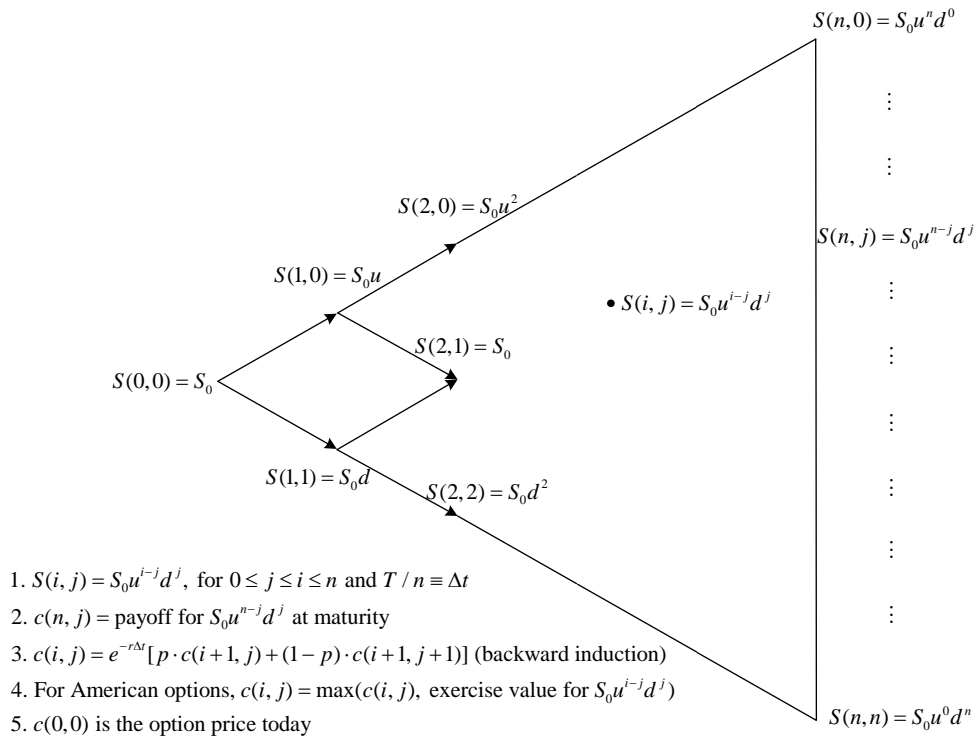
- Problems of the CRR model: 1) $p = \frac{e^{r\Delta t} - d}{u - d} = \frac{e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}}$ is not necessary in $[0, 1]$ unless Δt is small enough. 2) The approximations used to derive $\text{var}(S_{t+\Delta t})$ as well as u and d are vaild only when Δt is very small.

- Another binomial tree model: by considering the logarithmic stock price space and the constraint of $p = 1/2$, we can derive $p = \frac{1}{2}$, $u = e^{(r-\frac{\sigma^2}{2})\Delta t + \sigma\sqrt{\Delta t}}$, $d = e^{(r-\frac{\sigma^2}{2})\Delta t - \sigma\sqrt{\Delta t}}$.

Figure 4-6

$$\because \ln S_{t+\Delta t} \sim ND(\ln S_t + (r - \tfrac{\sigma^2}{2})\Delta t, \sigma^2 \Delta t)$$

$$\therefore E[\ln S_{t+\Delta t}] = \ln S_t + (r - \tfrac{\sigma^2}{2})\Delta t, \operatorname{var}(\ln S_{t+\Delta t}) = \sigma^2 \Delta t$$

Matching mean and variance of $\ln S_{t+\Delta t}$:

$$\Rightarrow \begin{cases} p(\ln S_t + \ln u) + (1-p)(\ln S_t + \ln d) = \ln S_t + (r - \tfrac{\sigma^2}{2})\Delta t & (1) \\ p(\ln S_t + \ln u)^2 + (1-p)(\ln S_t + \ln d)^2 - (\ln S_t + (r - \tfrac{\sigma^2}{2})\Delta t)^2 = \sigma^2 \Delta t & (2) \end{cases}$$

In this model, p is fixed to be 0.5, and only u and d are left as unknowns.

From Eq. (1), we can derive $0.5 \ln u + 0.5 \ln d = (r - \tfrac{\sigma^2}{2})\Delta t \Rightarrow \ln u = 2(r - \tfrac{\sigma^2}{2})\Delta t - \ln d$

$\|$ define $D = \ln d$ and $\mu = (r - \tfrac{\sigma^2}{2})\Delta t$

$\Rightarrow \ln u = 2\mu - D$

Replace $\ln u$ in Eq. (2)

$\Rightarrow 0.5(\ln S_t + 2\mu - D)^2 + 0.5(\ln S_t + D)^2 - (\ln S_t + \mu)^2 = \sigma^2 \Delta t$

$\Rightarrow (\ln S_t + 2\mu - D)^2 + (\ln S_t + D)^2 - 2(\ln S_t + \mu)^2 = 2\sigma^2 \Delta t$

$\Rightarrow D^2 - 2\mu D + \mu^2 - \sigma^2 \Delta t = 0$

$\Rightarrow D = \frac{2\mu \pm \sqrt{4\mu^2 - 4(\mu^2 - \sigma^2 \Delta t)}}{2} = \mu \pm \sigma\sqrt{\Delta t} = \ln d$

$\Rightarrow \ln u = \mu \mp \sigma\sqrt{\Delta t}$

Because $u > d \Rightarrow \begin{cases} u = e^{\mu + \sigma\sqrt{\Delta t}} = e^{(r-\frac{\sigma^2}{2})\Delta t + \sigma\sqrt{\Delta t}} \\ d = e^{\mu - \sigma\sqrt{\Delta t}} = e^{(r-\frac{\sigma^2}{2})\Delta t - \sigma\sqrt{\Delta t}} \end{cases}$

- Advantages of this alternative binomial tree model: 1) there is no approximation; 2) p maintains to be a positive number between 0 and 1; 3) the convergence rate is generally better than the CRR model for pricing plain vanilla options.

- Disadvantages of this alternative binomial tree model is due to $S_0 u d \neq S_0$: 1) Since there is no stock price layer coinciding with the specified barrier price, it is less effecient to price the family of barrier options. 2) Since there is no stock price layer equal to S_0, it is impossible to apply a quick method, known as the extended tree model, to estimate some Greek letters, like Δ, Γ, and Θ.

- Trinomial Tree $u = e^{\sigma \sqrt{3\Delta t}}, d = \frac{1}{u}, p_d = -\sqrt{\frac{\Delta t}{12\sigma^2}}(r - \frac{\sigma^2}{2}) + \frac{1}{6}, p_m = \frac{2}{3}$,

$p_u = \sqrt{\frac{\Delta t}{12\sigma^2}}(r - \frac{\sigma^2}{2}) + \frac{1}{6}$ (match (i) mean of $S_{t+\Delta t}$, (ii) variance of $S_{t+\Delta t}$, (iii) $\sum p_i = 1$, (iv) $d = \frac{1}{u}$, (v) $p_m = \frac{2}{3}$, where (iv) and (v) are arbitrarily imposed constraints.)

III. Estimation and Calibration of μ and σ

$E[\ln(S_T/S_0)] = (\mu - \frac{\sigma^2}{2})T$ $\qquad\qquad \neq \qquad \ln(E[S_T/S_0])$

$\qquad\qquad \uparrow$ because $\qquad\qquad\qquad\qquad\qquad\qquad \| $ assume

$\ln S_T - \ln S_0 \sim ND((\mu - \frac{\sigma^2}{2})T, \sigma^2 T)$ $\qquad\qquad\qquad KT$

Assume $S_1 = S_0 e^{\eta}$ $(T = 1$ year$)$ $\qquad\qquad \Rightarrow E[S_T/S_0] = e^{KT}$

(η: continuosly compounding return per annum) $\qquad \Rightarrow E[S_T] = S_0\, e^{KT}$

(S_1 is stochastic $\Rightarrow \eta$ follows a distribution) $\qquad (E[S_T]$ is a known number $\Rightarrow K$ is constant)

$\Rightarrow \eta = \ln \frac{S_1}{S_0} \sim ND(\mu - \frac{\sigma^2}{2}, \sigma^2)$ $\qquad\qquad \Rightarrow K = \mu = $ expected growth rate according to

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ the lognormal property of S_T on page 4-5

$\Rightarrow \mu - \frac{\sigma^2}{2}$ is the expected value of the continuously

$\quad$ compounding return per annum

i. Considering two trading days, t and $t+1$, $\qquad$ i. Considering two trading days, t and $t+1$,

$\quad$ we can derive $S_{t+1} = S_t \cdot e^{\eta_d}$ and $\eta_d = \ln \frac{S_{t+1}}{S_t}$ $\qquad$ then $E[S_{t+1}/S_t] = e^{K_d}$ and $K_d = \ln E[S_{t+1}/S_t]$

ii. Calculating the average of $\ln \frac{S_{t+1}}{S_t}$ for n days, $\qquad$ ii. Calculating the average of $\frac{S_{t+1}}{S_t}$ for n days and

$\quad$ the result is the estimation of $\mu_d - \frac{\sigma_d^2}{2}$. $\qquad\qquad$ taking the natural log, we can estimate $K_d = \mu_d$

$\frac{1}{n}(\ln \frac{S_1}{S_0} + \ln \frac{S_2}{S_1} + \cdots + \ln \frac{S_n}{S_{n-1}})$ $\qquad\qquad \ln(\frac{1}{n}(\frac{S_1}{S_0} + \frac{S_2}{S_1} + \cdots + \frac{S_n}{S_{n-1}}))$

$= \frac{1}{n}(\ln R_1 + \ln R_2 + \cdots + \ln R_n)$ $\qquad\qquad = \ln(\frac{1}{n}(R_1 + R_2 + \cdots + R_n))$

$= \frac{1}{n}(\ln(R_1 \cdot R_2 \cdots R_n))$

$= \ln(R_1 \cdot R_2 \cdots R_n)^{\frac{1}{n}}$

$\Rightarrow$ The geometric average of daily returns $\qquad \Rightarrow$ The arithemtic average of daily returns
$\quad$ estimate the continuously compounding $\qquad\quad$ estimates the daily expected growth rate μ_d
$\quad$ return $\mu_d - \frac{\sigma_d^2}{2}$

$\Rightarrow$ The standard deviation of the series of daily $\qquad \Rightarrow$ The standard deviation of the series of $\frac{S_{t+1}}{S_t}$
$\quad \ln \frac{S_1}{S_0}, \ln \frac{S_2}{S_1}, \ldots, \ln \frac{S_n}{S_{n-1}}$ generates the estimation $\qquad$ is NOT the estimation of σ_d
$\quad$ of σ_d

$*$ Note that the estimated results based on the daily data, i.e., μ_d and σ_d, need to be annualized to derive the corresponding annual results.

- Implied volatility (the calibration (校準) of σ)

  For any option pricing function $c(S_0, K, T, r, \sigma)$, S_0 is the stock price today, K and T can be found in the option contract, and r is the risk-free rate corresponding to the time to maturity T. As for σ, it is commonly estimated based on the historical stock prices.

  However, the market price of an option reflects the consensus of the forward-looking σ of market participants and may not equal the theoretical option value based on the historical σ. Through equalizing $c(S_0, K, T, r, \sigma)$ and the market option price, it is possible to calibrate σ from the forward-looking viewpoint.

  The value of σ^* satisfying $f(\sigma^*) \equiv c(S_0, K, T, r, \sigma^*) - \text{market option price} = 0$ is called the implied volatility. Here two root-finding algorithms are introduced to solve for the implied volatility.

- Bisection Method

  First find $[a_n, b_n]$ such that $f(a_n)f(b_n) < 0$. The iterative two steps to find $[a_{n+1}, b_{n+1}]$ are as follows.

  (i) Calcuate $x_n = a_n + \frac{b_n - a_n}{2}$

  (ii) If $f(a_n)f(x_n) < 0 \Rightarrow a_{n+1} = a_n, b_{n+1} = x_n$

  $\qquad\qquad$ else $\Rightarrow a_{n+1} = x_n, b_{n+1} = b_n$

- Newton's Method

  $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$

  $\Vert$ Based on the first-order Taylor series: $f(x) = f(x_n) + f'(x_n)(x - x_n)$
  $\Vert$ (Find the root of $f(x)$, i.e., solve $f(x) = 0$.)
  $\Vert$ $\Rightarrow -f(x_n) = f'(x_n)(x - x_n) \Rightarrow x = x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$

  $\Vert$ Quadratical convergence:
  $\Vert$ According to the second-order Taylor series, $f(x) = f(x_n) + f'(x_n)(x - x_n) + \frac{f''(\xi)}{2}(x - x_n)^2$,
  $\Vert$ where ξ is between x_n and x.
  $\Vert$ Solve $f(x) = 0$:
  $\Vert$ $\Rightarrow x - x_n + \frac{f(x_n)}{f'(x_n)} = -\frac{f''(\xi)}{2f'(x_n)}(x - x_n)^2$
  $\Vert$ $\Rightarrow x - x_{n+1} = -\frac{f''(\xi)}{2f'(x_n)}(x - x_n)^2$
  $\Vert$ (Suppose $|\frac{f''(\xi)}{2f'(x_n)}|$ is bounded by a finite number M, i.e., $|\frac{f''(\xi)}{2f'(x_n)}| \leq M < \infty$.)
  $\Vert$ $\Rightarrow |x - x_{n+1}| \leq M |x - x_n|^2$
  $\Vert$ (The error is smaller than the product of a finite number and the square of the error of
  $\Vert$ the n-th iteration.)

IV. Dividends and Option Pricing

- This section introduces how to modify the option pricing models if the dividend yield q or known cash dividends D at the pre-specified time point t are considered:

Table 4-1 Three Different Models for Dividend Payments

Suppose the time point today is 0	Black-Scholes Model	Binomial Tree Model	
	European	European	American
Model 1: Dividend yield q	$S_0 \leftarrow S_0 e^{-qT}$	$S_0 \leftarrow S_0 e^{-qT}$ or $p = \dfrac{e^{(r-q)\Delta t} - d}{u - d}$	$p = \dfrac{e^{(r-q)\Delta t} - d}{u - d}$
Model 2: Known cash dividends at t as a percentage δ of S_t	not available	Figure 4-8	Figure 4-8
Model 3: Known cash dividends D at t	$S_0 \leftarrow S_0 - De^{-rt}$	$S_0 \leftarrow S_0 - De^{-rt}$ or Figure 4-10	Figure 4-10

- Model 1: dividend yield q

  It is known that the distributions of S_T are the same under $\begin{cases} S_0 + \text{paying dividend yield } q \\ S_0 e^{-qT} + \text{no dividend payment} \end{cases}$

Figure 4-7

4-11
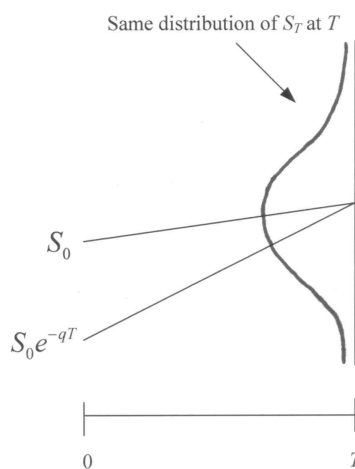
⊙ In the Black-Scholes formula, simply replace S_0 with $S_0 \cdot e^{-qT}$

$$\Rightarrow c = S_0 \cdot e^{-qT} \cdot N(d_1) - K \cdot e^{-rT} \cdot N(d_2)$$

where $d_1 = \frac{\ln(S_0/K) + (r - q + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}, d_2 = d_1 - \sigma\sqrt{T}$

As a matter of fact, the original $d\pi = -(\frac{\partial f}{\partial t} + \frac{1}{2} \cdot \frac{\partial^2 f}{\partial S^2} \cdot \sigma^2 \cdot S^2)dt$

$$\neq r \cdot \pi \cdot dt$$

new $d\pi^* = d\pi + q \cdot (\frac{\partial f}{\partial S} \cdot S \cdot dt) = r \cdot \pi \cdot dt$

(where $q \cdot (\frac{\partial f}{\partial S} \cdot S \cdot dt)$ is the dividend recived by the portfolio holder)

$$\Rightarrow \frac{\partial f}{\partial t} + \frac{\partial f}{\partial S}(r - q)S + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2 = rf$$

(where $(r - q)S$ replaces rS of the original PDE)

∗ It is worth noting that in the above PDE, when we consider the dividend yield q, it does not mean to replace r with $r - q$. In fact, for the underlying asset S, the expected growth rate is from r to become $r - q$, but the discount rate for the derivative f is still r because the right-hand side of the above PDE remains rf.

⊙ For the binomial tree model

Method 1: replace S_0 with $S_0 e^{-qT}$ + no dividend payment

Method 2: change the risk neutral probability p

$$p \cdot S_t u + (1 - p) \cdot S_t d = S_t e^{(r-q)\Delta t} \Rightarrow p = \frac{e^{(r-q)\Delta t} - d}{u - d}$$

∗ However, during the backward induction phase, we still use r to discount the expected option value at the next time point, i.e., $f = e^{-r\Delta t}[p \cdot f_u + (1 - p) \cdot f_d]$ (For European options, both above methods generate correct results, but for American options, only the second method can generate correct results.)

- Model 2: known cash dividends as a percentage δ of the stock price at the time point t
  (In practice, it is rare for companies to distribute cash dividends in this way.)

  ⊙ For the Black-Scholes formula, it is unavailable to deal with this problem.

  ⊙ For the binomial tree model, it is simple to deal with this problem (see Figure 4-8).
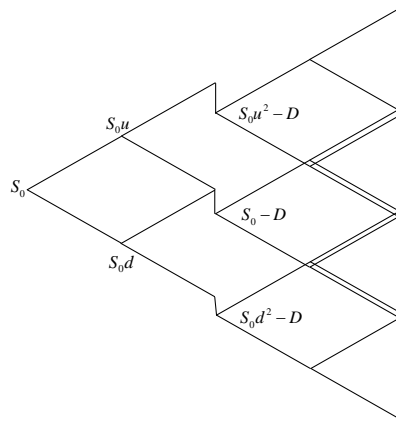
  Figure 4-8
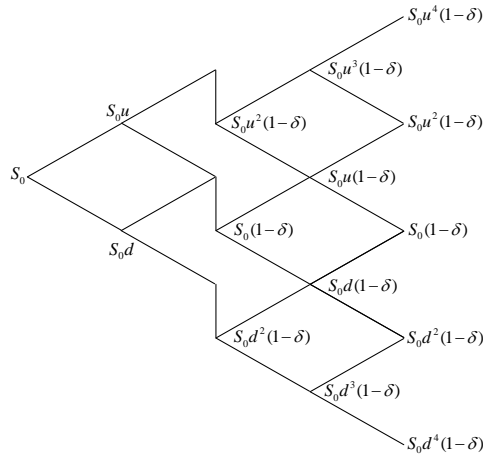
- Model 3: known cash dividends D at time point t

  ⊙ For the Black-Scholes formula, replace S_0 with $S_0 - De^{-rt}$.

  ⊙ For the binomial tree model,

  Method 1: replace S_0 with $S_0 - De^{-rt}$. However, this method only works for European options, but it cannot be applied to pricing American options.

  (This is because the distributions of S_τ ($\tau \geq t$) are the same under $\begin{cases} S_0 + \text{paying } D \text{ dollars at } t \\ S_0 - De^{-rt} + \text{ no dividend payment} \end{cases}$
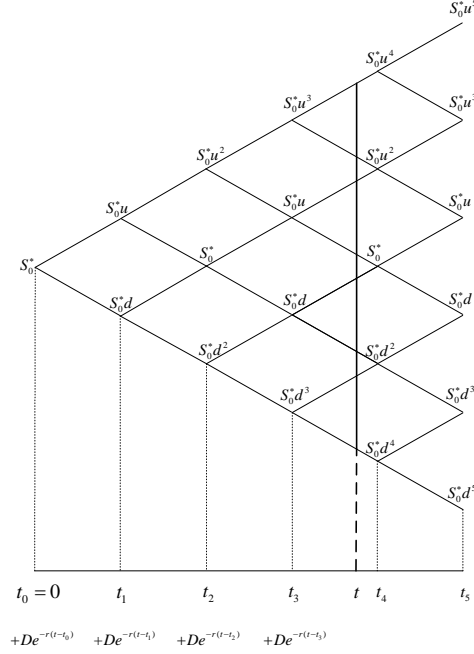
  Method 2: deduct the known dividend D from all stock prices on the divided payment date. This method cannot work because it makes the tree to be non-recombined.

  Figure 4-9

4-13

Method 3: this method can maintain the recombined feature of the binomial tree

Figure 4-10

Define $S_0^* = S_0 - De^{-r(t-t_0)} = S_0 - De^{-rt}$

Step 1: build S^*-tree following the traditional way, but it should be noted that in theory $\sigma^* = \sigma \cdot S_0/(S_0 - De^{-r(t-t_0)})$, where σ^* is the volatility for the stochastic part of the stock price and σ is the total volatility of the stock price.

(For a different volatility σ^*, we should derive new u^*, d^*, and p^* in theory. In practice, however, it is rare to calculate σ^*. The first reason is that the difference between σ^* and σ is minor. The second reason is that since σ is an estimated value based on the historical data, performing the above adjustment will not necessarily lead to a better estimation for σ^* in the forward-looking sense.)

Step 2: for $t_i < t$, replace $S_{t_i}^*$ with $S_{t_i}^* + De^{-r(t-t_i)}$

(Since $S_0^* = S_0 - De^{-r(t-t_0)} = S_0 - De^{-rt}$ in Step 1, S^*-tree is the same as the tree in Method 1 after t, which formulates the stock price process after the dividend payment appropriately. However, the stock prices before t are not consistent with the stock price process before the dividend payment, i.e., $S_0^* \neq S_0$ and $E[S_{t_i}^*] \neq E[S_{t_i}]$ for $t_i < t$. Therefore, the adjustment in Step 2 should be performed.)

4-14

- Currency option
  - ⊙ Replace the dividend yield q with the foreign risk-free rate r_f

    (Because the underlying asset S_t is a dollar of the foreign currency in domestic dollars, and holding the foreign currency can earn the foreign risk-free rate, a foreign currency is analogous to a stock paying a known dividend yield.)

  - ⊙ $c = S_0 e^{-r_f T} N(d_1) - K e^{-rT} N(d_2)$

    where $d_1 = \frac{\ln(S_0/K) + (r - r_f + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$, $d_2 = d_1 - \sigma\sqrt{T}$

- Futures options
  - ⊙ Call holders: have the right to enter a long position futures with the deliver price to be F_T and receive cash $F_T - K$ if $F_T \geq K$, where F_T is the latest settlement futures price before T (usually F_T is the closing price on the date T) and K is the strike price in futures options.

  - ⊙ Put holders: have the right to enter a short position of futures with the deliver price to be F_T and receive cash $K - F_T$ if $F_T \leq K$.

∗ Since a futures contract is worth zero when it is first created (in the above cases, it is at T), it can be concluded that the payoff of futures option is similar to that of exercising plain vanilla options, and the only difference is to replace S_T with F_T. This observation implies that we can develop the pricing formula for futures options by following the same way to develop the option formula for plain vanilla options.

⊙ Black-Scholes formula for futures options:

Suppose the futures price follows the geometric Brownian motion: $\frac{dF}{F} = \mu dt + \sigma dZ$

$c(F, t)$ is the call on futures, and according to the Itô's lemma

$\Rightarrow dc = (\frac{\partial c}{\partial t} + \frac{\partial c}{\partial F}\mu F + \frac{1}{2}\frac{\partial^2 c}{\partial F^2}\sigma^2 F^2)dt + \frac{\partial c}{\partial F}\sigma F dZ$

Construct portfolio $\pi = -c + (\frac{\partial c}{\partial F})F = -c$ (the initial value of futures is 0)

$d\pi = -dc + (\frac{\partial c}{\partial F})dF = -(\frac{\partial c}{\partial t} + \frac{1}{2}\frac{\partial^2 c}{\partial F^2}\sigma^2 F^2)dt$

$\qquad\qquad = r\pi dt = r(-c)dt$

$\Rightarrow \frac{\partial c}{\partial t} + \frac{1}{2}\frac{\partial^2 c}{\partial F^2}\sigma^2 F^2 = rc$

(Comparing with the PDE for stock options, these is no such term $\frac{\partial c}{\partial F}(r-q)F$.)

Thus, one can set $q = r$ and $S_0 = F_0$ in the Black-Scholes formula to derive the formula for futures options.)

$c = e^{-rT}[F_0 N(d_1) - KN(d_2)]$, where $d_1 = \frac{\ln(F_0/K) + \frac{\sigma^2 \cdot T}{2}}{\sigma\sqrt{T}}$ and $d_2 = d_1 - \sigma\sqrt{T}$

⊙ Binomial Tree

Because $q = r$, $p = \frac{e^{(r-q)\Delta t} - d}{u - d} = \frac{1-d}{u-d}$ (but we still use r as the discount rate)

Another way to derive p
Figure 4-11
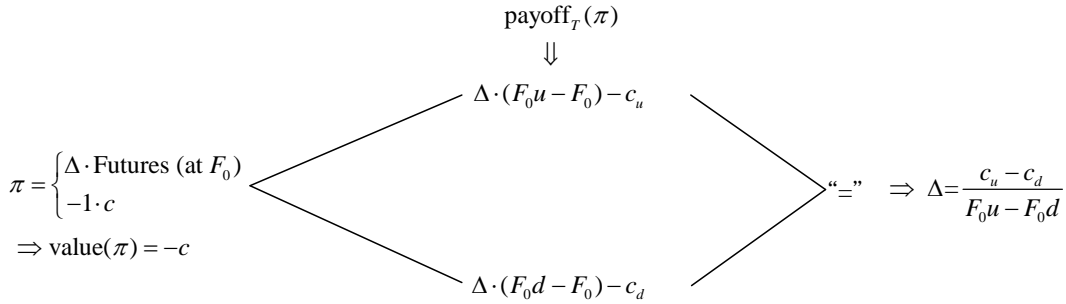
$\Rightarrow -c = [\Delta(F_0 u - F_0) - c_u]e^{-rT}$ (the initial value of the futures is 0)
$\Rightarrow c = e^{-rT}(p \cdot c_u + (1-p) \cdot c_d)$, where $p = \frac{1-d}{u-d}$

- The Black-Scholes model as well as the binomial tree model are versatile models:

Treat stock index, currency, and futures like a share of stock paying a dividend yield q

For stock index options: $q =$ average dividend yield on the index over the option life

For currency options: $q = r_f$

For futures options: $q = r$

V. Introduction of Combinatorial Method

- Combinatorics (組合數學) is a branch of pure mathematics concerning the study of discrete (and usually finite) objects. Aspects of combinatorics include "counting" the objects satisfying certain criteria, deciding when the criteria can be met, or finding "largest", "smallest", or "optimal" objects.

- Combinatorial method for European options

  Based on the binomial tree framework, applying the combinatorial method is far faster than the backward induction procedure. In fact, it is not necessary to build the binomial tree in the combinatorial method, which is another advantage of the combinatorial method because is can save memory space for computer programming.

Figure 4-12

T is partitioned
into n time steps
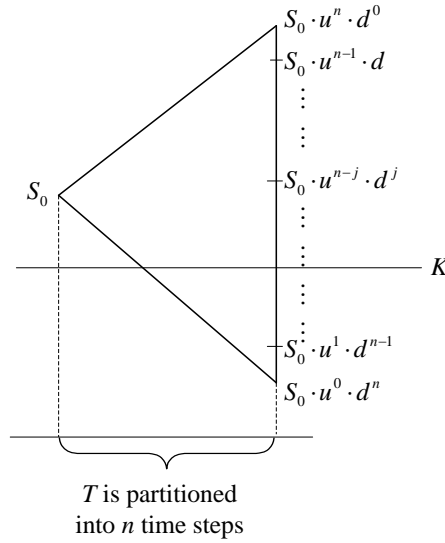
European option value $= e^{-rT} \sum\limits_{j=0}^{n} \binom{n}{j} p^{n-j}(1-p)^j \max(S_0 u^{n-j} d^j - K, 0)$, where $\binom{n}{j}$, also denoted as C_j^n, is the combination of j from n, $u = e^{\sigma\sqrt{\Delta t}}$, $d = e^{-\sigma\sqrt{\Delta t}}$, and $p = \frac{e^{(r-q)\Delta t} - d}{u-d}$.

∗ For the binomial tree model, its complexity is $O(n^2)$, whereas for the above combinatorial method, the complexity is $O(n)$. The difference of required computational time is substantial for a large number of n, e.g., $n > 5000$.

4-17

Appendix A. Binominal Tree Model for Jump-Diffusion Processes

• The content in this appendix belongs to the advanced content.

• Binomial tree model for the jump-diffusion process in Amin (1993):
  ⊙ Consider Merton's (1976) lognormal jump-diffusion process as follows.
$$\frac{dS}{S} = (r - q - \lambda K_Y)dt + \sigma dZ + (Y - 1)dq,$$
  where $K_Y = E[Y - 1]$ and $\ln Y \sim ND(\mu_J, \sigma_J^2)$. By the Itô's lemma,
$$d\ln S = (r - q - \tfrac{1}{2}\sigma^2 - \lambda K_Y)dt + \sigma dZ + \ln Y dq = \alpha dt + \sigma dZ + \ln Y dq$$

  ⊙ The lattice model proposed by Amin (1993)

Figure 4-13

Figure 4-13

$\ln S_0$ •

$\bullet$ $\ln S_{\Delta t} = \ln S_0 + \alpha \Delta t + 2\sigma\sqrt{\Delta t}$

$\bullet$ $\ln S_{\Delta t} = \ln S_0 + \alpha \Delta t + \sigma\sqrt{\Delta t}$

$\bullet$ $\ln S_{\Delta t} = \ln S_0 + \alpha \Delta t$

$\bullet$ $\ln S_{\Delta t} = \ln S_0 + \alpha \Delta t - \sigma\sqrt{\Delta t}$

$\bullet$ $\ln S_{\Delta t} = \ln S_0 + \alpha \Delta t - 2\sigma\sqrt{\Delta t}$

$\cdots$

$\bullet$ $\ln S_{i\Delta t} = \ln S_0 + \alpha i \Delta t + 2\sigma\sqrt{\Delta t}$

$\bullet$ $\ln S_{i\Delta t} = \ln S_0 + \alpha i \Delta t + \sigma\sqrt{\Delta t}$

$\bullet$ $\ln S_{i\Delta t} = \ln S_0 + \alpha i \Delta t$

$\bullet$ $\ln S_{i\Delta t} = \ln S_0 + \alpha i \Delta t - \sigma\sqrt{\Delta t}$

$\bullet$ $\ln S_{i\Delta t} = \ln S_0 + \alpha i \Delta t - 2\sigma\sqrt{\Delta t}$

$\cdots$

$0 \quad \Delta t \qquad \cdots \qquad i\Delta t \qquad \cdots \qquad n\Delta t = T$
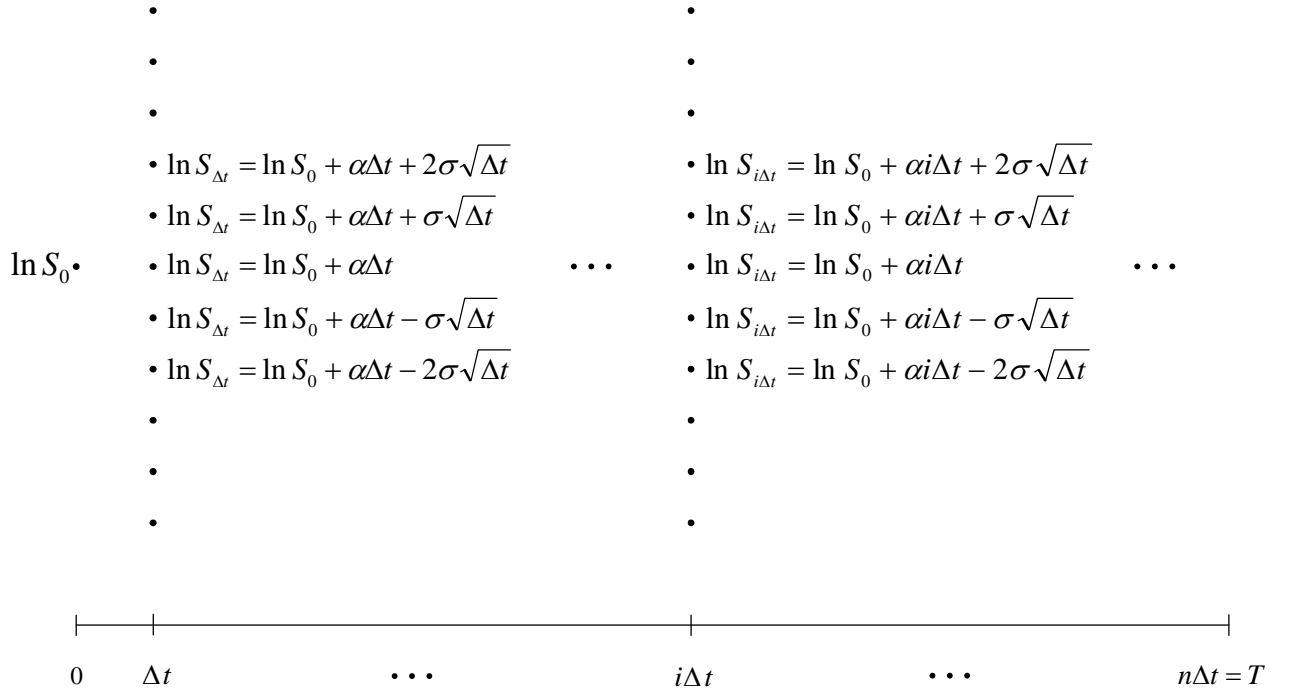
⊙ Backward induction:

Figure 4-14

$\ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t + M_J\sigma\sqrt{\Delta t}$

$\vdots$

$\ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t + 2\sigma\sqrt{\Delta t}$

$\ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t + \sigma\sqrt{\Delta t}$

$\ln S_t \qquad \ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t$

$\ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t - \sigma\sqrt{\Delta t}$

$\ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t - 2\sigma\sqrt{\Delta t}$

$\vdots$

$\ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t - M_J\sigma\sqrt{\Delta t}$
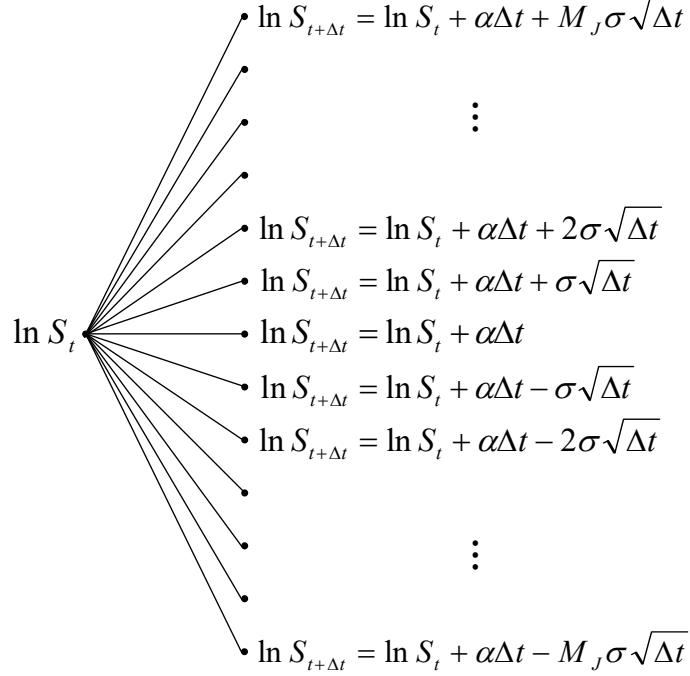
∗ Define the branching probabilities as follows.

$$p(k) = \begin{cases} \text{prob}(\ln S_{t+\Delta t} - \ln S_t = \alpha\Delta t + \sigma\sqrt{\Delta t}) & = (1-\lambda\Delta t)p_u + \lambda\Delta t\Delta N(k) \\ \text{prob}(\ln S_{t+\Delta t} - \ln S_t = \alpha\Delta t - \sigma\sqrt{\Delta t}) & = (1-\lambda\Delta t)p_d + \lambda\Delta t\Delta N(k) \\ \text{prob}(\ln S_{t+\Delta t} - \ln S_t = \alpha\Delta t + k\sigma\sqrt{\Delta t}) = \lambda\Delta t N(k), \ \text{for} \ -M_J \le k \le M_J \end{cases},$$

where $\Delta N(k) = N((\alpha\Delta t + (k+\frac{1}{2})\sigma\sqrt{\Delta t} - \mu_J)/\sigma_J) - N((\alpha\Delta t + (k-\frac{1}{2})\sigma\sqrt{\Delta t} - \mu_J)/\sigma_J)$ if $k \ne \pm M_J$, $M_J \equiv \text{int}(\frac{3\sigma_J}{\sigma\sqrt{\Delta t}})$, $\Delta N(M_J) = 1 - N((\alpha\Delta t + (M_J - \frac{1}{2})\sigma\sqrt{\Delta t} - \mu_J)/\sigma_J)$, and $\Delta N(-M_J) = N((\alpha\Delta t - (M_J - \frac{1}{2})\sigma\sqrt{\Delta t} - \mu_J)\sigma_J)$.

(Note that using $\lambda\Delta t\Delta N(k)$ in the above way can fully capture the effect of the jump component $\ln Y\,dq$ in the $d\ln S$ process.)

* Determine p_u and p_d by matching the mean and variance of the remaining diffusion process, $d \ln S = \alpha dt + \sigma dZ$

$$\begin{cases} p_u S_t u + p_d S_t d = S_t e^{\alpha \Delta t} \\ p_u S_t^2 u^2 + p_d S_t^2 d^2 - (p_u S_t u + p_d S_t d)^2 = S_t^2 \sigma^2 \Delta t, \\ p_u + p_d = 1 \end{cases}$$

where $u = e^{\alpha \Delta t + \sigma \sqrt{\Delta t}}$ and $d = e^{\alpha \Delta t - \sigma \sqrt{\Delta t}}$. Analogous with the CRR binomial tree model, we can infer

$$p_u = \frac{e^{\alpha \Delta t} - d}{u - d}.$$

* However if we consider the probabilities of $(1 - \lambda \Delta t)p_u$ and $(1 - \lambda \Delta t)p_d$ for the upward and downward branches, respectively, when the jump component is introduced. The mean and variance generated by the probabilities of $(1 - \lambda \Delta t)p_u$ and $(1 - \lambda \Delta t)p_d$ are no more $S_t e^{\alpha \Delta t}$ and $S_t^2 \sigma^2 \Delta t$.

$$\begin{cases} (1 - \lambda \Delta t)p_u S_t u + (1 - \lambda \Delta t)p_d S_t d = (1 - \lambda \Delta t)S_t e^{\alpha \Delta t} \\ (1 - \lambda \Delta t)p_u S_t^2 u^2 + (1 - \lambda \Delta t)p_d S_t^2 d^2 - [(1 - \lambda \Delta t)S_t e^{\alpha \Delta t}]^2 \approx (1 - \lambda \Delta t)S_t^2 \sigma^2 \Delta t, \\ (1 - \lambda \Delta t)p_u + (1 - \lambda \Delta t)p_d = (1 - \lambda \Delta t) \end{cases}$$

For the variance equation, it can be rewritten as $(1 - \lambda \Delta t)[p_u S_t^2 u^2 + p_d S_t^2 d^2 - (1 - \lambda \Delta t)(S_t e^{\alpha \Delta t})^2] \approx (1 - \lambda \Delta t) \cdot S_t^2 \sigma^2 \Delta t$, if Δt is small such that $(1 - \lambda \Delta t) \approx 1$.
Thus, the standard deviation become $\sqrt{(1 - \lambda \Delta t)}S_t \sigma \sqrt{\Delta t}$. By comparing with the mean, $(1 - \lambda \Delta t)S_t e^{\alpha \Delta t}$, one can find that the deviation of the mean from its true value is more serious than the deviation of the standard deviation from its true value because $\sqrt{(1 - \lambda \Delta t)}$ is closer to 1 than $(1 - \lambda \Delta t)$ when Δt is small.

* To offset the above undesirable effects, we should solve p_u and p_d in the following system of equations.

$$\begin{cases} p_u S_t u + p_d S_t d = \frac{S_t e^{\alpha \Delta t}}{1 - \lambda \Delta t} \\ p_u S_t^2 u^2 + p_d S_t^2 d^2 - (p_u S_t u + p_d S_t d)^2 = \frac{S_t^2 \sigma^2 \Delta t}{1 - \lambda \Delta t} \\ p_u + p_d = 1 \end{cases}$$

* In Amin (1993), only the adjustment for the mean is considered, and thus the probability p_u can be derived as

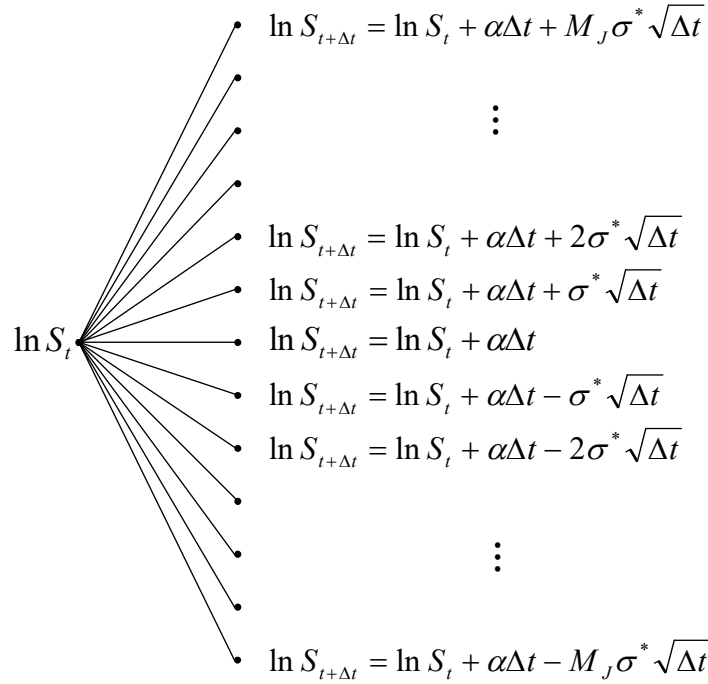$$p_u = \frac{\frac{e^{\alpha \Delta t}}{1 - \lambda \Delta t} - d}{u - d} .$$

∗ Option value of a node can be expressed as

Option value of $\ln S_t$

$$= \sum_{-M_J \le k \le M_J} p(k) \cdot (\text{option value for } \ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t + k\sigma\sqrt{\Delta t})$$

⊙ If the adjustment for the variance is also considered, one can adjust the grid size of the multinomial tree to achieve this goal. By defining $\sigma^* = \frac{\sigma}{\sqrt{1-\lambda\Delta t}}$, $u^* = e^{\alpha\Delta t + \sigma^*\sqrt{\Delta t}}$, $d^* = e^{\alpha\Delta t - \sigma^*\sqrt{\Delta t}}$, one can derive the following tree structure and the corresponding branching probabilities.

Figure 4-15

$\ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t + M_J\sigma^*\sqrt{\Delta t}$

$\vdots$

$\ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t + 2\sigma^*\sqrt{\Delta t}$

$\ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t + \sigma^*\sqrt{\Delta t}$

$\ln S_t$ → $\ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t$

$\ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t - \sigma^*\sqrt{\Delta t}$

$\ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t - 2\sigma^*\sqrt{\Delta t}$

$\vdots$

$\ln S_{t+\Delta t} = \ln S_t + \alpha\Delta t - M_J\sigma^*\sqrt{\Delta t}$

$$p^*(k) = \begin{cases} \text{prob}(\ln S_{t+\Delta t} - \ln S_t = \alpha\Delta t + \sigma^*\sqrt{\Delta t}) & = (1-\lambda\Delta t)p_u^* + \lambda\Delta t\Delta N(k) \\ \text{prob}(\ln S_{t+\Delta t} - \ln S_t = \alpha\Delta t - \sigma^*\sqrt{\Delta t}) & = (1-\lambda\Delta t)p_d^* + \lambda\Delta t\Delta N(k) \\ \text{prob}(\ln S_{t+\Delta t} - \ln S_t = \alpha\Delta t + k\sigma^*\sqrt{\Delta t}) = \lambda\Delta t N(k), \text{ for } -M_J \le k \le M_J \end{cases},$$

where $\Delta N(k) = N((\alpha\Delta t + (k+\frac{1}{2})\sigma^*\sqrt{\Delta t} - \mu_J)/\sigma_J) - N((\alpha\Delta t + (k-\frac{1}{2})\sigma^*\sqrt{\Delta t} - \mu_J)/\sigma_J)$ if $k \ne \pm M_J$, $M_J \equiv \text{int}(\frac{3\sigma_J}{\sigma^*\sqrt{\Delta t}})$, $\Delta N(M_J) = 1 - N((\alpha\Delta t + (M_J - \frac{1}{2})\sigma^*\sqrt{\Delta t} - \mu_J)/\sigma_J)$, $\Delta N(-M_J) = N((\alpha\Delta t - (M_J - \frac{1}{2})\sigma^*\sqrt{\Delta t} - \mu_J)/\sigma_J)$, and $p_u^* = \frac{\frac{e^{\alpha\Delta t}}{\sqrt{1-\lambda\Delta t}} - d^*}{u^* - d^*}$ and $p_d^* = 1 - p_u^*$.