

Ch 12. Interest Rate and Credit Models

I. Equilibrium Interest Rate Models

II. No-Arbitrage Interest Rate Models

III. Forward Rate Models

IV. Credit Risk Models

- This chapter introduces interest rate models first. Two categories of the interest rate models, the equilibrium and no-arbitrage models, will be introduced. In addition, the forward rate models is also discussed, in which the risk factor is the instantaneous forward rate rather than the instantaneous short rate in the interest rate models. Finally, two classical credit risk models, the reduced-form and structural models, are introduced.

I. Equilibrium Interest Rate Models

- After the emergence of the Black-Scholes model, some academics try to model the bond price following the lognormal or normal distribution and thus derive the pricing formulae of the bond options following the method in the Black-Scholes model. However, the critical problem of these model is that the bond price is neither lognormally nor normally distributed.
- Furthermore, for all bonds and bonds options, their prices are affected by the change of the term structure of interest rates, which implies that they share a common source of risk factor. So it is inappropriate to model the movement of the price of each bond separately, i.e., the price of each bond is assumed to follow an individual stochastic process.
- It is known that there exist a relationship between the bond price and the interest rate theoretically. Hence, some academics turn to consider the stochastic process of interest rates. They try to model the short rate process dr , which is by definition the instantaneous interest rate for an infinitesimal period of time.
- It is worth noting that any interest rate we can observe in the market is associated with a time to maturity and thus the interest rates we can observe in the market is called the long rates. Unlike the long rate, the short rate is unobservable.

- To model the short rate, there are two streams of models: the equilibrium and no-arbitrage models. In this section, two famous equilibrium interest rate models are introduced: the Vasicek and Cox-Ingersoll-Ross (CIR) models.

- Vasicek model

⊙ $dr = \beta(\mu - r)dt + \sigma dZ$, where μ is the long-term mean of r , and β measures the speed of mean reversion.

$$\left\| \begin{array}{l} \text{By performing the stochastic integral, we can express } r(T) \text{ as follows.} \\ r(T) = e^{-\beta(T-t)}(r(t) + \mu(e^{\beta(T-t)} - 1) + \sigma \int_t^T e^{\beta(\tau-t)} dZ(\tau)), \\ \text{where } r(t) \text{ is the level of short rate today.} \\ \\ \text{From the above equation, we know that } r(T) \text{ is normally distributed, and} \\ E_t[r(T)] = \mu + (r(t) - \mu)e^{-\beta(T-t)} \\ \text{var}_t(r(T)) = \frac{\sigma^2}{2\beta}(1 - e^{-2\beta(T-t)}) \end{array} \right.$$

Denote P to be the current price of a zero coupon bond with 1 dollar payoff at maturity, and P should be the function of r , t (the time point today), T (the maturity date). According to the Itô's Lemma

$$\begin{aligned} dP(t, r, T) &= \left(\frac{\partial P}{\partial t} + \beta(\mu - r) \frac{\partial P}{\partial r} + \frac{1}{2} \sigma^2 \frac{\partial^2 P}{\partial r^2} \right) dt + \sigma \frac{\partial P}{\partial r} dZ \\ &\equiv P \cdot \mu(t, T) dt + P \cdot \sigma(t, T) dZ \end{aligned}$$

(Note that the maturity date T is a constant in the bond contract, so it is not necessary to consider the differentiation of T in the Itô's Lemma.)

Construct a portfolio W by investing $\left\{ \begin{array}{l} \frac{\sigma(t, T_2)}{\sigma(t, T_2) - \sigma(t, T_1)} \text{ shares of } P(t, r, T_1) \\ \frac{\sigma(t, T_1)}{\sigma(t, T_2) - \sigma(t, T_1)} \text{ shares of } P(t, r, T_2) \end{array} \right.$

$$\Rightarrow dW = \frac{\mu(t, T_1)\sigma(t, T_2) - \mu(t, T_2)\sigma(t, T_1)}{\sigma(t, T_2) - \sigma(t, T_1)} dt = r(t) dt$$

$$\Rightarrow \frac{\mu(t, T_1)\sigma(t, T_2) - \mu(t, T_2)\sigma(t, T_1)}{\sigma(t, T_2) - \sigma(t, T_1)} = r(t)$$

$$\Rightarrow \mu(t, T_1)\sigma(t, T_2) - r(t)\sigma(t, T_2) = \mu(t, T_2)\sigma(t, T_1) - r(t)\sigma(t, T_1)$$

$$\Rightarrow \frac{\mu(t, T_1) - r(t)}{\sigma(t, T_1)} = \frac{\mu(t, T_2) - r(t)}{\sigma(t, T_2)} = \lambda$$

(For zero-coupon bonds with different maturity dates, the ratio of their excess returns over volatilities are identical. It is reasonable since all these bonds share a common risk factor dr and they should be with the same market price of risk, which is, by definition, the ratio of the excess return over volatility of any derivatives based on r .)

$$\Rightarrow \mu(t, T) - r(t) = \lambda \cdot \sigma(t, T)$$

Substitute $\mu(t, T)$ and $\sigma(t, T)$ of $dP(t, r, T)$ to derive

$$\Rightarrow \frac{(\frac{\partial P}{\partial t} + \beta(\mu - r)\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial^2 P}{\partial r^2})}{P} - r(t) = \frac{\lambda \cdot \sigma \frac{\partial P}{\partial r}}{P}$$

$$\Rightarrow \frac{\partial P}{\partial t} + (\beta(\mu - r) - \lambda\sigma)\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial^2 P}{\partial r^2} = rP$$

(Based on the Vasicek short rate model, the price of any zero coupon bond should satisfy the above partial differential equation. To solve the above partial differential equation, we can derive the theoretical value of any zero coupon bond as follows.)

$$P(t, r, T) = A(t, T)e^{-B(t, T)r(t)},$$

where $r(t)$ is the current short rate level, and

$$A(t, T) = \begin{cases} \exp[\frac{(B(t, T) - T + t)(\beta^2\mu - \frac{\sigma^2}{2})}{\beta^2} + \frac{\sigma \cdot \lambda}{\beta}(B(t, T) - (T - t)) - \frac{\sigma^2 B(t, T)^2}{4\beta}] & \text{if } \beta \neq 0 \\ \exp[\frac{\sigma^2(T-t)^3}{6}] & \text{if } \beta = 0 \end{cases}$$

$$B(t, T) = \begin{cases} \frac{1 - e^{-\beta(T-t)}}{\beta} & \text{if } \beta \neq 0 \\ T - t & \text{if } \beta = 0 \end{cases}$$

* Note that in the risk-neutral world, the market price of risk, λ , is zero, and the above pricing formula becomes identical to that presented in Hull (2011).

* The formula of the zero coupon bond price $P(t, r, T)$ makes it possible to derive the whole term structure based on the Vasicek short rate process. This is because given the zero coupon bond price $P(t, r, T)$ and the identity function $P(t, r, T) = \exp(-r(t, T)(T - t))$, the zero rate corresponding to the maturity date T , $r(t, T)$, can be derived through $r(t, T) = -\ln(P(t, r, T))/(T - t)$.

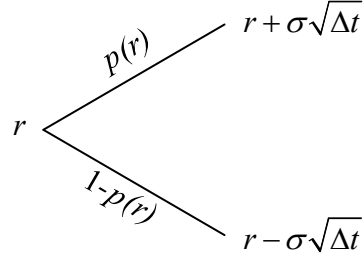
* In a word, once you specify the values of the parameters β , μ , and σ and the initial value of the short rate $r(t)$, you can derive a corresponding term structure. The Vasicek model can generate normal, inverted, and humped-shape term structures.

* One possible way to calibrate the estimations of β , μ , σ , and $r(t)$ is to employ the least squares approach such that the generated term structure can best fit the real term structure in the market.

⊙ The binomial tree to simulate the Vasicek interest rate process:

Since the volatility term is constant in the Vasicek model, it is straightforward to apply the method in Cox and Rubinstein (1985) to constructing the corresponding binomial tree.

Figure 12-1



$p(r) = \frac{1}{2} + \frac{\beta(\mu-r)\sqrt{\Delta t}}{2\sigma}$. If $p(r)$ is not in $[0, 1]$, consider a smaller Δt to fix this problem. When Δt approaches zero, $p(r)$ will approach 0.5.

- Cox-Ingersoll-Ross (CIR) model

$$dr = \beta(\mu - r)dt + \sigma\sqrt{r}dZ$$

* In comparison with the Vasicek model, the only difference is the volatility term, which is from the constant σ to $\sigma\sqrt{r}$. The modification decreases the volatility of the short rate when the interest rate is low and thus the short rate is always positive. Hence, the CIR model remedies the negative interest rate problem in the Vasicek model.

* It is worth noting that Cox, Ingersoll, and Ross (1985) do not develop their model by simply modifying the process of the Vasicek model. Instead, they develop an equilibrium model in which the short rate is assumed to be the sum of squared values of n Ornstein-Uhlenbeck processes as follows.

$$r(t) = \sum_{i=1}^n X_i^2(t), \text{ where } dX_i(t) = -\frac{1}{2}bX_i(t)dt + \frac{1}{2}\sigma dZ_i(t).$$

Under the above assumption, the short rate r follows a noncentral Chi-square distribution and thus is always positive.

Following the similar procedure in the Vasicek model, the partial differential equation of the zero coupon bond based on the CIR model is as follows:

$$\frac{\partial P}{\partial t} + \beta(\mu - r)\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2 r \frac{\partial^2 P}{\partial r^2} = rP.$$

Solve the above partial differential equation, the theoretical value of the zero-coupon bond today is

$$P(t, r, T) = A(t, T)e^{-B(t, T)r(t)},$$

where $r(t)$ is the short rate level today, and

$$A(t, T) = \left[\frac{2 \cdot v \cdot e^{(\beta+v)(T-t)/2}}{(\beta+v)[e^{v(T-t)} - 1] + 2v} \right] 2\beta\mu/\sigma^2,$$

$$B(t, T) = \frac{2[e^{v(T-t)} - 1]}{(\beta+v)[e^{v(T-t)} - 1] + 2v},$$

$$v = \sqrt{\beta^2 + 2\sigma^2}.$$

⊙ The binomial tree approach to simulate the CIR interest rate process:

Because the CIR model is with a non-constant volatility, we need to apply the method in Nelson and Ramaswamy (1990) to construct the corresponding binomial tree.

Consider $X = \frac{2\sqrt{r}}{\sigma} \Rightarrow r = f(X) \equiv \frac{X^2\sigma^2}{4}$.

Therefore, $dX = m(x)dt + dZ$, where $m(x) = \frac{2\beta\mu}{\sigma^2 X} - \frac{\beta X}{2} - \frac{1}{2X}$.

Finally, we construct the recombined X -tree first (see Figure 12-2) and then derive the r -tree by transforming the value of X into the corresponding value of r via $f(X)$ (see Figure 12-3).

Figure 12-2

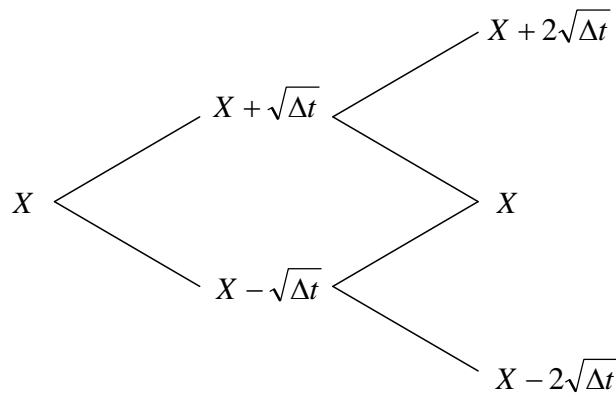
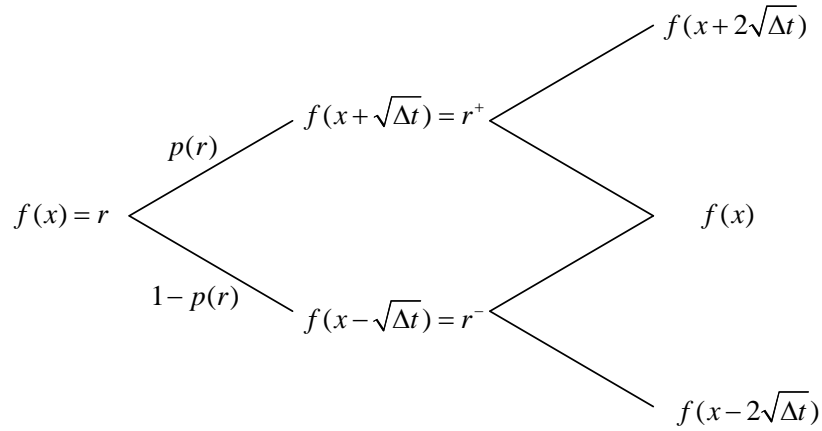


Figure 12-3



The probability for each node is $p(r) = \frac{\beta(\mu - r)\Delta t + r - r^-}{r^+ - r^-}$.

(Note that the outgoing probabilities for each node are not the same.)

- Multifactor CIR Model:

$$r = r_1 + r_2,$$

where

$$dr_1 = \beta_1(\mu_1 - r_1)dt + \sigma\sqrt{r_1}dZ_1,$$

$$dr_2 = \beta_2(\mu_2 - r_2)dt + \sigma\sqrt{r_2}dZ_2,$$

and $\text{corr}(dZ_1, dZ_2) = \rho$

* Since both r_1 and r_2 are short rates, the characteristics of this model are similar to those of the single-factor CIR model.

Equilibrium models vs. No-arbitrage models

	Equilibrium models	No-arbitrage models
Match today's term structure	no	yes
Today's term structure is for	output	input
Drift term is	a function of r (usually for mean reversion)	a function of both r and t (for matching today's term structure)
Volatility term is	a constant or a function of r	a constant or a function of t (for matching today's volatility term structure)
Disadvantages	<ul style="list-style-type: none"> (i) Cannot match today's term structure and thus the bond prices in the markets. (ii) For interest rate or bond options, the pricing error is unacceptable. (A 1% error in the price of the underlying asset may generally lead to a 25% error in the option price.) 	<ul style="list-style-type: none"> (i) Since the drift and volatility terms are specified to be functions of t in order to match the current interest rate and volatility term structures, its exactness of matching term structures disappears with the passage of time. (ii) Thus, this type of models is not suited for prediction.
Advantages	<ul style="list-style-type: none"> (i) It provides superior prediction for future interest rate movements. (ii) Thus, this type of models is more suited for the risk management of financial institutions. 	<ul style="list-style-type: none"> (i) Match the term structure today exactly and thus value interest-rate-relative products more accurately. (ii) This type of models is useful for traders, who need accurate prices for both spots and derivatives of interest rates or bonds.

II. No-Arbitrage Interest Rates Models

- This section introduces three no-arbitrage interest rate models. The first one is proposed by Black, Derman, and Toy (1990), in which the short rate follows the lognormal distribution. The second and third models are respectively Ho and Lee (1986) and Hull and White (1994), in which the short rate follows the normal distribution.

- Black-Derman-Toy (BDT) model: Similar to the stock price process under the Black-Scholes framework, the short rate is assumed to follow the lognormal distribution.

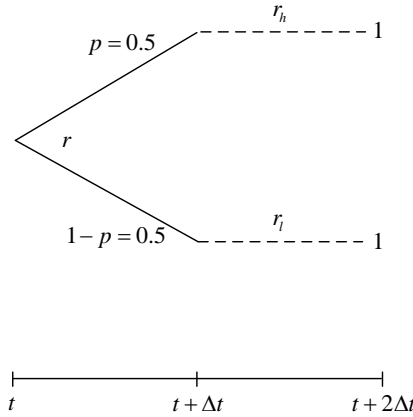
$$d \ln r = \theta(r, \sigma, t)dt + \sigma(t)dZ$$

* $\theta(r, \sigma, t) = \theta(t) + \frac{\sigma'(t)}{\sigma(t)} \ln r$ (see Hull and White (1990)), where $\theta(t)$ is introduced to match the term structure of interest rates today, and if $\sigma'(t) < 0$, the interest rate exhibits the mean reverting to $\theta(t)$.

* The advantage of this assumption is to ensure that the interest rate is nonnegative. However, lognormal models for the interest rates usually do not give analytical solutions for even basic fix-income securities.

⊙ As a consequence, a binomial tree method is employed to simulate the BDT interest rate process.

Figure 12-4



* Note that r is applied to the period $(t, t + \Delta t]$, and for the period $(t + \Delta t, t + 2\Delta t]$, there is 50% of probability that the interest rate is r_h and 50% of probability that the interest rate is r_l .

* Similar to the CRR binomial tree model,

$$r_h = re^{\sigma_t\sqrt{\Delta t}} \text{ and } r_l = re^{-\sigma_t\sqrt{\Delta t}}.$$

Then we can express the spacing parameter v to be the ratio of r_h over r_l .

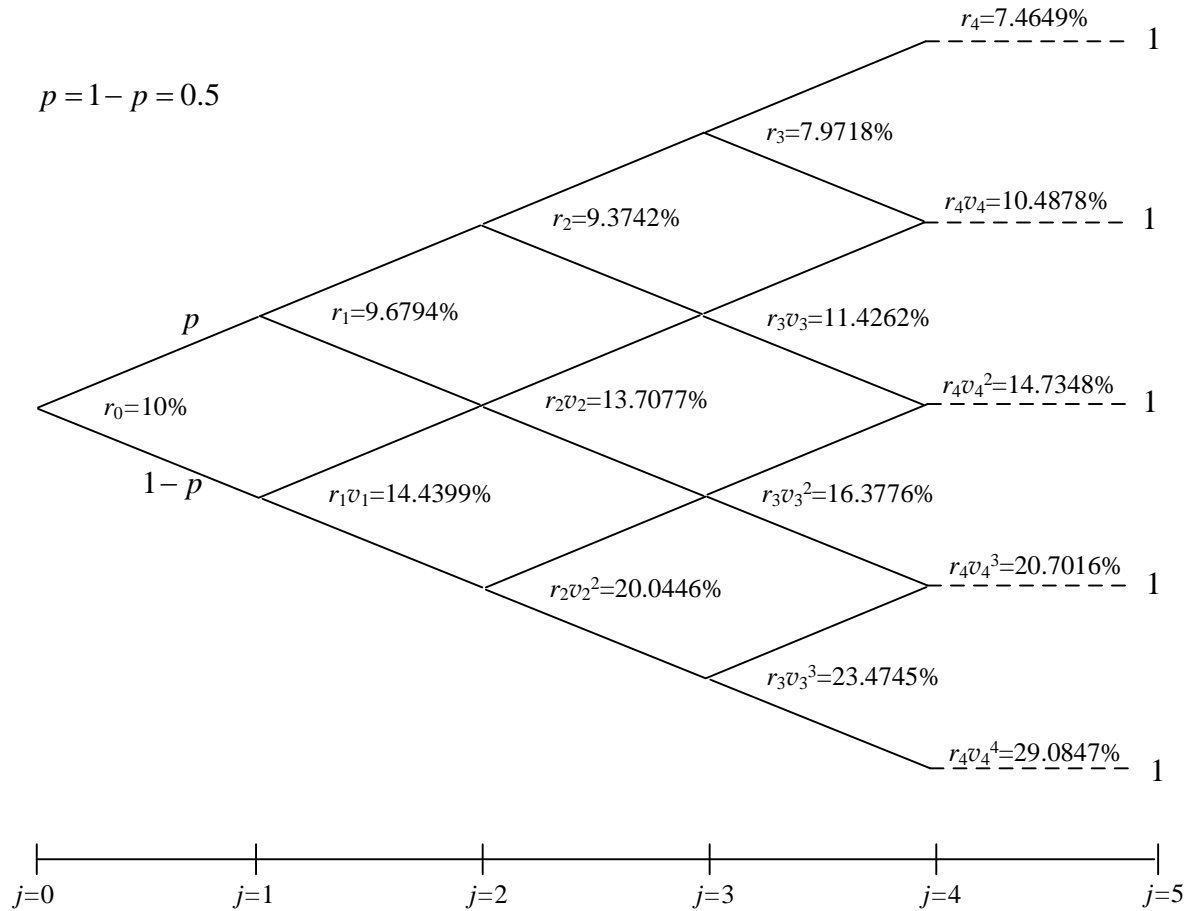
$$\frac{r_h}{r_l} = e^{2\sigma_t\sqrt{\Delta t}} = v_t$$

⊙ Given the 5-year information of the yield curve and the future local volatilities as follows ($\Delta t = 1\text{yr}$):

T (indexed by j)	0	1	2	3	4	5
$r(0, j)$		0.1	0.11	0.12	0.125	0.13
$F(j, j + 1)$	0.1	0.12	0.14	0.14	0.15	
σ_j of $\ln r_j$, which is the short rate applied in $(j\Delta t, (j + 1)\Delta t]$		0.2 σ_1	0.19 σ_2	0.18 σ_3	0.17 σ_4	

* In the above table, $r(0, j)$ denotes the interest rate for the time to maturity $j\Delta t$ (which is observable in the market), $F(j, j + 1)$ is the implied forward rate in the period of $(j\Delta t, (j + 1)\Delta t]$ based on $r(0, j)$, and σ_j is the future local volatility of the short rate applied in the period of $(j\Delta t, (j + 1)\Delta t]$.

Figure 12-5 The Binomial Interest Rate Tree for the BDT model.



* Note that given σ_j for each period, we already know $v_j = e^{2\sigma_j\sqrt{\Delta t}}$, which can match the current volatility term structure. In addition, from the current term structure of interest rates, $r_0 = r(0, 1) = 10\%$. The remaining job is to decide r_j such that the current yield curve can be matched exactly.

⊙ Solve the base line interest rate r_j by matching the expected forward rate based on the binomial tree and the implied forward rate based on the current yield curve.

$$\Rightarrow \sum_{i=0}^j 2^{-j} \binom{j}{i} r_j (v_j)^i = F(j, j+1), \text{ for } j = 1, \dots, n-1.$$

(In the period of $(j\Delta t, (j+1)\Delta t]$, the expected forward rate equals the sum of the product the short rate and the corresponding probability of each node at $j\Delta t$.)

$$\Rightarrow r_j \cdot 2^{-j} \cdot \sum_{i=0}^j \binom{j}{i} (v_j)^i = F(j, j+1), \text{ for } j = 1, \dots, n-1.$$

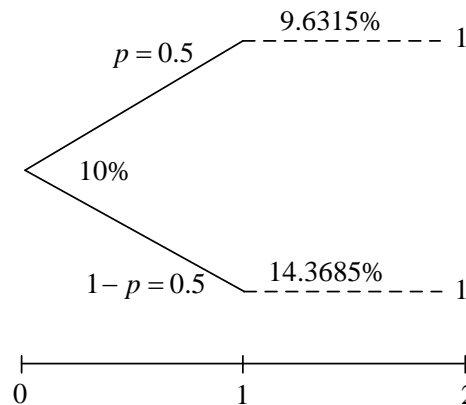
$$\left\| \sum_{i=0}^j \binom{j}{i} (v_j)^i \right. \\ \left. = \binom{j}{0} \cdot (v_j)^0 \cdot 1^j + \binom{j}{1} \cdot (v_j)^1 \cdot 1^{j-1} + \dots + \binom{j}{j} \cdot (v_j)^j \cdot 1^0 = (1 + v_j)^j \right.$$

$$\Rightarrow r_j = \left(\frac{2}{1+v_j}\right)^j \cdot F(j, j+1), \text{ for } j = 1, \dots, n-1.$$

⊙ The above method is fast, but matching the expected and real forward rates discretely cannot guarantee the BDT interest rate tree to match the current term structure of interest rates exactly.

* Based on the current term structure of interest rates, $r(0, j)$, we can derive the market price of the 2-year zero-coupon bond should be $P(0, 2) = \frac{1}{(1+r(0,2))^2} = \frac{1}{(1.11)^2} = 0.8116$. However, applying the tree constructed through the above fast method cannot generate the correct price of the 2-year zero-coupon bond (see Figure 12-6).

Figure 12-6



The expected forward rate for the second year is $0.5 \cdot 9.6315\% + 0.5 \cdot 14.3685\% = 12\%$.

The price of the 2-year zero-coupon bond based on the above binomial tree is

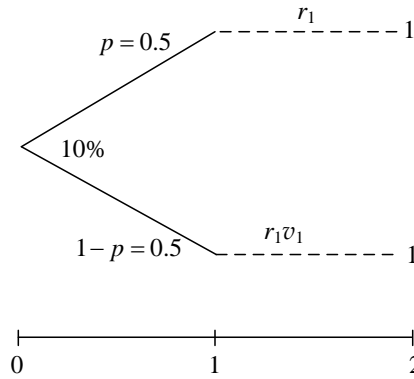
$$P(0, 2) = \frac{1}{1+0.1} \left[\frac{1}{2} \cdot \frac{1}{1+9.6315\%} + \frac{1}{2} \cdot \frac{1}{1+14.3685\%} \right] = 0.8121,$$

which does not equal the correct one based on the current term structure of interest rates. This result demonstrates that the binomial tree constructed by the fast method cannot match the current term structure of interest rates exactly.

⊙ The bootstrap method to calibrate r_j by matching the current yield curve.

Figure 12-7 Calibration for r_1

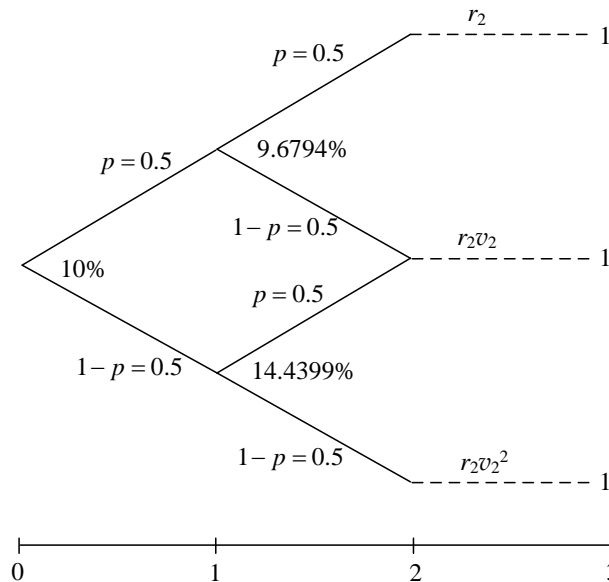
$$v_1 = e^{2\sigma_1\sqrt{\Delta t}} = 1.4918$$



$\Rightarrow \frac{1}{1+0.1} \left[\frac{1}{2} \cdot \frac{1}{1+r_1} + \frac{1}{2} \cdot \frac{1}{1+r_1 v_1} \right] = 0.8116$. Since v_1 is known, solving the equation to derive $r_1 = 9.6794\%$.

Figure 12-8 Calibration for r_2

$$v_2 = e^{2\sigma_2\sqrt{\Delta t}} = 1.4623$$



$\Rightarrow \frac{1}{1+0.1} \left\{ \frac{1}{2} \frac{1}{1+9.6794\%} \left[\frac{1}{2} \frac{1}{1+r_2} + \frac{1}{2} \frac{1}{1+r_2 v_2} \right] + \frac{1}{2} \frac{1}{1+14.4399\%} \left[\frac{1}{2} \frac{1}{1+r_2 v_2} + \frac{1}{2} \frac{1}{1+r_2 (v_2)^2} \right] \right\} = \frac{1}{(1+r(0,3))^3} = 0.7118$. Solve this equation for the only unknown, r_2 , which is 9.3742% .

* A similar interest rate model proposed by Black and Karasinski (1991) is $\ln r = [\theta(t) - a(t) \ln(r)]dt + \sigma(t)dZ$. However, the introduction of $a(t)$ makes this model more general but also more difficult to use. This is because how to calibrate the term $a(t)$ is a problem. Black and Karasinski suggest that in addition to the yield curve and the future local volatility $\sigma(t)$, the volatilities for the yields with different maturities are also needed for parameter calibration.

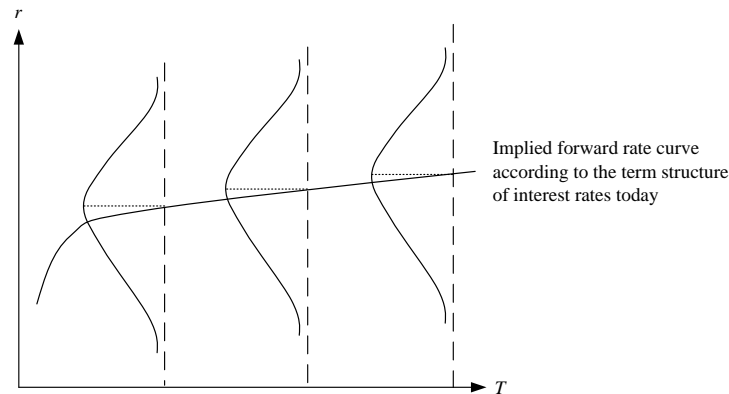
- Ho-Lee model: the first no-arbitrage model proposed in 1986.

$$dr = \theta(t)dt + \sigma dZ,$$

where $\theta(t)$ is determined by today's term structure of interest rates. Theoretically speaking $\theta(t) = F_t(0, t) + \sigma^2 t$, where $F(0, t)$ is the instantaneous forward rate at t based on the current term structure of interest rates. In fact, $\theta(t) \approx F_t(0, t)$ since the term $\sigma^2 t$ is relatively smaller.

* Note that $E[r_{t+\Delta t}] - E[r_t] = \theta(t)\Delta t$ by the definition of the Ho-Lee model. With $\theta(t) \approx F_t(0, t)$, it can be inferred that $E[r_{t+\Delta t}] - E[r_t] \approx F(0, t + \Delta t) - F(0, t)$. Therefore, the expected change in the short rate is approximately equal to the change of the instantaneous forward rate. Moreover, since $r(0) = F(0, 0)$ by definition, the mean level of $r(t)$ is approximately equal to the level of $F(0, t)$ as shown in Figure 12-9.

Figure 12-9 The Graphical Illustration of the Ho-Lee Model



* The details to derive the expression of $\theta(t)$ for various no-arbitrage models can refer to Hull and White (1993), "One-Factor Interest-Rate Models and the Valuation of Interest-Rate Derivative Securities," *Journal of Financial and Quantative Analysis* 28, pp. 235–254.

⊙ The theoretical value of the zero-coupon bond based on the Ho-Lee model is

$$P(t, r, T) = A(t, T)e^{-r(t)(T-t)},$$

where $\ln A(t, T) = \ln \frac{P(0, T)}{P(0, t)} + (T - t)F(0, t) - \frac{1}{2}\sigma^2 t(T - t)^2$.

* Ho and Lee (1986) also propose a binomial tree model to simulate the change of the term structure of interest rates. In addition, they assume the changes of the bond prices follow the implied forward bond price plus a perturbation. The details can refer to Ho and Lee (1986), “Term Structure Movements and Pricing Interest Rate Contingent Claims,” *Journal of Finance* 41, pp. 1011–1029.

* The disadvantage of the Ho and Lee’s (1986) model is the lack of the feature of mean reversion, which is commonly found in the movements of interest rates. Next, the Hull and White interest rate model is introduced, in which the Ho-Lee model is extended to incorporate the feature of mean reversion.

• Hull-White model:

$$\begin{aligned} dr &= [\theta(t) - ar]dt + \sigma dZ \\ &= a\left[\frac{\theta(t)}{a} - r\right]dt + \sigma dZ \end{aligned}$$

(The Hull-White model can be viewed as the Ho-Lee model plus the mean reversion with the speed a or as an extension of Vasicek model to be a no-arbitrage model.)

⊙ $\theta(t)$ is determined based on today’s term structure of interest rates, and its continuous-time equivalence is

$$\theta(t) = F_t(0, t) + aF(0, t) + \frac{\sigma^2}{2a}(1 - e^{-2at}).$$

⊙ The theoretical value of the zero-coupon bond is

$$P(t, r, T) = A(t, T)e^{-B(t, T)r(t)},$$

where $B(t, T) = \frac{1 - e^{-a(T-t)}}{a}$, and $\ln A(t, T) = \ln \frac{P(0, T)}{P(0, t)} + B(t, T)F(0, t) - \frac{1}{4a^3}\sigma^2(e^{-aT} - e^{-at})^2(e^{2at} - 1)$.

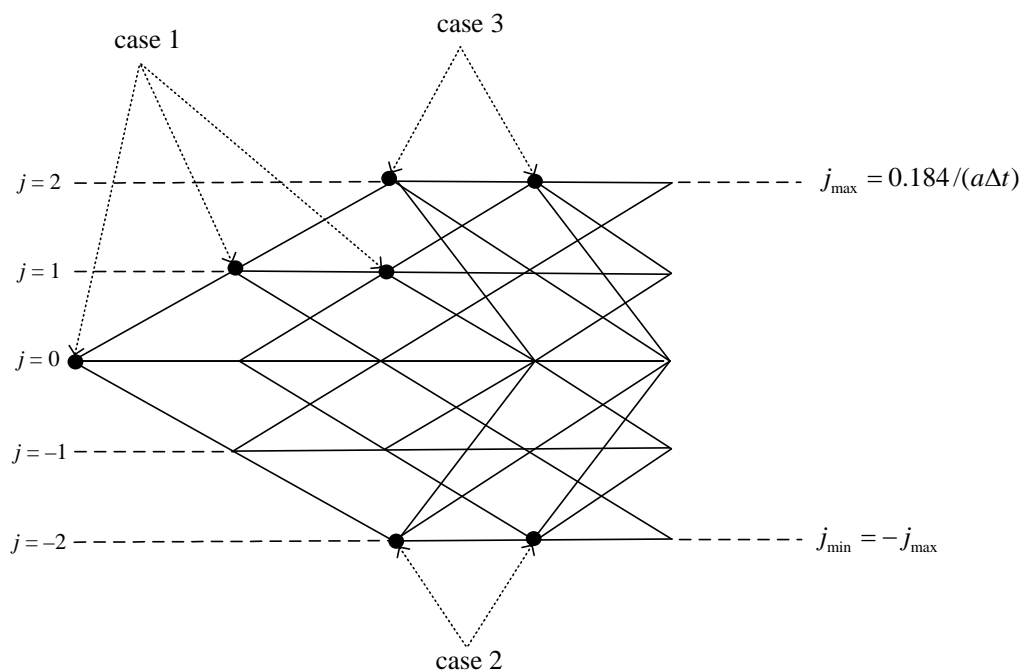
⊙ The trinomial tree approach to simulate the Hull-White interest rate process:

(The details of this trinomial tree model can refer to Hull and White (1994), “Numerical Procedure for Implementing Term Structure Models I: Single-Factor Models, *Journal of Derivatives* 2, pp.7–16)

Consider the discrete version $\Delta R^* = [\theta(t) - aR^*]\Delta t + \sigma\Delta Z$, in which the length of the time step is Δt .

First Stage: to build the trinomial tree for $\Delta R = -aR\Delta t + \sigma\Delta Z$. Note that the long term mean-reverting level for R is zero. The trinomial tree for R is illustrated in Figure 12-10.

Figure 12-10

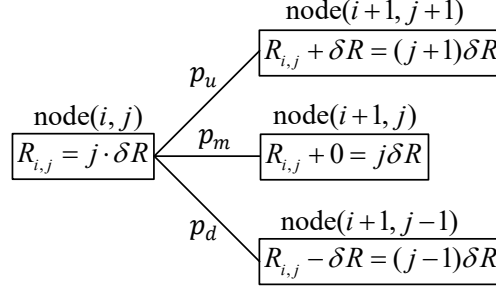


* The spacing parameter for the process R is defined as $\delta R \equiv \sigma\sqrt{3\Delta t}$ for error minimization. Due to the mean reverting feature, the trinomial tree is truncated when the level of $R(t)$ is too high or too low. Hull and White suggest that maximum deviation of $R(t)$ from zero (corresponding to the layer indexed by $j = 0$) is $j_{\max}\delta R$, where $j_{\max} \equiv 0.184/(a\Delta t)$.

* Define node(i, j) to be the node at $t = i\Delta t$ and the level of $R_{i,j}$ is $j\delta R$. Three branching cases on the Hull-White trinomial interest rate tree are as follows.

◦ Case 1: the normal situation.

Figure 12-11

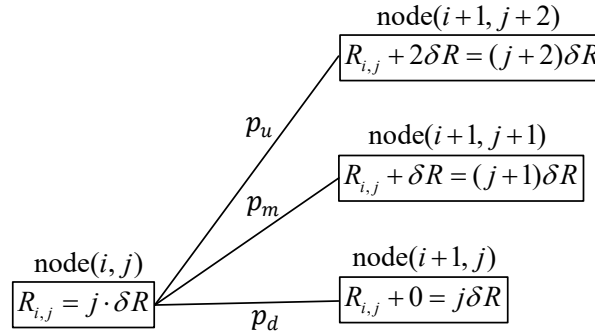


$$\begin{cases} p_u \cdot \delta R - p_d \cdot \delta R &= -aj\delta R\Delta t & \text{(matching mean)} \\ p_u \cdot (\delta R)^2 + p_d \cdot (\delta R)^2 &= \sigma^2\delta t + a^2j^2(\delta R)^2(\Delta t)^2 & \text{(matching variance)} \\ p_u + p_m + p_d &= 1 & \text{(total probability equal to 1)} \end{cases}$$

$$\Rightarrow \begin{cases} p_u &= \frac{1}{6} + \frac{a^2j^2(\Delta t)^2 - aj\Delta t}{2} \\ p_m &= \frac{2}{3} - a^2j^2\Delta t^2 \\ p_d &= \frac{1}{6} + \frac{a^2j^2(\Delta t)^2 + aj\Delta t}{2} \end{cases}$$

◦ Case 2: $R_{i,j} = j\delta R$ is negatively too small such that the mean of ΔR , $-aj\delta R\delta t$, is too large for the normal case.

Figure 12-12

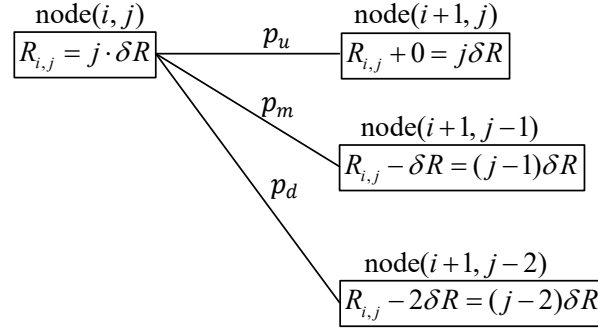


$$\begin{cases} p_u \cdot (2\delta R) + p_m \cdot \delta R &= -aj\delta R\Delta t & \text{(matching mean)} \\ p_u \cdot (2\delta R)^2 + p_m \cdot (\delta R)^2 &= \sigma^2\delta t + a^2j^2(\delta R)^2(\Delta t)^2 & \text{(matching variance)} \\ p_u + p_m + p_d &= 1 & \text{(total probability equal to 1)} \end{cases}$$

$$\Rightarrow \begin{cases} p_u &= \frac{1}{6} + \frac{a^2j^2(\Delta t)^2 + aj\Delta t}{2} \\ p_m &= -\frac{1}{3} - a^2j^2(\Delta t)^2 - 2aj\Delta t \\ p_d &= \frac{7}{6} + \frac{a^2j^2(\Delta t)^2 + 3aj\Delta t}{2} \end{cases}$$

◦ Case 3: $R_{i,j} = j\delta R$ is too high such that the mean of ΔR , $-aj\delta R\delta t$, is too negative for the normal case.

Figure 12-13



$$\begin{cases} p_m \cdot (-\delta R) + p_d \cdot (-2\delta R) &= -aj\delta R\Delta t & \text{(matching mean)} \\ p_m \cdot (-\delta R)^2 + p_d \cdot (-2\delta R)^2 &= \sigma^2\delta t + a^2j^2(\delta R)^2(\Delta t)^2 & \text{(matching variance)} \\ p_u + p_m + p_d &= 1 & \text{(total probability equal to 1)} \end{cases}$$

$$\Rightarrow \begin{cases} p_u &= \frac{7}{6} + \frac{a^2j^2(\Delta t)^2 - 3aj\Delta t}{2} \\ p_m &= -\frac{1}{3} - a^2j^2(\Delta t)^2 + 2aj\Delta t \\ p_d &= \frac{1}{6} + \frac{a^2j^2(\Delta t)^2 - aj\Delta t}{2} \end{cases}$$

Second Stage: to consider the difference between R^* and R .

Define $\alpha(t) = R^*(t) - R(t)$, and assume the continuous limits of ΔR^* and ΔR are as follows.

$$\begin{cases} dR^* &= [\theta(t) - aR^*]dt + \sigma dZ \\ dR &= -aRdt + \sigma dZ \end{cases}$$

$$\begin{aligned} \Rightarrow d\alpha &= dR^* - dR \\ &= [\theta(t) - aR^* + aR]dt \\ &= [\theta(t) - a(R^* - R)]dt \\ &= [\theta(t) - a\alpha]dt \end{aligned}$$

Since we already know $\theta(t) = F_t(0, t) + aF(0, t) + \frac{\sigma^2}{2a}(1 - e^{-2at})$, we can solve $\alpha(t)$ from the above differential equation.

$$\begin{cases} \alpha(t) = F(0, t) + \frac{\sigma^2}{2a^2}(1 - e^{-at})^2 & \text{if } a \neq 0 \\ \alpha(t) = F(0, t) + \frac{\sigma^2 t^2}{2} & \text{if } a = 0 \end{cases}$$

* If we add $\alpha(t)$, which is a function of t , to each time- t node on the R -tree, we can derive the R^* -tree. However, due to the discretization error for the trinomial tree model, the resulting R^* -tree cannot match the current term structure of interest rates exactly.

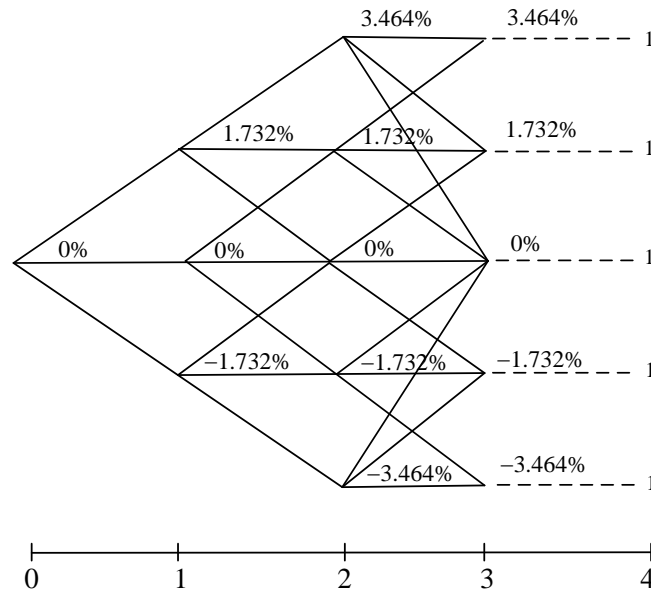
* Moreover, if we ignore the second term in the formula of $\alpha(t)$, it can be inferred that $\alpha(t)$ is approximately equal to the instantaneous forward rate $F(0, t)$. In other words, we can derive R^* -tree approximately by adding the forward rate of each period to the corresponding nodes on the R -tree.

* In order to achieve the exact match with the current term structure of interest rates, the following method is considered.

⊙ The bootstrap method to calibrate $\alpha(t)$ by matching the current yield curve.

Consider a numerical example in which $\sigma = 0.01$, $a = 0.1$, $\Delta t = 1$, and we can derive $\delta R = 0.01\sqrt{3} = 1.732\%$, $j_{\max} = \lceil \frac{0.184}{0.1} \rceil = 2 = -j_{\min}$. After the first stage, the tree for R is in Figure 12-14.

Figure 12-14 R -tree after the first stage

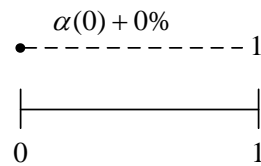


Suppose the term structure of interest rates today as follows.

T	1	2	3	4
$r(0, T)$	3.824%	4.512%	5.086%	5.566%

The following goal is to find $\alpha(t)$ by which is the upward shifting level of $R(t)$ such that the resulting R^* -tree can match today's term structure of interest rates exactly.

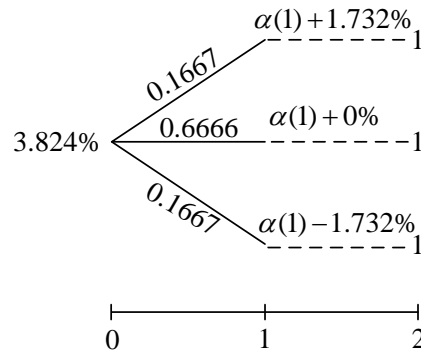
Figure 12-15 Calibration of $\alpha(0)$



Solve $\alpha(0)$ by matching the 1-year zero-coupon bond prices based the yield curve and based on the trinomial tree.

$$P(0, 1) = e^{-3.824\% \cdot 1} = e^{-\alpha(0) \cdot 1} \Rightarrow \alpha(0) = 3.824\%$$

Figure 12-16 Calibration of $\alpha(1)$



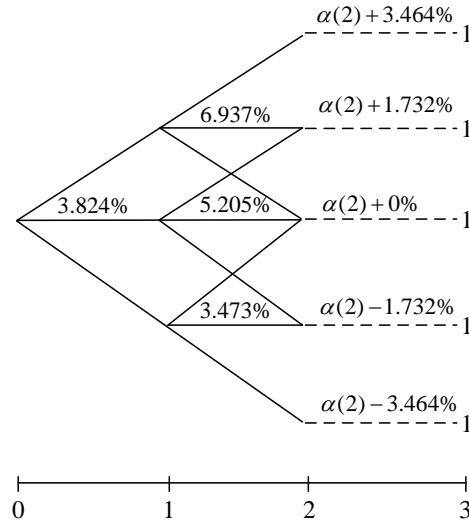
Match the 2-year zero-coupon bond prices to derive

$$P(0, 2) = e^{-4.512\% \cdot 2} = e^{-3.824\%} [0.1667e^{-(\alpha(1)+1.732\%)} + 0.6666e^{-(\alpha(1)+0\%)} + 0.1667e^{-(\alpha(1)-1.732\%)}].$$

Solve the above equation to derive $\alpha(1) = 5.205\%$.

(Note that $\alpha(1)$ is very close to the forward rate in the second year, which is 5.200%.)

Figure 12-17 The Calibration of $\alpha(2)$



Considering the 3-year zero coupon bond price $P(0, 3)$, you can solve for $\alpha(2) = 6.252\%$. As to $\alpha(t)$ for $t > 2$, the desired results can be derived by performing the similar bootstrap method recursively.

(Note that $\alpha(2)$ is very close to the forward rate in the third year, which is 6.234%.)

⊙ The approximation of $\theta(t)$ based on $\alpha(t)$ derived in the bootstrap method.

Note that the interest rate levels of the midpoint nodes of the R^* -tree are $\alpha(0)$, $\alpha(1)$, and so on. Thus the drift term of these nodes is, by definition, $(\theta(t) - ar)\Delta t = (\theta(t) - a\alpha(t))\Delta t$. As a consequence, we can obtain,

$$\begin{aligned} (\theta(t) - a\alpha(t))\Delta t &= \alpha(t + \Delta t) - \alpha(t) \\ \Rightarrow \hat{\theta}(t)_{\text{mid}} &= \frac{\alpha(t + \Delta t) - \alpha(t)}{\Delta t} + a\alpha(t), \end{aligned}$$

where $\hat{\theta}(t)_{\text{mid}}$ is the approximation of $\theta(t)$ based on the midpoint nodes because the above equation is true only for the midpoint nodes. Since the tree is symmetric around the midpoint nodes and thus the midpoint nodes are the most representative nodes, $\hat{\theta}(t)_{\text{mid}}$ can be a practical approximation of $\theta(t)$.

Comparing with the continuous limit formula of $\theta(t) = F_t(0, t) + aF(0, t) + \frac{\sigma^2}{2a}(1 - e^{-2at})$, the above approximation is in effect to consider the discrete approximation of the first two terms of the continuous limit formula of $\theta(t)$ because it is known that $\alpha(t) \approx F(0, t)$.

III. Forward Rate Models

- Heath, Jarrow, and Morton (HJM) model:

Instead of considering the short rate process, Heath, Jarrow, and Morton (1992) focus on the instantaneous forward rate process. Note that the instantaneous forward rate is more general than the short rate due to its additional maturity dimension. Therefore, the short rate can be viewed as the special case of the instantaneous forward rate matured immediately.

⊙ Define $f(t, T_1, T_2)$ to be the forward rate for the period of $[T_1, T_2]$ observed at t , and $F(t, T)$ to be the instantaneous forward rate for the period of $[T, T + dt]$ observed at t . By definition,

$$F(t, T) = \lim_{\substack{T_1 \rightarrow T \\ T_2 \rightarrow T+dt}} f(t, T_1, T_2).$$

Moreover, denote $r(t)$ as the short rate in the following infinitesimal period $[t, t + dt]$. According to the definition of $F(t, T)$, we can derive

$$r(t) = F(t, t).$$

- The most important result of the HJM model is to discover the relationship of the drift and volatility terms of the instantaneous forward rate process, $dF(t, T)$, and their connection with the volatility of the zero-coupon bond price process.
- ⊙ Suppose the price process of the zero-coupon bond $P(t, T)$ under the risk neutral measure as follows.

$$dP(t, T) = r(t)P(t, T)dt + v(t, T)P(t, T)dZ,$$

where $v(t, T)$ is the volatility of the bond price. Since the bond price at maturity is known, we must have

$$v(t, t) = v(T, T) = 0.$$

* The forward rate process can be expressed with the zero-coupon bond price process as follows.

$$df(t, T_1, T_2) = \frac{d \ln(P(t, T_1)) - d \ln(P(t, T_2))}{T_2 - T_1},$$

where the processes $d \ln(P(t, T_1))$ and $d \ln(P(t, T_2))$ can be derived by the Itô's Lemma as follows.

$$\begin{aligned} d \ln(P(t, T_1)) &= [r(t) - \frac{v(t, T_1)^2}{2}]dt + v(t, T_1)dZ, \\ d \ln(P(t, T_2)) &= [r(t) - \frac{v(t, T_2)^2}{2}]dt + v(t, T_2)dZ. \end{aligned}$$

As a result,

$$df(t, T_1, T_2) = \frac{v(t, T_2)^2 - v(t, T_1)^2}{2(T_2 - T_1)} dt + \frac{v(t, T_1) - v(t, T_2)}{T_2 - T_1} dZ.$$

* According to the straightforward differential calculation, we can derive

$$\frac{1}{2} \frac{\partial [v(T, T)^2]}{\partial T} = v(t, T) v_T(t, T),$$

where $v_T(t, T)$ denotes the partial derivative with respect to T . Next we consider $T_1 = T$ and $T_2 = T + dt$, and thus can rewrite the process $df(t, T_1, T_2)$ as follows.

$$dF(t, T) = v(t, T) v_T(t, T) dt - v_T(t, T) dZ.$$

If we define $s(t, T)$ as $-v_T(t, T)$, then the drift term $v(t, T) v_T(t, T)$ can be expressed as follows.

$$v(t, T) v_T(t, T) = \left(- \int_t^T s(t, \tau) d\tau \right) (-s(t, T)) = s(t, T) \int_t^T s(t, \tau) d\tau.$$

⊙ The key result and the problems of the HJM model:

$$dF(t, T) = \left(s(t, T) \int_t^T s(t, \tau) d\tau \right) dt + s(t, T) dZ.$$

⊙ From the above equation, we can find the drift term of the forward rate process depends on the historical of its volatility terms. Hence, the forward rate as well as the corresponding short rate processes are non-Markovian, which makes the HJM model difficult to use in practice. Moreover, the volatility term $s(t, T)$ is not constant, and thus the tree model for the HJM is nonrecombined. It is well known that due to the constraint of the memory space of computers, the nonrecombined tree is infeasible when the number of time steps is large.

⊙ Another disadvantage of the HJM model is that the instantaneous forward rate is unobservable in financial markets. This disadvantage inspires the development of the Brace, Gatarek, and Musiela (1997) model, in which the process of $df(t, T_1, T_2)$ rather than $dF(t, T)$ is considered.

- Brace, Gatarek, and Musiela (BGM) model

Brace, Gatarek, and Musiela propose this (long) forward rate model in 1997, and this model is also termed the LIBOR market model (LMM) since it is widely applied for pricing LIBOR-based derivatives. The BGM model is extremely complicated. More details can refer to Brace, Gatarek, and Musiela (1997), “The Market Model of Interest Rate Dynamics,” *Mathematical Finance* 7, pp. 127–155. Or you can read Chapter 31 of Hull (2011), in which Hull introduces a simpler version of the BGM model and the derivation in Hull (2011) is much easier to understand. Also due to its complexity, there is no other available numerical method to implement the BGM interest rate model except the Monte Carlo simulation.

IV. Credit Risk Models

- This section introduces two categories of credit risk (or default) models: the reduced-form and structural models. In reduced-form models, the default is triggered by a Poisson jump and thus occurs unpredictably. In contrast, the structural models takes the asset and liability values of a firm into consideration and thus the default probability is correlated with the level and the volatility of the liability ratio of a firm.

- Reduced-form models:

- ◉ Default probability and default intensity:

$$\lambda_t = 1 - e^{-\xi_t \Delta t},$$

where λ_t is the default probability in $(t, t + \Delta t]$, ξ_t is the default intensity and usually modeled as a Poisson process.

- ◉ Based on the reduced-form model, there exists an explicit relationship between the default probability and the credit spread of bonds:

* Consider a 1-year risky zero coupon bond with the face value of \$1. Its default probability is λ in the future one year. If the risky bond defaults, the investor loses L out of the \$1 principal at maturity. Thus, the value of this 1-year risky zero coupon bond can be expressed as follows.

$$\begin{aligned} e^{-R} &= (1 - \lambda) \cdot e^{-r} + \lambda \cdot e^{-r}(1 - L) \\ \Rightarrow 1 - R &\approx (1 - \lambda)(1 - r) + \lambda(1 - r)(1 - L) \\ &= 1 - \lambda - r + \lambda r + \lambda - \lambda r - \lambda L + \lambda r L \\ \Rightarrow 1 - R &= 1 - r - \lambda L + \lambda r L \quad (\lambda r L \text{ is relatively smaller and thus ignored}) \\ \Rightarrow R &\approx r + \lambda L \end{aligned}$$

* As a result, the credit spread $S = R - r$ equals approximately the product of the default probability, λ , and the loss rate given default, L . In other words, if you are given the information of the credit spread from the market and the estimation of the loss rate given default, the default probability of the risky bond can be implied from the above equation.

- A tree approach to simulate the riskless and risky interest rate processes:

Here I introduce a tree-based reduced-form model to simulate the riskless and risky interest rate processes. This method is proposed by Jarrow and Turnbull (1995), "Pricing Derivatives on Financial Securities Subject to Credit Risk," *Journal of Finance* 50, pp. 53–85.

⊙ Given the market prices of the riskless and risky zero-coupon bonds, $P(0, T)$ and $V(0, T)$ for different maturity dates T , and the recovery rate $\delta = 1 - L$, one can derive the default probability of the reference firm for each period.

Figure 12-18 One-Period Risk-free Interest Rate Tree

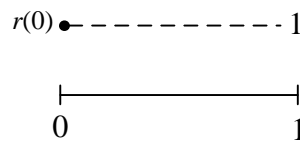
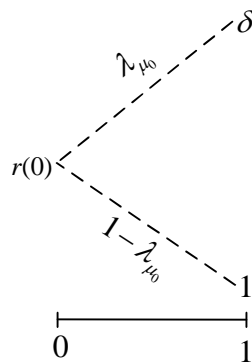


Figure 12-19 One-Period Risky Interest Rate Tree



$$\begin{aligned} V(0, 1) &= e^{-r(0)}[\lambda_{\mu_0} \cdot \delta + (1 - \lambda_{\mu_0}) \cdot 1] \\ &= P(0, 1)[\lambda_{\mu_0} \cdot \delta + (1 - \lambda_{\mu_0}) \cdot 1] \end{aligned}$$

Given the market prices of $V(0, 1)$ and $P(0, 1)$, solve λ_{μ_0} , the default probability of the first period (year), from the above equation.

Figure 12-20 Two-Period Risk-free Interest Rate Tree

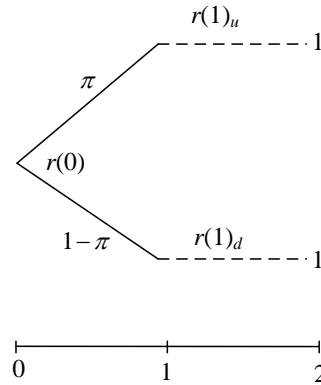
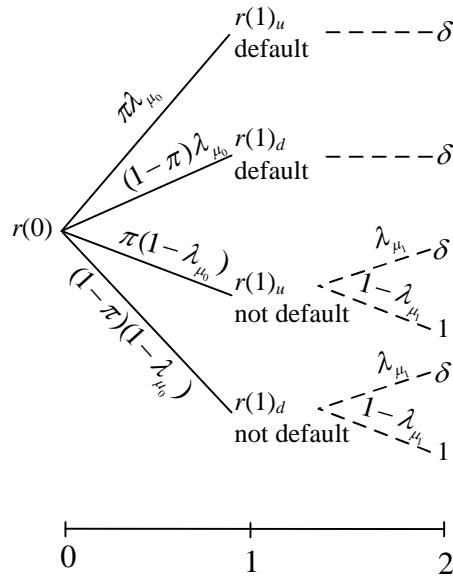


Figure 12-21 Two-Period Risky Interest Rate Tree



$$\begin{aligned}
 V(0, 2) &= e^{-r(0)} \{ e^{-r(1)_u} \cdot \pi \cdot \lambda_{\mu_0} \cdot \delta + \\
 &\quad e^{-r(1)_d} \cdot (1 - \pi) \cdot \lambda_{\mu_0} \cdot \delta + \\
 &\quad e^{-r(1)_u} \cdot \pi \cdot (1 - \lambda_{\mu_0}) \cdot [\lambda_{\mu_1} \cdot \delta + (1 - \lambda_{\mu_1}) \cdot 1] + \\
 &\quad e^{-r(1)_d} \cdot (1 - \pi) \cdot (1 - \lambda_{\mu_0}) \cdot [\lambda_{\mu_1} \cdot \delta + (1 - \lambda_{\mu_1}) \cdot 1] \} \\
 &= P(0, 2) \{ \lambda_{\mu_0} \cdot \delta + (1 - \lambda_{\mu_0}) \cdot [\lambda_{\mu_1} \cdot \delta + (1 - \lambda_{\mu_1}) \cdot 1] \},
 \end{aligned}$$

where $P(0, 2) = e^{-r(0)} [e^{-r(1)_u} \cdot \pi \cdot 1 + e^{-r(1)_d} \cdot (1 - \pi) \cdot 1]$ can be derived according to Figure 12-20.

Since $P(0, 2)$, $V(0, 2)$, and λ_{μ_0} is known, it is straightforward to find the only unknown λ_{μ_1} by solving the above equation. This bootstrap method to solve the default probability of each period, λ_{μ_t} , can be performed recursively given different T .

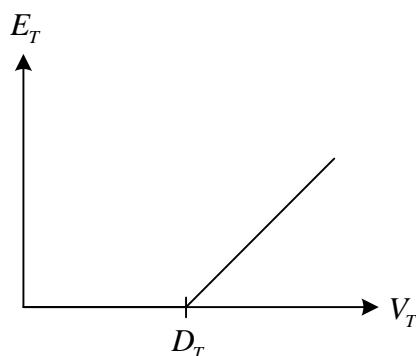
- Structural model

- ⊙ One of the most classical structural model is the KMV credit model, which was proposed by Kealhofer, McQuown, and Vasicek in 1989. This credit risk model was acquired by the credit rating agency Moody's in 2002 and thus renamed as Moody's KMV.

- ⊙ The basic idea of the KMV model is simple: the firm equity value can be viewed as a call option on the firm asset value.

- * Since the priority of the debt holder's claim on the firm asset value is higher than the equity holder, on the maturity debt of the corporate debt, what the equity holder can receive is the firm asset value in excess of the firm debt. Once the firm asset value is lower than the face value of the firm debt at maturity, the default occurs.

Figure 12-22



- * The payoff of the equity holder is depicted in the above figure, which is the same as the the payoff of the call option if we treat the firm asset value as the underlying asset. Since the firm equity value can be viewed as a call option on the firm asset value, it can be priced with the Black-Scholes model.

- ⊙ Notation system:

V_0 : the firm asset value today

E_0 : the firm equity value today (as a call option on V_T with a strike price D_T)

D_T : the face value of the firm debt needed to repay at T

σ_E : the volatility of the firm equity value

σ_V : the volatility of the firm asset value

Based on the Black-Scholes formula, we have

$$E_0 = V_0 N(d_1) - D_T e^{-rT} N(d_2), \quad (1)$$

where $d_1 = \frac{\ln(\frac{V_0}{D_T}) + (r + \frac{\sigma_V^2}{2})T}{\sigma_V \sqrt{T}}$, and $d_2 = d_1 - \sigma_V \sqrt{T}$.

* In the real world, we know E_0 and can estimate σ_E from historical stock prices. However, the value of V_0 is not observable and thus σ_V cannot be estimated in the traditional way. To solve for these two unknowns, we need one more equation in addition to Equation (1).

Suppose the processes of V and E follow the geometric Brownian motions (GBM) under the risk neutral measure as

$$dV = rVdt + \sigma_V V dZ,$$

$$dE = rEdt + \sigma_E E dZ.$$

Since E should be a function of V and t , according to the Itô's Lemma, we have

$$dE = (\dots)dt + \left(\frac{\partial E}{\partial V} V \sigma_V\right)dZ.$$

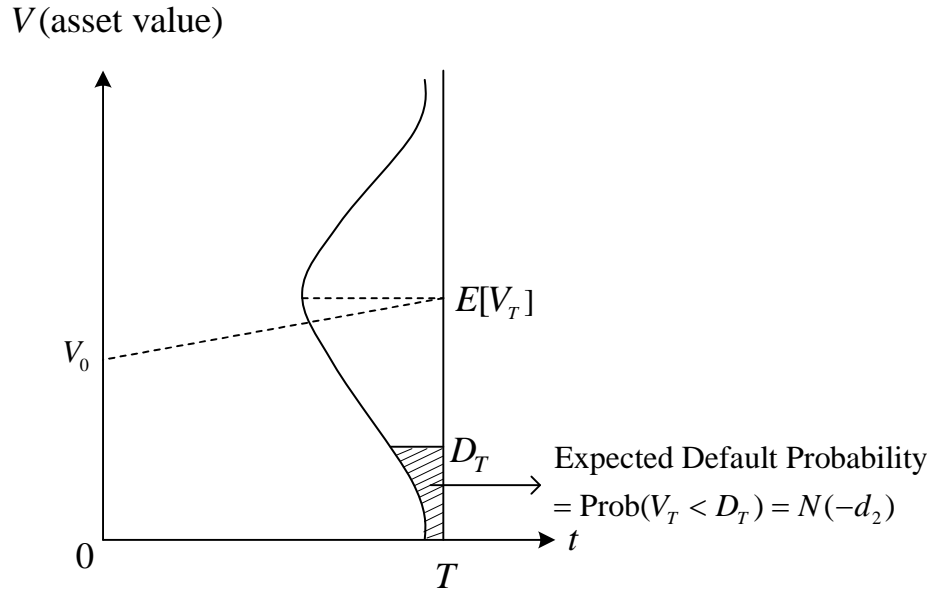
Compare with the two GBM processes for E , the following equation can be derived

$$\sigma_E E = \frac{\partial E}{\partial V} V \sigma_V \Rightarrow \sigma_E = \frac{\frac{\partial E}{\partial V} V \sigma_V}{E} = \frac{N(d_1) V \sigma_V}{E}. \quad (2)$$

Finally, we can solve the system consisting of Equations (1) and (2) for V_0 and σ_V .

Based on the results of V_0 and σ_V , we can estimate the default probability at T :

Figure 12-23 The Graphical Illustration for the KMV Model



$$\begin{aligned}
 \text{Expected default probability} &= \text{Prob}(V_T \leq D_T) = \text{Prob}(\ln V_T \leq \ln D_T) \\
 &= \text{Prob}(\ln V_0 + (r - \frac{\sigma_V^2}{2})T + \sigma_V \sqrt{T} \varepsilon_V \leq \ln D_T) \\
 &= \text{Prob}(\varepsilon_V \leq \frac{\ln(\frac{D_T}{V_0}) - (r - \frac{\sigma_V^2}{2})T}{\sigma_V \sqrt{T}}) \\
 &= \text{Prob}(\varepsilon_V \leq -d_2) = N(-d_2)
 \end{aligned}$$

* The determination of D_T : since the corporate bonds with different time to maturities may exist concurrently, a rule is needed to decide the value of D_T .

The KMV model suggests the consideration of T is set to be 1 year and the value of D_T equals the sum of the face value of the current liability and the half of the face value of the long-term liability.

⊙ Disadvantages of the KMV model:

(1) The generated expected default probabilities are significantly lower than the historical default probabilities. (risk-neutral vs. risk-averse default probabilities)

(2) The expected default probabilities for the very near future are almost zero. If the firm has not defaulted until now, it can be inferred that $V_0 \geq D_T$. Since the volatility $\sigma_V \sqrt{T}$ is almost zero when $T \rightarrow 0$, thus $V_T \geq D_T$ is almost sure such that the $\text{Prob}(V_T \leq D_T) \rightarrow 0$.

(3) The expected default probabilities approach zero when T approach infinity. This is because the firm asset value is with a constant growth rate of r , but the paid off debt level is a constant. With the passage of time, the increasing expected value of the firm asset will be in general higher than the debt level. As a result, the expected default probabilities approach zero when T approach infinity.

* To solve the second problem, usually the jump of the firm asset value is taken into account. For the third problem, a stationary debt ratio process is introduced in Collin-Dufresne and Goldstein (2001), “Do Credit Spreads Reflect Stationary Leverage Ratios,” *Journal of Finance* 56, pp. 1929–1957.

* Two solutions for the first problem are introduced.

(i) The widely adopted solution in practice is to take the default distance (DD) into consideration and next map the value of DD to the historical default probability.

$$\text{Default Distance (DD)} = \frac{E[V_T] - D_T}{E[V_T] \sigma_V}$$

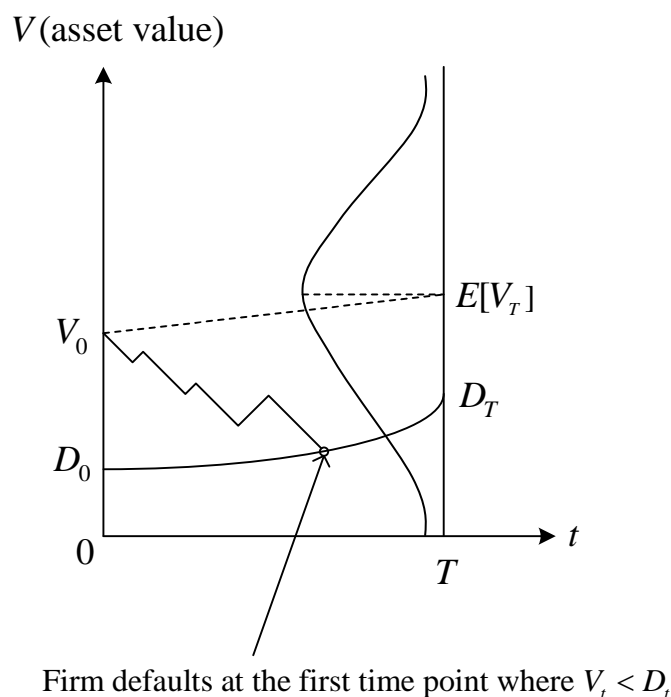
The relationship between the DD value of the KMV model and the historical default probability :

Default Distance	0-1	1-2	2-3	3-4	4-5	5-6
Historical Default Probability	80%	30%	10%	4.3%	0.7%	0.4%
No. of defaults	720	450	200	150	28	17
No. of firm samples	900	1500	2000	35000	40000	42000

* Note that the figures in this table are not real.

(ii) Another solution is to consider the possibility that the firm could default before the maturity of the debt. This method is more appropriate to describe the default behavior in the real world, and in the meanwhile, this method theoretically generates a higher expected default probability than the KMV model does.

Figure 12-24 The Graphical Illustration of the First Passage Time Model



Note that D_t is interpreted as the default threshold at t . The default threshold is highly correlated but not necessarily equal to the debt level of the firm.

There exist closed-form solutions for the expected default probability based on the first passage time model with a constant or exponential default threshold by regarding the equity price as the value of a barrier call option on the firm asset value, for example, see Black and Cox (1976) or Longstaff and Schwartz (1995).