Financial Mathematics 33000

Lecture 6

Roger Lee

2022 November 9

Fundamental theorem in continuous time

Black-Scholes

Fundamental theorem for continuous time Itô processes

$\boxed{\exists \text{ equivalent martingale measure } \mathbb{P} \Rightarrow \text{No arb}}$

Proof:

- Given: under $\mathbb{P}$, each member of the discounted asset prices

$$\frac{1}{B_t}\mathbf{X}_t := \left(\frac{X_t^1}{B_t}, \frac{X_t^2}{B_t}, \cdots, \frac{X_t^N}{B_t}\right)$$

  is a martingale Itô process. Need to prove: No arbitrage

- For any self-financing trading strategy $\boldsymbol{\Theta}_t$ with value V_t,

  we'll show that V/B is a martingale, where $V_t := \boldsymbol{\Theta}_t \cdot \mathbf{X}_t$.

  After we show this, we're done, by the familiar argument:

  If $V_0 = 0$, then $V_0/B_0 = 0$, hence $\mathbb{E}(V_T/B_T) = 0$.

  If also $V_T \geq 0$, then $V_T/B_T \geq 0$, so $V_T/B_T = 0$, hence $V_T = 0$

  Conclusion: $\boldsymbol{\Theta}$ is not an arbitrage.

Fundamental theorem

- To see that V/B is a martingale, let $A_t := 1/B_t$. Then

$$
\begin{aligned}
\mathrm{d}(V_t/B_t) = \mathrm{d}(A_tV_t) &= V_t\mathrm{d}A_t + A_t\mathrm{d}V_t + \mathrm{d}A_t\mathrm{d}V_t \\
&= \boldsymbol{\Theta}_t \cdot \mathbf{X}_t\mathrm{d}A_t + A_t(\boldsymbol{\Theta} \cdot \mathrm{d}\mathbf{X}_t) + (\mathrm{d}A_t)(\boldsymbol{\Theta}_t \cdot \mathrm{d}\mathbf{X}_t) \\
&= \boldsymbol{\Theta}_t \cdot (\mathbf{X}_t\mathrm{d}A_t + A_t\mathrm{d}\mathbf{X}_t + \mathrm{d}A_t\mathrm{d}\mathbf{X}_t) \\
&= \boldsymbol{\Theta}_t \cdot \mathrm{d}(A_t\mathbf{X}_t) = \sum_{n=1}^{N} \theta_t^n \mathrm{d}(A_tX_t^n)
\end{aligned}
$$

  Since each $A_tX_t^n$ is a martingale, V/B is a martingale also.

- Idea: A martingale is the cumulative PnL from betting on zero-$\mathbb{E}$ games. Varying your bet size across games and across time still produces, collectively, a zero-expectation game. Can't risklessly make something from nothing by playing zero-expectation games.

Fundamental theorem

$p(A) =$ PRICE of contract that pays: $\begin{cases} \$1 \text{ if } A \\ \$0 \text{ if not } A. \end{cases}$

$\boxed{\text{No arb} \Rightarrow \exists \text{ equivalent martingale measure } \mathbb{P}\text{:}}$

Intuition of proof: Same as in L2, L3.

- Define $\mathbb{P}$ by defining the $\mathbb{P}_t$-probability of an event to be the time-t **price**, in units of B, of an ("Arrow-Debreu") asset that pays 1 unit of B at time T if the event occurs, else 0. (But what if the A-D asset does not exist and can't be replcated?)
- Martingale property holds because any asset X can be replicated by portfolio of $X_T(\omega_j)/B_T(\omega_j)$ units of the A-D asset for each ω_j. Value portfolio by summing quantity $\times$ price.

$$\frac{X_0}{B_0} = \sum_{j=1}^{J} \frac{X_T(\omega_j)}{B_T(\omega_j)} \cdot \mathbb{P}(\omega_j) = \mathbb{E}\left(\frac{X_T}{B_T}\right)$$

Fundamental theorem: Comments

- Idea: The $\mathbb{P}$ probability of an event is simply the *price* (in units of B) of a asset that pays 1 unit of B iff that event occurs.
- Note: In this entire proof, we never assumed that B is the bank account, and never assumed that it is riskless. It is enough to assume that B is some asset with positive price process.

  In some applications, it may be easier to normalize using some such asset (some *numeraire*) that is *not* the bank account.

  By default, if we say risk-neutral or martingale measure without specifying the numeraire, it is understood to be the bank account.

Option pricing

In L5, we did this by *replication*.

In L6, let's do it by martingale methods: Option price equals the *expected discounted payoff*, under a martingale measure $\mathbb{P}$. Why?

- By the Fundamental theorem.

How do we calculate $\mathbb{P}$-expectations (denoted by $\mathbb{E}$)?

- In many cases, a model is already specified under risk-neutral measure. Then simply work directly under the given measure.

But what if the model is specified under physical measure?

- We know how S behaves with respect to physical measure P. How does S behave wrt $\mathbb{P}$? All risk driven by W. So let's see what changing measure does to W, then find what it does to S.

Girsanov's theorem

Theorem: If W is a Brownian motion under P,

and if $\mathbb{P}$ is a probability measure on $\mathcal{F}_T^W$ that is equivalent to P,

then there exists an adapted process λ such that for all $t \in [0, T]$,

$$\tilde{W}_t := W_t + \int_0^t \lambda_s \mathrm{d}s$$

is Brownian motion under $\mathbb{P}$. Therefore:

- $\mathrm{d}\tilde{W}_t = \mathrm{d}W_t + \lambda_t \mathrm{d}t$, and $\tilde{W}$ is BM under $\mathbb{P}$ but not under P
- $\mathrm{d}W_t = \mathrm{d}\tilde{W}_t - \lambda_t \mathrm{d}t$, and W is BM under P but not under $\mathbb{P}$

Girsanov: an analogy

skip

No proof, but here is an *analogy* on a sample space $\Omega = \{\omega_1, \ldots, \omega_6\}$.

Let $X(\omega_1) = X(\omega_2) = X(\omega_3) = 10$, $X(\omega_4) = X(\omega_5) = X(\omega_6) = 25$.

- Let $\mathsf{P}(\omega) = 1/6$ for each ω. Then $X \sim \text{Uniform}\{10, 25\}$ under P.
- But if $\mathbb{P}$ assigns probability

  $1/12$ to each of $\omega_1, \omega_2, \omega_3$,

  and $1/4$ to each of $\omega_4, \omega_5, \omega_6$,

  then X is *not* $\text{Uniform}\{10, 25\}$ under $\mathbb{P}$.
- However, $\tilde{X} := X + \lambda$ *is* $\text{Uniform}\{10, 25\}$ under $\mathbb{P}$,

  where $\lambda(\omega_4) := 15$ and $\lambda(\omega) := 0$ for $\omega \neq \omega_4$.

X under $\mathbb{P}$ does not have the same distribution as X under P.

But X *plus drift* under $\mathbb{P}$ has the same distribution as X under P.

Girsanov: some intuition

SKIP

No proof, but here is some *intuition*:

- W is BM under P. After changing measure to $\mathbb{P}$, the W may not still be BM, but it is plausible that it is a martingale plus drift:

$$\mathrm{d}W_t = \lambda_t \mathrm{d}t + \sigma_t \mathrm{d}B_t$$

  where B is a BM under $\mathbb{P}$, and σ_t is some adapted process. So

$$(\mathrm{d}W_t)^2 = (\lambda_t \mathrm{d}t + \sigma_t \mathrm{d}B_t)^2$$

  hence $\mathrm{d}t = \sigma_t^2 \mathrm{d}t$, so $\sigma_t = \pm 1$. Define $\tilde{W}$ by $\mathrm{d}\tilde{W}_t = \sigma_t \mathrm{d}B_t$.

- Then W can be shown to be $\mathbb{P}$-BM. And, as claimed,

$$\mathrm{d}W_t = \lambda_t \mathrm{d}t + \mathrm{d}\tilde{W}_t.$$

Fundamental theorem in continuous time

Black-Scholes

Black-Scholes via martingale approach

Black-Scholes dynamics

$$\mathrm{d}B_t = rB_t\mathrm{d}t \qquad\qquad B_0 = 1$$

$$\mathrm{d}S_t = \mu S_t\mathrm{d}t + \sigma S_t\mathrm{d}W_t \qquad\qquad S_0 > 0$$

where W is BM under physical measure P, and $\sigma > 0$.

No arb implies that $\exists\ \mathbb{P}$, equivalent to P, such that S/B is a $\mathbb{P}$-MG.

Hence by Girsanov, $\exists\ \lambda$ such that $\tilde{W}_t := W_t + \int_0^t \lambda_s \mathrm{d}s$ is $\mathbb{P}$-BM.

Substitute $\mathrm{d}\tilde{W}_t = \mathrm{d}W_t + \lambda_t\mathrm{d}t$ into the SDE of S:

$$\begin{aligned}\mathrm{d}S_t &= \mu S_t\mathrm{d}t + \sigma S_t(\mathrm{d}\tilde{W}_t - \lambda_t\mathrm{d}t)\\ &= (\mu - \lambda_t\sigma)S_t\mathrm{d}t + \sigma S_t\mathrm{d}\tilde{W}_t\end{aligned}$$

But can we say anything about $\mu - \lambda_t\sigma$?

Under $\mathbb{P}$, every tradeable asset X has drift rX

This page does not assume that X is a GBM.

Assume only that X is an Itô process.

- Under $\mathbb{P}$, the discounted price X/B is a MG, hence has zero drift.
- By Itô's rule, X/B has dynamics

$$\begin{aligned} \mathrm{d}(X_t/B_t) = \mathrm{d}(e^{-rt}X_t) &= e^{-rt}\mathrm{d}X_t - re^{-rt}X_t\mathrm{d}t + \mathrm{d}(e^{-rt})\mathrm{d}X_t \\ &= e^{-rt}(\mathrm{d}X_t - rX_t\mathrm{d}t), \end{aligned}$$

  so $\mathrm{d}X_t - rX_t\mathrm{d}t$ has no drift term.
- Therefore the drift term of $\mathrm{d}X_t$ must be $rX_t\mathrm{d}t$.

Under $\mathbb{P}$, the GBM S is still GBM, but with drift r

- Applying this to S, we have $(\mu - \lambda_t \sigma) S_t = r S_t$, and

$$\boxed{\mathrm{d}S_t = r S_t \mathrm{d}t + \sigma S_t \mathrm{d}\tilde{W}_t}$$

  where $\tilde{W}$ is $\mathbb{P}$-BM.

  Rate of growth changes from μ to r. Volatility stays the same.

- By L4, therefore, under $\mathbb{P}$, conditional on $\mathcal{F}_t^W$

$$\boxed{\log S_T \sim \text{Normal}(\log S_t + (r - \sigma^2/2)(T - t),\ \sigma^2 (T - t))}$$

  Compare: under P, conditional on $\mathcal{F}_t^W$,

$$\log S_T \sim \text{Normal}(\log S_t + (\mu - \sigma^2/2)(T - t),\ \sigma^2 (T - t))$$

Lognormal distribution

Here's a more general calculation, allowing different rates for growth and discounting, on an underlying X, not necessarily a stock price.

- Let $t < T$. Let R_{grow} and r be constants.
- Assume that (conditional on the time-t information $\mathcal{F}_t$) the random variable X_T has lognormal $\mathbb{P}$-distribution

  $$\log X_T \sim \text{Normal}(\log X_t + (R_{grow} - \sigma^2/2)(T-t),\ \sigma^2(T-t))$$

  where $X_t > 0$, and $\sigma > 0$ is a constant.
- One way that this distribution could arise is from the dynamics

  $$\mathrm{d}X_t = R_{grow} X_t \mathrm{d}t + \sigma X_t \mathrm{d}W_t \qquad X_0 > 0$$

  where W is $\mathbb{P}$-BM.

Conclusion: the Black-Scholes call price formula

Then, letting $\mathbb{E}$ denote expectation wrt $\mathbb{P}$,

$$\boxed{e^{-r(T-t)}\mathbb{E}_t(X_T - K)^+ = C^{BS}(X_t, t, K, T, R_{grow}, r, \sigma)}$$

where the function C^{BS} is defined for $X > 0, K > 0, \sigma > 0, t < T$ by

$$C^{BS}(X, t, K, T, R_{grow}, r, \sigma) := e^{-r(T-t)}\left[FN(d_1) - KN(d_2)\right],$$

and

$$F := Xe^{R_{grow}(T-t)} = \mathbb{E}_t X_T$$

and

$$d_{1,2} := d_{+,-} := \frac{\log(F/K)}{\sigma\sqrt{T-t}} \pm \frac{\sigma\sqrt{T-t}}{2}.$$

Proof of formula: decompose $(X_T - K)^+$

$\mathbb{1}_A = \begin{cases} 1 & \text{if } A \\ 0 & \text{if not } A \end{cases}$

$\mathbb{E}\mathbb{1}_A = \mathbb{P}(A)$

Proof: Using the fact that $x^+ = x\mathbf{1}_{x>0}$ for all x,

$$\begin{aligned}
\mathbb{E}_t(X_T - K)^+ &= \mathbb{E}_t(X_T - K)\mathbf{1}_{X_T>K} \\
&= \mathbb{E}_t(X_T\mathbf{1}_{X_T>K}) - K\mathbb{E}_t\mathbf{1}_{X_T>K} \\
&= F_t\underbrace{\mathbb{E}_t((X_T/F_t)\mathbf{1}_{X_T>K})}_{\text{will see this is } N(d_1)} - K\underbrace{\mathbb{P}_t(X_T > K)}_{\text{will see this is } N(d_2)},
\end{aligned}$$

where

$$F_t := X_t e^{R_{grow}(T-t)} = \mathbb{E}_t X_T.$$

To proceed, first note that the event $X_T > K$ is equivalent to $\log(X_T/F_t) > \log(K/F_t)$, and that (conditional on time-t information)

$$\log(X_T/F_t) \sim \text{Normal}(-(\sigma^2/2)(T-t),\ \sigma^2(T-t)).$$

Proof of formula: Second term of decomposition

Now consider separately each term in the decomposition.

For the second term,

$$
\begin{aligned}
\mathbb{P}_t(X_T > K) &= \mathbb{P}_t\Big(\log(X_T/F_t) > \log(K/F_t) \Big) \\
&= \mathbb{P}_t\left(\frac{\log(X_T/F_t) + (\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \right. \\
&\qquad\quad \left. > \frac{\log(K/F_t) + (\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \right) \\
&= N\left(- \frac{\log(K/F_t) + (\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \right) \\
&= N(d_2)\Big|_{X=X_t}
\end{aligned}
$$

Proof of formula: First term of decomposition

In the first term, we have $\mathbb{E}_t((X_T/F_t)\mathbf{1}_{X_T>K})$

$$
\begin{aligned}
&= \mathbb{E}_t[e^{\log(X_T/F_t)}\mathbf{1}_{\log(X_T/F_t)>\log(K/F_t)}] \\
&= \int_{\log(K/F_t)}^{\infty} e^y \frac{1}{\sqrt{2\pi\sigma^2(T-t)}} \exp\left(-\frac{(y+(\sigma^2/2)(T-t))^2}{2\sigma^2(T-t)}\right)\mathrm{d}y \\
&= \int_{\log(K/F_t)}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2(T-t)}} \exp\left(-\frac{(y-(\sigma^2/2)(T-t))^2}{2\sigma^2(T-t)}\right)\mathrm{d}y \\
&= \int_{\frac{\log(K/F_t)-(\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)\mathrm{d}z \\
&= N\left(-\frac{\log(K/F_t)-(\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}\right) = N(d_1)\Big|_{X=X_t},
\end{aligned}
$$

as claimed. The substitution was $z = \frac{y-(\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$ so $\mathrm{d}z = \frac{1}{\sigma\sqrt{T-t}}\mathrm{d}y$

Probabilistic interpretation of $N(d_1)$ and $N(d_2)$

In summary,

$$C_t = e^{-r(T-t)}\left(F_t \mathbb{E}_t\left[\frac{X_T}{F_t}\mathbf{1}_{S_T>K}\right] - K\mathbb{P}_t(S_T > K)\right)$$

$$= e^{-r(T-t)}\left(F_t N(d_1) - K N(d_2)\right)$$

The $N(d_2)$ is the $\mathbb{P}_t$ that the call option expires in the money.

(So $e^{-r(T-t)}N(d_2) =$ time-t price of K-strike T-expiry binary call.)

The $N(d_1)$ is the "share measure" $\mathbb{P}_t^S$ that the call expires ITM.

And time-t price of asset-or-nothing call paying $X_T\mathbf{1}_{X_T>K}$ is $e^{-r(T-t)}F_tN(d_1)$. If $X = S$ is a no-dividend stock, this is $S_tN(d_1)$

Probabilistic analysis of effect of r

Increasing r, while keeping everything else fixed, has what effect on time-0 call prices?

Martingale methods make the answer clear:

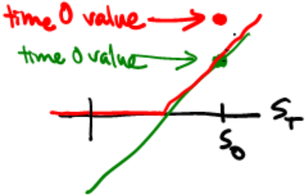
$$
\begin{aligned}
C_0 &= e^{-rT}\mathbb{E}C_T \\
&= e^{-rT}\mathbb{E}(S_T - K)^+ \\
&= e^{-rT}\mathbb{E}(S_0 e^{(r-\sigma^2/2)T+\sigma\tilde{W}_T} - K)^+ \\
&= \mathbb{E}(S_0 e^{(-\sigma^2/2)T+\sigma\tilde{W}_T} - Ke^{-rT})^+
\end{aligned}
$$

If $r\uparrow$ then $C_0\uparrow$

Probabilistic intuition about impact of σ

The *vega*, at time-t, of an asset or portfolio with value $C_t = C(S_t, t; \sigma)$ is $\frac{\partial C}{\partial \sigma}(S_t, t; \sigma)$. For a call or put in the B-S model,

$$\text{vega} := \frac{\partial C^{BS}}{\partial \sigma} = S\sqrt{T-t}N'(d_1) > 0$$

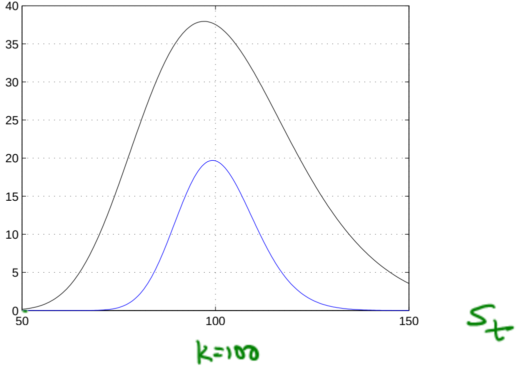


Why positive? Let's take $r = 0$.

- For a linear payout $a + bS_T$, the time-0 value is $a + bS_0$ regardless of σ. So vega for a linear payout is zero.
- For a Convex payout $f(S_T)$, such as a call, the payout dominates the linear tangent to f at S_0. So the contract's time-0 value is bigger than $f(S_0)$. By how much? Depends on σ. The larger the σ, the larger the Convex contract's time-0 value, because the larger the chance that S goes to where $f >$ linear. Large $\sigma \Rightarrow$ more Convexity is reachable.

Vega of a call

Vega of a call, plotted against S_t, for $T - t = 1$, and $T - t = 0.25$.



Under B-S dynamics, vega of a call is positive.

Probabilistic analysis of joint effect of σ and T

Halving σ and quadrupling T, while keeping everything else fixed, has what effect on time-0 call prices?

$\log \frac{S_T}{S_0} \sim \text{Normal}\left(\left(r - \frac{\sigma^2}{2}\right)T,\ \sigma^2 T \right)$

$\sigma \to \frac{\sigma}{2}$

$T \to 4T$

$\left(r - \frac{(\sigma/2)^2}{2}\right)4T$

$= \left(4r - \frac{\sigma^2}{2}\right)T$

$\left(\frac{\sigma}{2}\right)^2 (4T) = \sigma^2 T$

Also $e^{-rT} \to e^{-4rT}$

$\boxed{\text{Equivalent to } r \to 4r. \quad C_0 \uparrow \text{ if } r > 0}$

Appendix

On an interval I a function $f : I \to \mathbb{R}$ is said to be *convex* if its graph lies on or below all of its chords: for all $x, y \in I$, all $\alpha \in [0, 1]$,

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

Properties:

- A twice differentiable function f is convex iff $f'' \geq 0$ everywhere.
- A convex function's graph lies on or above all of its tangents.
- 

  Jensen's inequality:

  If f is convex on I and X is an integrable random variable taking values in I then $\mathbb{E}f(X) \geq f(\mathbb{E}X)$.

  Johan Jensen (1859-1925)