Stochastic processes

- These are stochastic variables which evolve over time.
- Some of you may know about these from
  - time series econometrics,
  - other applications in microeconomics or macroeconomics.
- Purpose here: Analyze prices of stocks and options.
- Binomial tree example of stochastic process in discrete time.
- "Discrete time:" Process only defined at certain time points.
- Black-Scholes-Merton option values based on another process.
- In continuous time, i.e., stock values S_t change continuously.
- (Although we typically observe only at some points in time.)
- Also continuous-valued, i.e., S_t can be any positive number.
- (In typical markets, S_t only has two or three decimals.)
- Could just define that process directly.
- Will instead follow Hull, ch. 12.
- First some rather simple, motivating points.
- Will then develop motivation for more complications.

The Markov property

- S_t called a *Markov* process if (the *Markov* property:) the probability distribution of all $S_{t+\Delta t}$ for all later dates $t+\Delta t$, as seen from date t, depends on S_t only.
- For instance, if S_t is a given number, knowledge of particularly high outcomes for S_{t-2} and S_{t-1}, or for $S_{t-0.2}$ and $S_{t-0.1}$, will not affect the probability distribution of $S_{t+0.1}$ or $S_{t+0.2}$ or ....
- Alternatively, we could think that the probability distribution of $S_{t+\Delta t}$ could depend on the whole history of S's, or some part of it, say S_{t-2}, S_{t-1}, S_t. Not Markov.
- One possible type of dependence, called momentum, is that a falling sequence $S_{t-2} > S_{t-1} > S_t$ increases the probability of an outcome S_{t+1} less than S_t. This is not Markov. For a Markov process, a rising sequence $S_{t-2} < S_{t-1} < S_t$ will, if it has the same value for S_t, imply exactly the same probability distribution for S_{t+1} as the falling sequence $S_{t-2} > S_{t-1} > S_t$.
- Exist many types of processes are Markov process, with many different types of probability distributions for, e.g., S_{t+1} conditional on S_t.
- "Markov processes" should thus be viewed as a wide class of stochastic processes, with one particular common characteristic, the Markov property.
- Remark on Hull, p. 259: "present value" in the first line of section 12.1 means "current value," "today's value".

The Markov property, economic implications

- Connection to weak-form market efficiency.
- All available information reflected in today's S_t.
- Probabilities of future $S_{t+\Delta t}$ depend on S_t.
- But historical S values cannot matter.
- Implication of $S_{t-\Delta t}$ for $S_{t+\Delta t}$? Already in S_t.

Implications of Markov property for variance

- Markov: $S_2 - S_1$ is stochastically independent of $S_1 - S_0$.
- Also $S_3 - S_2$, etc.
- Assume we are at time 0, know S_0
- Can write $S_2 = S_0 + (S_1 - S_0) + (S_2 - S_1)$.
- As seen from time 0, S_0 has no variance.
- Then $\text{var}(S_2) = \text{var}[(S_2 - S_1) + (S_1 - S_0)] = \text{var}(S_2 - S_1) + \text{var}(S_1 - S_0)$.
- The last equality is due to stochastic independence.
- Assume all changes $S_{t+1} - S_t$ have same variance.
- Then $\text{var}(S_2) = \text{var}(S_2 - S_1) + \text{var}(S_1 - S_0) = 2\,\text{var}(S_{t+1} - S_t)$.
- More precisely, introduce conditional variance, given S_0.
- $\text{var}(S_2|S_0) = 2\,\text{var}(S_{t+1} - S_t)$.
- Likewise: $\text{var}(S_3|S_0) = 3\,\text{var}(S_{t+1} - S_t)$.
- Generally: $\text{var}(S_T|S_0) = T\,\text{var}(S_{t+1} - S_t)$.
- (Conditional) variance proportional to time.
- Standard deviation proportional to square root of time.

Wiener processes (also called Brownian motion)

- So far, in addition to the Markov property, have assumed the variance of changes is the same for different periods.
- Assume now that $\text{var}(S_{t+1} - S_t | S_t)$ equals 1, and that the expected change $E(S_{t+1} - S_t | S_t)$ equals 0.
- (A bit like looking at a standardized distribution, like $N(0, 1)$. Will call this process z_t (or sometimes $z(t)$), not S_t.)
- This gives us a particular type of Markov process called a *Wiener* process, defined by two properties. z_t is a Wiener process if and only if both are satisfied:
    - The change Δz during a short time interval Δt is $\Delta z = \epsilon\sqrt{\Delta t}$, where ϵ has a standard normal (Gaussian) distribution (with $E(\epsilon) = 0, \ \ \text{var}(\epsilon) = 1$).
    - The values of Δz for non-overlapping intervals Δt are stochastically independent.
- Over a longer interval, the change $z(T) - z(0)$ is normally distributed, the sum of N changes over intervals of length Δt, i.e., $N\Delta t = T$, and $z(T) - z(0) = \Sigma_{i=1}^{N} \epsilon_i \sqrt{\Delta t}$.
- This implies $E(z(T) - z(0)) = 0$, $\text{var}(z(T) - z(0)) = N\Delta t = T$, and the standard deviation of $z(T) - z(0)$ is $\sqrt{T}$. These do not depend on the length of Δt, which is reassuring, since it was not defined in any precise way.
- In limit when $\Delta t \to 0$, dz is change during dt; $\text{var}(dz) = dt$.
- Illustrated in Figure 12.1 in Hull, p. 262.

Generalized Wiener processes; Itô processes

- First multiply the Wiener process dz by a constant, b.
- $b\ dz$ has variance $b^2 \operatorname{var}(dz) = b^2 dt$.
- Then allow for an expected change different from zero,

$$dx = a\ dt + b\ dz$$

- This amounts to adding a non-stochastic linear growth path to the stochastic $b\ dz$, and is illustrated in Figure 12.2 in Hull, p. 264.
- The *generalized Wiener process* X is normally distributed with

$$E(X(T) - X(0)) = aT,$$
$$\operatorname{var}(X(T) - X(0)) = b^2 T.$$

- The process is also called *Brownian motion with drift.*
- A further generalization: Allow a and b to depend on (x, t),

$$dx = a(x,t)dt + b(x,t)dz.$$

- This is called an *Itô process.* In general not normally distributed.
- Over a small time interval Δt we get

$$\Delta x = a(x,t)\Delta t + b(x,t)\epsilon\sqrt{\Delta t}.$$

- For non-overlapping intervals the changes in x are stochastically independent, so all Itô processes are Markov processes.

Stochastic process for a stock price

- Looking for something more realistic than the binomial tree.
- Expected change will not be zero, so cannot use Wiener process.
- Could we use generalized Wiener process?
- Expected change over interval of length T is aT.
- Suppose $S_0 = 10$, $a = 0.1$, and that T is measured in years.
- Expected stock price in ten years is $E(S_{10}|S_0 = 10) = 20$.
- Expected stock price ten years later, $E(S_{20}|S_0 = 10) = 30$.
- Also, if S_{10} equals its expectation, $E(S_{20}|S_{10} = 20) = 30$.
- But the expected growth rate over the time interval $(10, 20)$ is substantially lower than the expected growth rate over $(0, 10)$, since growth rates are relative numbers, and $30/20 < 20/10$.
- More likely shareholders require constant expected growth rate.
- Need exponential expected path, not linear expected path.
- Will obtain this by letting $E(dS) = \mu\ S\ dt$.
- For the non-stochastic part (or, if $\sigma = 0$): $\frac{dS}{dt} = \mu\ S$.
- Integrating between 0 and T: $S_T = S_0 e^{\mu T}$ when $\sigma = 0$.
- This leads to a suggestion of

$$dS = \mu\ S\ dt + \sigma\ dz$$

  or, better,

$$dS = \mu\ S\ dt + \sigma\ S\ dz.$$

Stochastic process for stock price, contd.

- From previous page: a suggestion of

$$dS = \mu\, S\, dt + \sigma\, dz$$

or

$$dS = \mu\, S\, dt + \sigma\, S\, dz.$$

- Choose the latter so that a relative change in S not only has a constant expected value, $\mu\, dt$, but also a constant variance, $\sigma^2 dt$,

$$\frac{dS}{S} = \mu\, dt + \sigma\, dz.$$

- This stock price process process is basis for the most widespread option pricing theories, like the one in Chapter 13 of Hull, Black-Scholes-Merton.
- The process is called *geometric Brownian motion with drift.*
- Since S appears on right-hand side in dS formula: Not a generalized Wiener process, but a bit more complicated.
- dS is an Itô process, with $a(S,t) = \mu S$ and $b(S,t) = \sigma S$.
- Different stocks will differ in μ and/or σ.
- Hull discusses these variables in section 12.4.
- Remember: Hull's book does not rely on the CAPM.
- Imprecise discussion of how μ depends on r_f and risk.
- Footnote 4 means μ depends on covariance, not on σ.

Functions of Itô processes

- When x is an Itô process, $dx = a(x,t)dt + b(x,t)dz$:
  - Is a function G of x also an Itô process?
  - If yes, what happens to the functions $a(x,t)$ and $b(x,t)$?
  - Put differently: G will also have functions like these.
  - What do the two functions look like for G?
- Motivation: Call option value as function of S.
- Find this via a general rule, Itô's lemma.
- A bit more complicated than suggested above.
- Call option not only function of S; also of t.
- Option's value depends on time until expiration.
- For some given S, different t's give different c's.
- Thus, the more general questions are:
  - If x is an Itô process, is $G(x,t)$ an Itô process?
  - If yes, what do the "a and b functions" look like for G?
- The answers are given by *Itô's lemma.*
- Will not prove this mathematically.
- But will show how and why it differs from usual differentiation.

Itô's lemma

- Assume x is an Itô process:
- $dx = a(x,t)dt + b(x,t)dz$, where z is a Wiener process.
- Then $G(x,t)$ is also an Itô process:

$$dG = \left(\frac{\partial G}{\partial x}a + \frac{\partial G}{\partial t} + \frac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2\right)dt + \frac{\partial G}{\partial x}b\;dz.$$

- We recognize the general form of an Itô process.
- The expression above is Hull's equation (12.12).
- In fact, this is short-hand, dropping arguments.
- Contains six different functions of (x,t).
- Both a, b, G, and the partial derivatives of G.
- Right-hand side should really be written like this:

$$\left(\frac{\partial G(x,t)}{\partial x}a(x,t) + \frac{\partial G(x,t)}{\partial t} + \frac{1}{2}\frac{\partial^2 G(x,t)}{\partial x^2}[b(x,t)]^2\right)dt + \frac{\partial G(x,t)}{\partial x}b(x,t)dz.$$

- Perhaps this looks complicated, but:
- In our applications, G, a, and b are fairly simple.

Why not use ordinary differentiation? Hull, p. 275f

- Approximation of a function by its tangent:

$$\Delta G \approx \frac{dG}{dx}\Delta x$$

  when G is a function of one variable, x.

- Holds precisely in limit as $\Delta x \to 0$.
- As long as $\Delta x \neq 0$, can use Taylor series expansion:

$$\Delta G = \frac{dG}{dx}\Delta x + \frac{1}{2}\frac{d^2G}{dx^2}\Delta x^2 + \frac{1}{6}\frac{d^3G}{dx^3}\Delta x^3 + \ldots$$

- As $\Delta x \to 0$, higher-order terms vanish.
- $G(x, y)$, two dimensions, a tangent plane:

$$\Delta G \approx \frac{\partial G}{\partial x}\Delta x + \frac{\partial G}{\partial y}\Delta y.$$

- When both Δx and $\Delta y \neq 0$, can use Taylor series:

$$\Delta G = \frac{\partial G}{\partial x}\Delta x + \frac{\partial G}{\partial y}\Delta y + \frac{1}{2}\frac{\partial^2 G}{\partial x^2}\Delta x^2 + \frac{\partial^2 G}{\partial x \partial y}\Delta x\ \Delta y + \frac{1}{2}\frac{\partial^2 G}{\partial y^2}\Delta y^2 + \ldots$$

- Again, precisely in limit as $\Delta x \to 0$ and $\Delta y \to 0$:

$$dG = \frac{\partial G}{\partial x}dx + \frac{\partial G}{\partial y}dy.$$

- Want to find a similar expression for Itô processes.
- But all higher-order terms do not vanish.

Itô's lemma vs. ordinary differentiation

- Assume x is an Itô process:
- $dx = a(x,t)dt + b(x,t)dz$, where z is a Wiener process.
- Let G be a function $G(x,t)$, and use Taylor expansion:
  $$\Delta G = \frac{\partial G}{\partial x}\Delta x + \frac{\partial G}{\partial t}\Delta t + \frac{1}{2}\frac{\partial^2 G}{\partial x^2}\Delta x^2 + \frac{\partial^2 G}{\partial x \partial t}\Delta x\ \Delta t + \frac{1}{2}\frac{\partial^2 G}{\partial t^2}\Delta t^2 + \dots$$
- Only novelty here: Have called second variable t, not y.
- When $\Delta x \to 0$, need to observe the following.
- $\Delta x = a\ \Delta t + b\epsilon\sqrt{\Delta t}$ implies:
- $(\Delta x)^2 = b^2\epsilon^2\ \Delta t$ + terms of higher order.
- Since Δx contains a $\sqrt{\Delta t}$ term, normal rules don't work.
- Must include extra term with second-order partial derivative.
- The extra term contains ϵ^2, and ϵ is stochastic.
- Hull explains why $E(\epsilon^2 \Delta t) = \Delta t$.
- Hull also explains that $\text{var}(\epsilon^2 \Delta t)$ is of order $(\Delta t)^2$.
- Variance approaches zero fast as $\Delta t \to 0$.
- Thus: In limit $\epsilon^2 \Delta t$ is nonstochastic, $= \Delta t$.
- This gives us the following formula in the limit:
  $$dG = \frac{\partial G}{\partial x}dx + \frac{\partial G}{\partial t}dt + \frac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2 dt$$
- Insert for dx from above to find the form we used above:
  $$dG = \left(\frac{\partial G}{\partial x}a + \frac{\partial G}{\partial t} + \frac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2\right)dt + \frac{\partial G}{\partial x}b\ dz.$$

Example of application of Itô's lemma

- Consider the stock price process from p. 8 above:
- Assume $dS = \mu S\ dt + \sigma S\ dz$; z is a Wiener process.
- What kind of process is $\ln S$?
- Natural question; deterministic part of S is exponential in t.
- Might believe that deterministic part of $\ln S$ is linear in t.
- Observe this application of Itô's lemma is fairly simple:
  - "$a(S, t)$ function" of S process is μS. Simple, and no t.
  - "$b(S, t)$ function" of S process is σS. Simple, and no t.
  - The $G(S, t)$ function is $\ln S$. Fairly simple, and no t.
- Know from Itô's lemma that $\ln S$ is an Itô process.
- But what are the "a and b functions" of the G process?
- Will turn out that they are very simple. Constants, no S, no t.
- But slightly less simple than one might have thought.
- The constant which multiplies dt is not μ.
- Would be natural suggestion based on deterministic $S_T = S_0 e^{\mu T}$.

Example, contd.; lognormal property, Hull, sect. 12.6

- With $G(S,t) \equiv \ln S$, need three partial derivatives:

$$\frac{\partial G}{\partial S} = \frac{1}{S}, \quad \frac{\partial^2 G}{\partial S^2} = -\frac{1}{S^2}, \quad \frac{\partial G}{\partial t} = 0.$$

- Then Itô's lemma says that:

$$dG = \left(\frac{1}{S}\mu S + 0 + \frac{1}{2}\left(-\frac{1}{S^2}\right)(\sigma S)^2\right) dt + \frac{1}{S}\sigma S\ dz$$
$$= \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma\ dz.$$

- Implies that $\ln S$ is a generalized Wiener process.
- Can use formulae from p. 6 above.
- The change $\ln S_T - \ln S_0$ is normally distributed:

$$\ln S_T - \ln S_0 \sim \phi\left[\left(\mu - \frac{\sigma^2}{2}\right)T, \sigma^2 T\right],$$

  which implies (by adding the known $\ln S_0$)

$$\ln S_T \sim \phi\left[\ln S_0 + \left(\mu - \frac{\sigma^2}{2}\right)T, \sigma^2 T\right].$$

- $\ln S$ is normally distributed.
- By definition then, S is lognormally distributed.
- Not transparent earlier, but by using Itô's lemma.

The lognormal distribution of stock prices

- On p. 7, required an exponential expected path, $S_T = S_0 e^{\mu T}$.
- Could thus not use the generalized Wiener process for S.
- (Would have implied S having a normal distribution.)
- Found instead something similar for relative changes in S,

$$\frac{dS}{S} = \mu\, dt + \sigma\, dz.$$

- This implies S is lognormal, $\ln(S)$ is normal.
- Relation between these two distributions may be confusing.
- Remember that $\ln(S)$ is not linear, thus $E[\ln(S)] \neq \ln[E(S)]$:
  - $E[\ln(S_T)|S_0] = \ln(S_0) + (\mu - \sigma^2/2)T$,
  - $E(S_T|S_0) = S_0 e^{\mu T}$ so that $\ln[E(S_T|S_0)] = \ln(S_0) + \mu T$.
- The variance expression is simpler for $\ln(S_T)$ than for S_T:
  - $\text{var}[\ln(S_T)|S_0] = \sigma^2 T$,
  - $\text{var}(S_T|S_0) = S_0^2 e^{2\mu T}(e^{\sigma^2 T} - 1)$.
- Footnote 2 on p. 279 in Hull refers to a note on this:

`http://www.rotman.utoronto.ca/~hull/TechnicalNotes/TechnicalNote2.pdf`

- $S_T = S_0 e^{xT}$ defines continuously-compounded rate of return x.
- Its distribution is $x \sim \phi\left(\mu - \frac{\sigma^2}{2}, \frac{\sigma^2}{T}\right)$.