STOCHASTIC PROCESSES, ITÔ CALCULUS, AND APPLICATIONS IN ECONOMICS

Timothy P. Hubbard & Yigit Saglam[†]
Department of Economics
University of Iowa

March 3, 2006

Abstract

This document provides an introduction to stochastic processes and Itô calculus with emphasis on what an economist needs to understand to do research on optimal control and dynamic programming problems involving randomness. Specifically we show why learning this material is important by illustration with an example in the first section. We then discuss stochastic processes and, in particular, Wiener processes and Brownian motion. In the third section we introduce Itô calculus, a tool for dealing with stochastic integrals and stochastic differential equations. In the last section of the manuscript we discuss some applications from the field of natural resource economics.

[†] Department of Economics, University of Iowa, W210 Pappajohn Business Building, Iowa City, IA 52242, timothy-hubbard@uiowa.edu, yigit-saglam@uiowa.edu

1. Introduction

Economists often want to model problems in a way that incorporates the notion that today's choices impact future decisions (there exists a trend component) but also allow for the fact that there will be shocks that will alter your state in future periods (there is a random component). Take, for example, a stochastic growth model as in Stokey and Lucas with Prescott (1989). In their example, output y, which is $f(k)z$ in a stochastic one-period growth model, is determined by the size of the current capital stock, k, and a stochastic technology shock, z. Assuming z is distributed independently and identically over time, the objective can be expressed in the following Bellman equation

$$v(k,z) = \max_{k' \in [0, f(k)z]} \left\{ u\left[f(k)z - k'\right] + \beta \mathcal{E}\left[v(k', z')\right] \right\}. \tag{1.1}$$

The Bellman equation shows that production depends on the shock, z, and hence so does the agents utility today. Notice also, that unlike in deterministic dynamic programming we now have discounted expected utility entering into our optimization problem. What does this new $\mathcal{E}$ operator mean? Suppose that $z \in \{z_1, ..., z_N\}$ are the possible values that the shock z could take and that $\mathsf{Prob}(z_i)$ is π_i where

$$\sum_{i=1}^{N} \pi_i = 1.$$

Then (1.1) can be written as

$$v(k,z) = \max_{k' \in [0, f(k)z]} \left\{ u\left[f(k)z - k'\right] + \beta \sum_{i=1}^{N} v(k', z_i)\pi_i \right\}.$$

In general we want to construct problems where we can allow π_i to depend on today's shock z such that if $z \in [a, b]$ and $[c, d] \subseteq [a, b]$,

$$\pi(z) = \mathsf{Prob}\left\{z \in [c, d]\right\} = \int_c^d \pi(z) dz.$$

In this case we now rewrite (1.1) as

$$v(k,z) = \max_{k' \in [0, f(k)z]} \left\{ u\left[f(k)z - k'\right] + \beta \int_a^b v(k', z')\pi(z') dz' \right\}. \tag{1.2}$$

To understand what this equation means we first need to learn some concepts from measure theory and stochastic calculus. This document serves as a basic introduction to these ideas.

2. Stochastic Processes

A stochastic process can be thought of as a system that evolves over time in a random manner. Mathematically this means that a stochastic process maps an element from a probability space into a state space. If the time parameter t can take on any value in $\mathcal{R}_+$, then the process is a continuous-time stochastic process, whereas if $t \in \mathcal{Z}_+$, where $\mathcal{Z}_+$ is the set of nonnegative integers, then the process is a discrete-time process. Some specific examples of stochastic processes include a Wiener process (explained below), a random walk, or a Markov process (a Markov chain is the corresponding discrete-time process). A Wiener process, $W(\cdot)$, is a continuous-time stochastic process defined to have the following properties:

1. $W(t) \sim N(0, t)$ for all $t \geq 0$,
2. if $0 \leq s < t$ then $[W(t) - W(s)] \sim N(0, t - s)$, which results because $W(\cdot)$ is a Gaussian stochastic process,
3. if $[s, t)$ and $[u, v)$ do not overlap then $[W(t) - W(s)]$ and $[W(v) - W(u)]$ are independent random variables.

2.1. Brownian Motion

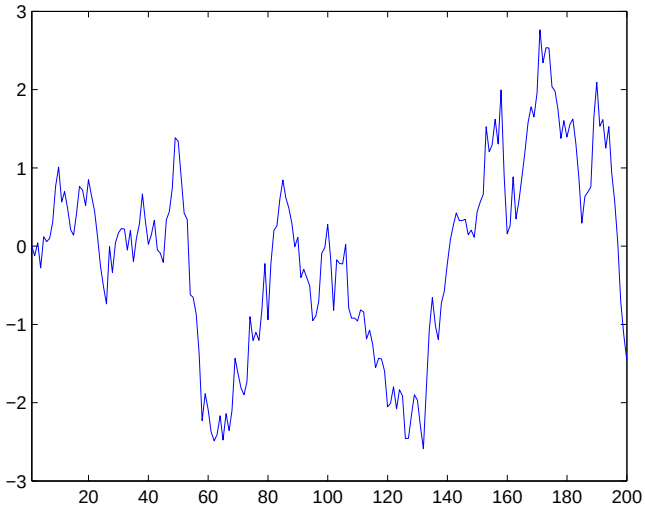
Brownian motion is a Wiener process in which a quantity of interest is constantly undergoing small, random fluctuations. The differential form of Brownian motion can be written

$$dS(t) = \mu dt + \sigma dW(t) \tag{2.1}$$

where $W(t)$ is a Wiener process, μdt is $\mathcal{E}\left[dS(t)\right]$ and $\sigma^2 dt$ is $\text{var}\left[dS(t)\right]$. The process was named for Scottish botanist Robert Brown who first studied such fluctuations in observing pollen grains suspended in water. Given an initial starting point one notices that a particle drifts over time and it is possible to compute the probability of a particle moving a specific distance in any direction during a certain time interval in a specific medium; *i.e.*, the probability a particle will move more than a specified distance (think survivor function). Albert Einstein later concluded that a smaller particle, less viscous fluid, and higher temperature all increased the amount of motion one would expect to observe. Let $S(t)$ denote the position of the particle at time t, then the position of the particle at time $t + dt$ is distributed as follows:

$$[S(t + dt)|S(t) = g] \sim N(g + \mu dt, \sigma^2 dt), \tag{2.2}$$

Figure 1



where μ is the mean of the increments (drift) and σ^2 the variance of the increments (this results from (2.1) and is shown later after we introduce stochastic integration). This formula specifies the probability density function for how far the particle moves. In Figure 1 we have traced the realized path for a quantity over 200 time periods given it follows a Brownian motion and assuming the initial quantity is zero. We note without proof that, as it is clear in the figure, Brownian motion is almost surely continuous and almost surely nowhere differentiable in t. Another important property is that Brownian motion is a Markovian process, that is the probability of being in state $S(t)$ at time t, given all states up to time $t-1$, depends only on the previous state, $S(t-1)$, at time $t-1$. Said differently, the prediction of what states will occur in the future depends only on the current state.

Stokey (1998) discusses a way to numerically approximate a Brownian motion to within any specified degree of accuracy using a discrete-time random walk process as an approximation. The relationship between a random walk process and Brownian motion is fairly straightforward. Consider a process S that, for each time increment Δ, increases by h with probability p and decreases by h with probability $(1-p)$ (this process can be modeled as a binomial tree). Then process S has the following properties:

$$\mathcal{E}\left[S\left(t+\Delta\right)-S(t)\right]=ph-(1-p)h=h(2p-1)$$

and

$$\mathcal{E}\left[S\left(t+\Delta\right)-S(t)\right]^2 = ph^2 + (1-p)h^2 = h^2.$$

Define $Q(N)$ to be $\sum_{i=1}^{N} S(i)$ which represents the position of a process that follows a random walk after N steps, where we assume p equals $\frac{1}{2}$. If the N steps take T periods then Δ is defined to be T/N. Since we assumed p is $\frac{1}{2}$, this implies $\mathcal{E}\left[S(t)\right]$ is zero and $\text{var}\left[S(t)\right]$ is one. Consequently $\mathcal{E}\left[Q(N)\right]$ is zero and $\text{var}\left[Q(N)\right]$ is N, by independence. By the Central Limit Theorem, as $N \to \infty$, $Q(N)/\sqrt{N} \sim^{D} N(0,1)$. Thus if we scale each $Q(t)$ by $\sqrt{N}$ we obtain the sequence of a random process that converges to Brownian motion as $N \to \infty$. It can also be shown that one can obtain a Brownian motion with parameters (μ, σ) by setting $\mu\Delta$ equal to $h(2p-1)$ and $\sigma^2\Delta$ equal to h^2.

2.2. Geometric Brownian Motion

Often authors assume geometric Brownian motion (GBM) which has the property that the logarithm of a randomly varying quantity follows a Brownian motion. This assumption is frequently employed when the randomly varying quantity can only take on values that are positive. A stochastic process $S(t)$ follows a GBM if it satisfies

$$dS(t) = uS(t)dt + vS(t)dW(t) \tag{2.3}$$

where $W(t)$ is a Wiener process, u is the expected change in $S(t)$, and v is the variance of the change. Note that $\log\left[S(t)/S(0)\right] \sim N\left[(u - \frac{1}{2}v^2)t, v^2t\right]$; that is, increments of a GBM are normal relative to the values of the variable in the current state. As an example consider a model for stock price behavior, then (2.3) can be interpreted as the Black-Scholes formula where $S(t)$ is the stock price at time t, u is the instantaneous rate of return on a riskless asset, and v is the volatility of the stock. As (2.3) hints at, we are often more interested in the process $dS(t)$ rather than the process $S(t)$, but studying it requires the use of Itô calculus, which we discuss next.

3. Itô Calculus

In this section we use concepts from stochastic processes to introduce the idea of a stochastic integral and illustrate some important properties.

3.1. Solving an Itô Integral

To begin talking about how to interpret and solve (2.3) we need to find a method for dealing with stochastic differential equations. Our approach is similar to the pedagogical method one would use to teach Riemann integrals in that we use a limiting approach to help illustrate the intuition of stochastic calculus and Itô integrals. We begin by considering simple functions, which is a step closer to stochastic integrals in the sense that they are like the rectangular step functions one uses to approximate the area under a function, with the exception that simple functions are like random step functions. Knowing a little measure theory will help in understanding this analogy but is not required.

Let $\mathbf{1}_{A_i}(s)$ be an indicator function that takes on a value of one if $s \in A_i$ and zero otherwise, where $A_i \subseteq A$. A simple function, $\phi(s)$ can be defined as

$$\phi(s) = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}(s) \tag{3.1}$$

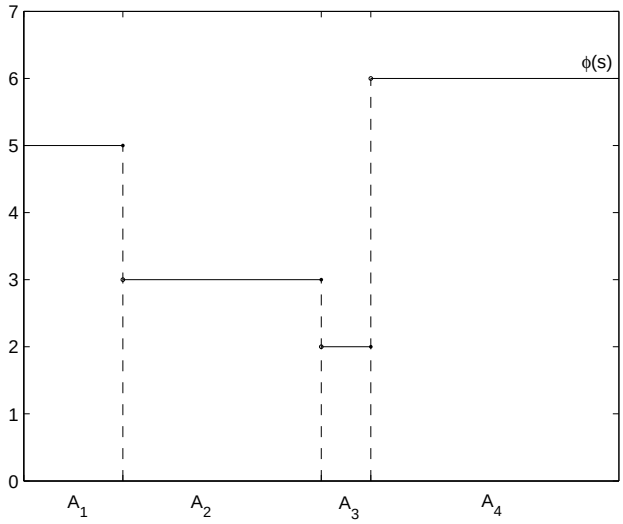
where a_i is the value the function attains if $s \in A_i$. Now consider a simple function $\phi(s)$ and a measure space $(A, \mathcal{F}, \mu)$ where A is a set, $\mathcal{F}$ is a $\sigma-$algebra, and μ is a measure mapping $\mathcal{F}$ into R. Then the integral of $\phi(\cdot)$ with respect to measure μ is

$$\int_s \phi(s)\mu(ds) = \sum_{i=1}^{n} a_i \mu(A_i). \tag{3.2}$$

The interpretation is that in computing the stochastic integral we want to weight things that are more likely to happen. This integral is like a typical Riemann integral in the sense that we are still adding together the area of rectangles, except now the idea is that we weight the regions by how likely they are according to measure μ. The following example will help illustrate this idea. Consider the set $A = \mathcal{R}_+$ with subsets $\{A_1, A_2, A_3, A_4\}$, such that $\bigcup_i A_i$ is $\mathcal{R}$ and $A_i \bigcap A_j$ is empty for all $i \neq j$, and let $\mu(A_1)$ equal .3, $\mu(A_2)$ equal .2, $\mu(A_3)$ equal .1, and $\mu(A_4)$ equal .4 (note that μ is a probability measure as $\sum_i \mu(A_i)$ is one). Also let a_1, the value that $\phi(s)$ attains for any element in A_1, be 5. Likewise let a_2, a_3, and a_4 take on the values 3, 2, and 6, respectively. For this example, the simple function $\phi(s)$ is shown in Figure 2. In this example,

$$\sum_{i=1}^{4} a_i \mu(A_i) = 5(.3) + 3(.2) + 2(.1) + 6(.4) = 4.7.$$

Figure 2



While (3.2) brings us one step closer (compared to a Riemann sum) to solving stochastic integrals we need some definitions and conditions that will allow us to solve stochastic integrals of the form

$$I_S(t) = \int_0^t S(t)dW(t),\ \forall\ t \geq 0 \tag{3.3}$$

where here, again suppressing arguments for convenience, I_S is the integral of S with respect to W. We maintain the assumption that the integral is bounded by imposing the following restriction

$$\mathcal{E}\left[\int_0^t S(x)^2 dx\right] < \infty,\ \forall\ t > 0. \tag{3.4}$$

To begin solving stochastic integrals as in (3.3) define a simple process to be a stochastic process that has a countable ordered sequence $\{t_k\}_{k=0}^K$ such that

$$S(r) = S(t_{k-1}),\ \forall\ r \in [t_{k-1}, t_k)\,. \tag{3.5}$$

To solve the stochastic integral of a simple process with steps $\{t_k\}_{k=0}^K$ we slightly adjust our integral for simple functions discussed above to account for the stochastic (Wiener) process W such that

$$\int_0^t S(t)dW(t) \equiv \sum_{i=0}^{n-1} S(i)\left[W(i+1) - W(i)\right] + S(t_n)\left[W(t) - W(t_n)\right] \tag{3.6}$$

where $0 < t_n < t$. Note that in this equation the state value at each point in time is multiplied by $[W(i+1) - W(i)] \sim N(0,1)$ (by definition of a Wiener process). Thus the state variable amplifies the random fluctuations. In Figure 2 we can think of $S(t)$ as the height of the simple function $\phi(s)$ and the $W(i)$'s as the endpoints of the subsets A_i.

To extend the concept of the Itô integral to a broader class of integrands we need to show that there exists a unique sequence of simple processes that approximate the function or, in the case where there is more than one sequence, that all approximating sequences converge to the same value. Thus we approximate a general integrand $\psi(t)$ by a sequence of simple processes and use the integral defined in (3.6). We state the following without proving existence and uniqueness of the sequence of simple processes as the proof is beyond the scope of this document and is not useful for applications.

> *Proposition:* If $\psi(t)$ satisfies (3.4) then $\psi(t)$ is integrable and there exists a stochastic process such that (3.3) is satisfied.

Thus our definition of an Itô integral for a simple process in (3.6) provides a way to approximate stochastic integrals for a general function so long as (3.4) is satisfied.

Given our definition of the Itô integral as the limit of a sequence of integrals for simple processes we can now define some properties. First note that

$$\mathcal{E}\left[\int_t S(t)dW(t)\right] = \mathcal{E}\left\{\sum_{i=0}^{n} S(i)\left[W(i+1) - W(i)\right]\right\} = 0$$

where the last equality results because the $S(i)$ are independent of $[W(i+1) - W(i)]$. This says that the expected value of a stochastic integral is zero! In terms of variance,

$$\text{var}\left[\int_t S(t)dW(t)\right] = \mathcal{E}\left[\int_t S(t)dW(t)\right]^2 = \int \mathcal{E}\left[S(t)^2\right] dt$$

where the last integral is a standard (Riemann) integral which results from the fact that dW^2 is dt (this is shown explicitly in the next subsection). Along the same lines if we consider two simple processes, $S(t)$ and $P(t)$ then

$$\begin{aligned}\text{cov}\left[\int_t S(t)dW(t)\int_t P(t)dW(t)\right] &= \mathcal{E}\left[\int_t S(t)dW(t)\int_t P(t)dW(t)\right] \\ &= \mathcal{E}\left[\int_t S(t)P(t)dt\right]\end{aligned}$$

where again the last integral is a Riemann integral and we have substituted dt for dW^2. Not surprisingly, if we again consider two simple processes and two constants, b and c, then

$$\int_t [aS(t) + bP(t)]\, dW(t) = a \int_t S(t)dW(t) + b \int_t P(t)dW(t)$$

which implies the stochastic integral of a weighted sum is the weighted sum of the stochastic integrals.

3.2. Itô's Lemma

An often used result in applications involving stochastic differential equations is Itô's lemma. One key to proving this lemma derives from property 2 of the definition for a Wiener process. Specifically, since $W(t) - W(s) \sim N(0, t-s)$ we can write ΔW as $\epsilon\sqrt{\Delta t}$, where $\epsilon \sim N(0,1)$. Letting Δt get infinitesimally small we write dW as $\epsilon\sqrt{dt}$, with the following properties

$$\mathcal{E}(dW) = \mathcal{E}\left(\epsilon\sqrt{dt}\right) = \sqrt{dt}\mathcal{E}(\epsilon) = 0$$

and,

$$\text{var}(dW) = \mathcal{E}(dW^2) - \mathcal{E}^2(dW) = \mathcal{E}\left(\epsilon^2 dt\right) = \text{var}(\epsilon)dt = dt.$$

Now assuming a function $f(S,t)$ is $\mathcal{C}^1$ in t and $\mathcal{C}^2$ in S we can approximate the total differential df (omitting arguments for shorthand) by a Taylor series expansion as follows

$$\begin{aligned} df &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial S}dS + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}dS^2 + \mathcal{O}\left(dS^2\right) \\ &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial S}\left(\mu dt + \sigma dW\right) \\ &\qquad + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\left[\mu^2 dt^2 + 2\mu\sigma dt dW + \sigma^2 dW^2\right] + \mathcal{O}\left(dS^2\right). \end{aligned}$$

If we drop terms of order higher than dt or dW^2 then this simplifies to

$$df = \frac{\partial f}{\partial t}dt + \mu\frac{\partial f}{\partial S}dt + \sigma\frac{\partial f}{\partial S}dW + \sigma^2\frac{1}{2}\frac{\partial^2 f}{\partial S^2}dW^2.$$

Using the fact (shown above) that $\mathcal{E}\left(dW^2\right)$ is dt, we obtain Itô's lemma

$$df = \left(\frac{\partial f}{\partial t} + \mu\frac{\partial f}{\partial S} + \sigma^2\frac{1}{2}\frac{\partial^2 f}{\partial S^2}\right)dt + \sigma\frac{\partial f}{\partial S}dW \tag{3.7}$$

which can be thought of as a chain rule for stochastic functions.

Consider, as an example which is commonly used, the case where $S(t)$ follows Brownian motion with parameters μ and σ as in (2.1) and let $R\left[S(t), t\right]$ be defined as $\exp\left[S(t)\right]$. Then we have the following

$$\frac{\partial R}{\partial t} = 0,$$
$$\frac{\partial R}{\partial S} = \exp(S) = R,$$

and

$$\frac{\partial^2 R}{\partial S^2} = \exp(S) = R.$$

Now applying Itô's lemma we obtain

$$dR = \left(\mu + \frac{1}{2}\sigma^2\right)Rdt + \sigma RdW \tag{3.8}$$

which shows that $R(t)$ follows GBM as in (2.3) where u is $\left(\mu + \frac{1}{2}\sigma^2\right)$ and v is σ. This is a particularly important example in economics as often we restrict parameters (like $R(t)$ here) to take on positive values.

3.3. Hamilton-Jacobi-Bellman Equation

Assume that $X(t)$ follows a Brownian motion with an initial value $X(0)$ of zero and parameters (μ, σ^2). Then

$$X(t) = X(0) + \mu t + \sigma W(t),\ \forall\ t,$$

where W is a Wiener process. This can be written in differential form as

$$dX(t) = \mu dt + \sigma dW(t),\ \forall\ t.$$

Now, assume that $F(t,x)$ is differentiable at least once in t and twice in x. Taking the total differential and using a Taylor series expansion yields

$$dF = \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial X}dX + \frac{1}{2}\frac{\partial^2 F}{\partial X^2}dX^2 + \dots$$

After some manipulation and after dropping higher order terms we obtain

$$dF = \frac{\partial F}{\partial t}dt + \mu\frac{\partial F}{\partial X}dt + \frac{1}{2}\sigma^2\frac{\partial^2 F}{\partial X^2}dX^2.$$

Since $\mathcal{E}(dW)$ is zero and $\mathcal{E}[(dW)^2]$ is dt after taking expectations, we find the mean and variance of dF

$$\mathcal{E}(dF) = \left[\frac{\partial F}{\partial t} + \mu\frac{\partial F}{\partial X} + \frac{1}{2}\sigma^2\frac{\partial^2 F}{\partial X^2}\right]dt$$

and

$$\text{var}(dF) = \mathcal{E}\left[dF - \mathcal{E}(dF)\right]^2 = \sigma^2\left(\frac{\partial F}{\partial X}\right)^2 dt.$$

Let $v(x_0)$ be the expected discounted value of a stream of returns, $\pi\left[X(t)\right]$, given the initial state $X(0)$ is x_0, which implies

$$v(x_0) = \mathcal{E}\left\{\int_0^\infty e^{-\rho t}\pi\left[X(t)\right]dt \mid X(0) = x_0\right\}. \tag{3.9}$$

Assuming $\pi(\cdot)$ is continuous and bounded, the Riemann integral in (3.9) is well defined. Now, for any small interval of time Δt, (3.9) has the following Bellman-type property

$$v(x_0) \approx \pi(x_0)\Delta t + \frac{1}{1+\rho\Delta t}\mathcal{E}\left\{v[X(0+\Delta t)]|X(0) = x_0\right\}.$$

After some manipulation, as $\Delta t \to 0$, we can derive that

$$\begin{aligned}\rho v(x_0) &= \lim_{\Delta t\to 0}\left\{\pi(x_0)(1+\rho\Delta t) + \frac{1}{\Delta t}\mathcal{E}\left[\Delta v|X(0) = x_0\right]\right\} \\ &= \pi(x_0) + \frac{1}{dt}\mathcal{E}\left[dv|X(0) = x_0\right],\end{aligned} \tag{3.10}$$

where Δv is defined as $\left\{v[X(\Delta t, w)] - v(x_0)\right\}$. Since we can write the second term on the right hand side as

$$\mathcal{E}(dv) = \left[\mu v'(x) + \frac{1}{2}\sigma^2 v''(x)\right]dt,$$

we can rewrite (3.10) as

$$\rho v(x) = \pi(x) + \mu(x)v'(x) + \frac{1}{2}\sigma^2(x)v''(x), \ \forall \ x, \tag{3.11}$$

which is the HJB equation. The interpretation of the last equation is that the return on the asset, the term on the left hand side, is equal to the sum of the dividend and the capital gain, the latter of which is captured by the last two terms on the right hand side. Note that in a deterministic setting $\sigma^2(x)$ is zero and (3.11) thus simplifies to the standard Bellman equation.

4. Applications

This section discusses papers that use stochastic environments in the models discussed.

4.1. Clarke & Reed (1989)

The standard deterministic tree cutting problem prescribes cutting down the tree when the proportional growth rate $\dot{R(t)}/R(t)$ at time t equals the market discount rate δ. Clarke and Reed extend this model to allow the asset's intrinsic value to evolve according to a continuous-time stochastic process. In a tree-cutting problem randomness can be introduced into the model in two ways: through the physical size of a biological asset and through the intrinsic value of a unit (price). The authors construct a model in which the physical size is subject to random variability in growth (size fluctuates according to a stochastic differential equation) and price follows GBM. Specifically, the following functional forms are assumed:

$$dq = bdt + \sigma_q dW_q \tag{4.1}$$

and

$$dy = g(t)dt + \sigma_y dW_y. \tag{4.2}$$

Note that $q(t) \sim N\left[q(0) + bt, \sigma_q^2 t\right]$, thus equation (4.1) says the asset's price, $P(t)$ which equals $\exp\left[q(t)\right]$, evolves according to GBM with constant drift, b, and constant variance σ_q^2. The stochastic differential equation in (4.2) shows that $\log X(t)$, which equals $y(t)$, where $X(t)$ is the asset's size at time t, behaves like Brownian motion

with time-dependent drift $g(t)$ and variance σ_y^2. Note that the drift, $g(t)$ depends only on the asset's age, t, and not the asset's size, in which case the drift would be $h(y)$. Optimal harvest rules with size dependent policies are more complicated as size can fluctuate over an asset's lifetime, whereas age is deterministic. The authors discuss the former in Reed and Clarke (1990) and the latter in this paper.

By applying Itô's lemma to (4.1) and using the definition of $P(t)$ we obtain the following

$$dP = \left(b + \frac{1}{2}\sigma_q^2\right) P dt + \sigma_q P dW_q \tag{4.3}$$

where $P(t)$ is lognormally distributed. What kind of implications does this have? Since $q(t)$ is defined as $\log P(t)$, the expectation of $P(t)$ increases exponentially over time. Note that

$$\begin{aligned} \mathsf{Prob}\left\{P(t) < \mathcal{E}\left[P(t)\right]\right\} &= \mathsf{Prob}\left[P(t) < P(0)\exp\left(bt + \frac{1}{2}\sigma_q^2 t\right)\right] \\ &= \mathsf{Prob}\left[q(t) < q(0) + bt + \frac{1}{2}\sigma_q^2 t\right] \\ &= \Phi\left(\frac{1}{2}\sigma_q\sqrt{t}\right), \end{aligned} \tag{4.4}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. Thus, as t approaches ∞ almost all sample asset price paths will end up below the exponential curve. Likewise applying Itô's lemma to (4.2) and using the definition of $X(t)$ we derive

$$dX = \left[g(t) + \frac{1}{2}\sigma_y^2\right] X dt + \sigma_y X dw_y. \tag{4.5}$$

If we assume risk neutrality then the market value $W(y, q, t)$ is the expected present value given that we choose to cut the tree down at the optimal time

$$W(y, q, t) = \sup_{\tau \geq t} \mathcal{E}\left\{\exp\left[-\delta(\tau - t) + q(\tau) + y(\tau)\right]\right\}. \tag{4.6}$$

Clarke and Reed then use an optimal stopping rule and assume a myopic-look-ahead (MLA) rule to simplify the problem. A MLA rule in a discrete-time model says to stop when the actual gain from cutting the tree is greater than the expected gain from waiting one-period longer. Analogously, in a continuous-time model we want

the assets value at τ, $R(\tau)$ which is $\exp[y(\tau)+q(\tau)]$, to equal the discounted expected value at an infinitesimal time later, $\tau + d\tau$; *i.e.*, we need

$$\exp[y(\tau)+q(\tau)] = \mathcal{E}_\tau \left\{ \exp(-\delta d\tau) \exp[q(\tau)+dq+y(\tau)+dy] \right\}.$$

This condition and our specified processes in (4.1) and (4.2) imply the following MLA rule

$$g(\tau_M) = \delta - b - \frac{1}{2}\left(\sigma_y^2 + \sigma_q^2 + 2\sigma_{yq}\right) \tag{4.7}$$

which implies with age-dependent growth the optimal rule is to harvest the tree at a fixed age, τ_M, where the right-hand side of the MLA rule is the nonrandom component of the asset's proportional growth. Note that the policy is independent of the price at τ_M! This latter conclusion results because price follows GBM. Consider now applying Itô's lemma and using the definition of $R(t)$ to obtain

$$\mathcal{E}(dR) = \left\{ [b+g(t)] + \frac{1}{2}\left(\sigma_y^2 + \sigma_q^2 + 2\sigma_{yq}\right) \right\} R dt + \mathcal{O}(dt)$$

which, using our MLA rule (4.7) implies

$$\mathcal{E}_{\tau_M}\left[\frac{dR_{\tau_M}}{R_{\tau_M}}\right] = \delta d\tau_M$$

just as in the standard deterministic optimal tree cutting case. The authors proceed to show that the MLA rule is the optimal harvest rule in the age-dependent growth model by showing that (4.6) satisfies the HJB equation and all boundary conditions.

Clarke and Reed then derive some comparative statics results that show price and growth uncertainty increase τ_M and that either an increase in price drift or a reduction in the discount rate will increase τ_M and W_M (the current market value). They also show that increases in volatility σ_y (σ_q) increase the value of the asset, which simply obtains from the implication that $\mathcal{E}[X(t)]$ ($\mathcal{E}[P(t)]$) increase exponentially in σ_y^2 (σ_q^2).

4.2. Reed & Clarke (1990)

The variables in the biological asset valuation model are defined as follows:

- $P(t)$ = the unit price of the asset at time t,

- $q(t) = \log[P(t)]$,
- $X(t) =$ the size of the asset at time t,
- $y(t) = \log[X(t)]$,
- $R(t) = P(t)X(t)$, the aggregate intrinsic value (revenue yielded by harvesting) at time t,
- $V(X, P, t) =$ the asset's market value when its size is X and the price is P at time t,
- $\delta =$ positive, constant, instantaneous discount rate.

PRICE: The price follows a stochastic process called geometric Brownian motion

$$\frac{dp}{P} = \mu dt + \sigma_q dw_q.$$

The Brownian motion for prices indicates that the prices evolve with a drift, μ, and a variance, σ_q^2

$$dq = bdt + \sigma_q dw_q,$$

where

$$b = \mu - \frac{1}{2}\sigma_q^2$$

if Itô Calculus is used.

SIZE: We might have two different geometric Brownian motions for the size of the biological asset

$$dy = g(t)dt + \sigma_y dw_y \tag{4.8}$$

or

$$dy = h(y)dt + \sigma_y dw_y. \tag{4.9}$$

The first case is called pure age-dependent growth, in which the proportional growth in size depends on the current age t, whereas the second case is called pure size-dependent growth, in which the proportional growth in size depends on the current size of the biological asset y. The choice of size does not only depend on the biological asset analyzed in a study, but also the availability of the information about the asset. The pure size-dependent growth is assumed throughout this paper. The authors assume $h(y)$ is decreasing, representing "*compensatory density-dependent growth.*"

Assuming risk-neutrality and ignoring costs

$$V(X,P,t) = \sup_{\tau \geq t} \mathcal{E}\left\{\exp(-\delta(\tau - t))P(\tau)X(\tau)|P(t) = P, X(t) = X\right\}.$$

If we use y and q instead of X and P, by assuming stationarity, we get

$$V(X,P,t) = W(y,q) = \sup_{\tau \geq t} \mathcal{E}\left\{\exp[-\delta\tau + q(\tau) + y(\tau)]|q(0) = q, y(0) = y\right\}. \quad (4.10)$$

STOPPING RULE: Given that we have our objective function, now we can talk about the *optimal stopping rule*. A stopping rule partitions the $y - q$ space into two regions: stopping, and continuation regions. The process is stopped, i.e. the maximization problem is solved when the pair $\{y(t), q(t)\}$ is in the continuation region. Given the initial conditions and the objective function and using a stopping rule S, we can use the HJB equation to find

$$\delta W^s(y,q) = h(y)W^s_y + bW^s_q + \frac{1}{2}\sigma^2_y W^s_{yy} + \frac{1}{2}\sigma^2_q W^s_{qq}. \quad (4.11)$$

The *free-boundary problem* can be defined by the HJB equation, with the initial conditions:

$$W^s(y,q) = e^{(y+q)},$$

and the smooth-pasting conditions, which are:

$$W^s_y(y,q) = \frac{\partial}{\partial y}e^{(y+q)} = e^{(y+q)},$$

and

$$W^s_q(y,q) = \frac{\partial}{\partial q}e^{(y+q)} = e^{(y+q)}.$$

BARRIER RULE: A simple possible solution, which turns out to give be optimal, is the *barrier rule*. A barrier rule basically says that whenever y exceeds some barrier value $\bar{y}$, the process will stop. In this case, the solution will give us

$$\bar{W}(y,q) = \varepsilon\left\{exp(-\delta T_{y,\bar{y}} + q(T_{y,\bar{y}}) + \bar{y})|q(0) = q, Y(0) = y\right\}, \quad (4.12)$$

where $T_{y,\bar{y}}$ is the first passage time for the $\{y(t)\}$ process to reach $\bar{y}$ from an initial state y.

MYOPIC-LOOK-AHEAD RULE: This rule basically keeps the process on the continuation region as long as the expected discounted intrinsic value of the asset an infinitesimal time ahead of the current time, is larger than the current intrinsic value of the asset. Then, the stopping condition will be

$$R(\tau) = \varepsilon_\tau \left[e^{-\delta\tau} R(\tau + d\tau) \right], \tag{4.13}$$

where τ shows that the expectation is conditional on the state variables at time τ.

In Clarke and Reed (1989), discussed above, we showed that when the Brownian motion for the size is pure age-dependent, the MLA rule is an optimal harvest rule. However, in the case of the pure size-dependent evolution of size variable, the MLA rule is not optimal.

ONGOING vs SINGLE ROTATIONS: Instead of single rotation, *i.e.*, harvesting the land only once, there may be ongoing rotations, *i.e.*, the land becomes available for a new forest growth after harvesting. In the second case, the net present value of the revenues derived from a piece of land from all future harvests is known as *the land expectation value*, and in the case of a deterministic model; it is given as

$$R(T)/(1 - e^{-\delta T}).$$

In the stochastic case, it is going to look like

$$L(P_0) = \sup \varepsilon \left\{ \sum_{i=1}^{\infty} e^{-\delta T_i} P(T_i) \widehat{X}(T_i) | P(0) = P_0, X(0) = x_0 \right\}, \tag{4.14}$$

where the supremum is taken over $\{T_i\}_{i=1:2,\ldots}$; and $\widehat{X}_i$ is $X(T_i - T_{i-1})$, the size of a stand at absolute time T_i, and X_0 is the initial size of a newly planted stand.

The result in the ongoing rotation case is that the optimal harvest time is shorter than that of the single rotation case, which is a generalization of the tree-cutting rules of the Faustmann and Wicksell types, respectively.

B. Bibliography

Clarke, Harry R. and William J. Reed. (1989): "The Tree-Cutting Problem in a Stochastic Environment," *Journal of Economic Dynamics & Control*, 13, 569–595.

Mitchell, Matthew F. (2004): Class notes, University of Iowa.

Reed, William J. and Harry R. Clarke. (1990): "Harvest Decisions and Asset Valuation for Biological Resources Exhibiting Size-Dependent Stochastic Growth." *International Economic Review*, 31, 147–169.

Ross, Sheldon M. (1997): *Introduction to Probability Models, 6e,* San Diego, CA: Academic Press.

Stokey, Nancy L. (1998): Course Materials, `http://home.uchicago.edu/ nstokey/`, University of Chicago.

Stokey, Nancy L. and Robert E. Lucas, Jr. with Edward C. Prescott. (1989): *Recursive Methods in Economic Dynamics,* Cambridge, MA: Harvard University Press.