Very simple notes (need to add references). It is NOT meant to be a substitute for a real course in stochastic calculus, just listing heuristic derivations of the stuff most oftenly used in economics. Ito calculus is a lot more than only dealing with Poisson jumps and Wiener processes. Some abuses of notation included without clarification.

1. Stochastic Process

A stochastic process is a collection of random variables (measurable functions)

$$\{X_t : t \in T\}, \quad X_t : \Omega \to S,$$

ordered in t (time), along with a measurable space (S, Σ). The probability space $(\Omega, \mathcal{F}, P)$ denotes, respectively, the "state space" (set of all possible histories), the σ-algebra that contains all possible sets (Borel sets) of histories induced by Ω, and the probability measure over $\mathcal{F}$. The space (S, Σ) contains the range of the function $X_t : \Omega \mapsto S$ and its corresponding σ-algebra. For example, for most of our applications $X_t \in \mathbf{R}$ or $\mathbf{R}_+$.

If X_t is measurable, any process induces a measure P_t that we can construct using the original probability space. This calls for the notion of a *filtration*: a weakly increasing collection of Borel sets on Σ, $\{\mathcal{F}_t, t \in T\}$, s.t. for all $s < t \in T$,

$$\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}.$$

The process X is adapted to the filtration $\{\mathcal{F}_t\}_{t \in T}$ if X_t is F_t-measurable. This just means that for any X_t, I can compute the probability only using $\mathcal{F}_t$ and not all of $\mathcal{F}$. Hence, a well defined stochastic process is always adapted to its *natural filtration*

$$\mathcal{F}_t = \sigma\left(\left\{X_s^{-1}(A) : s \leq t, A \in \Sigma\right\}\right).$$

This just means that for any history of X_t up to time t, all possibly realizable trajectories can be mapped backed into a subset of $\mathcal{F}_t$, so that I can compute its probability for all points up to time t. This generates an *induced* probability measure over X.

EXAMPLE 1 *Let $\Omega = [0,1]^\infty$. Then any $\omega \in \Omega$ is just a coordinate on the infinite dimensional unit cube. If we let $X_t : \Omega \mapsto S$ denote the t-th coordinate, S is just the unit interval $[0,1]$. If we construct, say, the probability measure so that $P = P_1 \times \cdots \times P_\infty$, where each P_t is the uniform distribution, X_t is i.i.d. uniform.*

2. Poisson (Jump) Process

Let N_t be the random variable equal to the number of "hits" up to time t. The (adapted) state space is $\mathbf{R}_{[0,t]}$ and range is all right-continuous paths that increase by 1. Now define the probability measure over ω as Poisson:

$$P\{N_t - N_s = n\} = \frac{(\lambda(t-s])^n}{n!} \cdot \exp(-\lambda(t-s])$$

where λ is the rate of arrival. This is what is usually called the Poisson process. (Not to be confused with what we use more often in economics: X_t is a Compound Poisson Process (CPP) if if it changes to some value at rate λ_t, studied below. In fact this is a new random variable in which X_t changes to some value if $N_t > N_s$ for all $s < t$, and you could redefine the probability space to the histories of N_t rather than $R_{[0,t]}$. This is the set of all right-continuous paths that increase by 1.)

More typically, the Poisson process is defined as a counting process:

Definition 1 *A continuous stochastic process N_t is Poisson if*

1. *N_t is a counting process:*
   (a) *N_t lives in $(\mathbf{Z}_+, 2^{Z_+})$, for all $t \geq 0$,*
   (b) *$N_s \leq N_t$ for all $s \leq t$,*
   (c) *$\lim_{s\downarrow t} N_s \leq \lim_{s\uparrow t} N_s$ for all $t \geq 0$; that is, no hit can happen simultaneously.*
2. *$N_0 = 0$ a.s.,*
3. *N is a stochastic process with stationary, independent increments*

The two definitions are equivalent; there are many other definitions as well but I refer you to the internet. It is easier to show that the earlier definition implies the counting process; by definition, increments are independent. The probability of getting 0, 1, or 2 or more hits in a time interval $dt > 0$ is

$$\begin{aligned}
P(N_{t+dt} - N_t = 0) &= \exp(-\lambda dt) && \approx 1 - \lambda dt + o(dt) \\
P(N_{t+dt} - N_t = 1) &= \lambda dt \cdot \exp(-\lambda dt) && \approx \lambda dt - \lambda^2 dt^2 + o(dt) \approx \lambda dt \\
P(N_{t+dt} - N_t \geq 2) &= (\lambda dt)^2 \cdot e^{-\lambda dt}/2 + o(dt) = o(dt).
\end{aligned}$$

Clearly, the actual probability that something happens in any interval dt (and $(t, t+dt]$, since the increments are independent) is 0. Conversely, one way to make sense of the counting process is to realize that stationarity implies

$$\mathbb{E}[N(T)/T] = \lim_{T\to\infty} N(T)/T = \lambda$$

and instead of sending T to infinity, send the number of intervals dt in $(0, T]$ to infinity to get that the expected number of hits in any given time interval is λ:

$$\begin{aligned}
\mathbb{E}[dN_t] &= \mathbb{E}[N(dt)] \\
&= \lambda dt = 0 \cdot P(N(dt) = 0) + 1 \cdot P(N(dt) = 1) + \sum_{n=2}^{\infty} n \cdot P(N(dt) = n)
\end{aligned}$$

$$= P(N(dt) = 1)$$

since two hits cannot occur at the same time. This is important later when we derive the stochastic HJB equation.

2.1 Compound Poisson Process

Now define a jump process over the underlying Poisson process: Let X_t be a r.v. that is γ_a if N_t is even and γ_b if N_t is odd. Heuristically,

$$\begin{aligned}
\mathbb{E}[dX_t] &= 0 \cdot \exp(-\lambda dt) + (\gamma_b - \gamma_a) \cdot \lambda dt \cdot \exp(-\lambda dt) + 0 \cdot (\lambda dt)^2 \cdot \exp^{-\lambda dt}/2 \\
\mathbb{E}[\dot{X}_t] &= \lim_{dt \to 0} (\gamma_b - \gamma_a) \cdot \lambda \cdot \exp(-\lambda dt) = \lambda(\gamma_b - \gamma_a)
\end{aligned}$$

More generally, let $\{Z_k\}_{k \geq 1}$ be an i.i.d. ordered sequence of random variables with measure $G_z(z)$, independent of the Poisson process N_t. Let X_t be a continuous stochastic process that is a function of (N_t, Z_k), and define

$$X_t = \sum_{k=1}^{N_t} Z_k.$$

Then

$$\begin{aligned}
EX_t &= \sum_{n=1}^{\infty} \left\{ \mathbb{E}\left[\left(\sum_{k=1}^{n} Z_k | N_t = n \right) \mathbb{P}(N_t = n) \right] \right\} = \sum_{n=1}^{\infty} \frac{e^{-\lambda t}(\lambda t)^n}{n!} \\
&= \mathbb{E}\left[\left(\sum_{k=1}^{n} Z_k \right) \right] = \lambda t \mu_Z \sum_{n=1}^{\infty} \frac{e^{-\lambda t}(\lambda t)^{n-1}}{(n-1)!} \quad = \mu_Z \cdot \lambda t
\end{aligned}$$

and

$$dX_t = Z_{N_t} dN_t = \int_0^t Z_{N_t} dN_t$$

assuming $X_0 = 0$.

2.2 Stochastic Integral with Poisson

First consider a function $f(N_t)$. The integral is easy to write as

$$\begin{aligned}
f(N_t) - f(0) &= \sum_{k=1}^{N_t} [f(k) - f(k-1)] = \int_0^t [f(1 + N_{s^-}) - f(N_{s^-})] \, dN_s \\
&= \int_0^t [f(N_s) - f(N_s - 1)] \, dN_s = \int_0^t [f(N_s) - f(N_{s^-})] \, dN_s,
\end{aligned}$$

where N_{s-} is the left limit of the Poisson process, and only one jump occurs in an dt by definition (or construction) of the Poisson process.

For the compound process, recall that the waiting time for the kth hit of the Poisson process, T_k, is also a random variable s.t. that the event $\{T_k > t\} \Leftrightarrow \{N_t \geq k-1\}$; in particular this means that $T_k - T_{k-1}$ is an i.i.d. process by definition.

For $k = 1$, the waiting time follows an exponential distribution. For $k > 1$,

$$\mathbb{P}(T_k > t) = \lambda \int_t^\infty \frac{e^{-\lambda s}(\lambda s)^{n-1}}{(n-1)!} ds, \tag{1}$$

since

$$\begin{aligned}
\mathbb{P}(T_k > t) &= \mathbb{P}(T_k > t \geq T_{k-1}) + \mathbb{P}(T_{k-1} > t) \\
&= \mathbb{P}(N_t = n-1) + \lambda \int_t^\infty \frac{e^{-\lambda s}(\lambda s)^{n-2}}{(n-2)!} ds \\
&= \frac{e^{-\lambda t}(\lambda t)^{n-1}}{(n-1)!} + \lambda \int_t^\infty \frac{e^{-\lambda s}(\lambda s)^{n-2}}{(n-2)!} ds
\end{aligned}$$

and integration by parts leads to (1). Using waiting times, the stochastic integral of a function of a compound Poisson process can be written

$$\begin{aligned}
f(Y_t) - f(0) &= \sum_{k=1}^{N_t} [f(Y_{T_k^-} + Z_k) - f(Y_{T_k^-})] = \int_0^t [f(Y_{s^-} + Z_{N_s}) - f(Y_{s^-})] \, dN_s \\
&= \int_0^t [f(Y_s) - f(Y_{s^-})] \, dN_s \\
&= \int_0^t [f(Y_s) - f(Y_{s^-})] \, (dN_s - \lambda ds) + \lambda \int_0^t [f(Y_s) - f(Y_{s^-})] \, ds.
\end{aligned}$$

3. Wiener Process (Brownian Motion)

DEFINITION 2 *A Wiener process is defined by four properties:*

1. *$W_0 = 0$ a.s.*
2. *Independent increments: $W_t - W_s$ is independent of $\mathcal{F}_s$ for all $s \leq t$*
3. *Normality: $W_t - W_s \sim \mathcal{N}(0, t-s)$*
4. *W_t is continuous a.s.*

We could spend the whole semester just talking about this, which we won't. Basically, think of Brownian motion as a random walk in continuous time: the best predictor of dX_t is 0, with Gaussian errors. So clearly, W_t is a particular type of a martingale ($\mathbb{E}[W_t|\mathcal{F}_s] = W_s$ a.s., for all $0 \leq s < t < \infty$).

Most commonly you will encounter a Brownian motion with drift, a geometric Brownian motion, or a generic (Ito) diffusion process:

$$dX_t = \mu dt + \sigma dW_t, \quad dX_t = \mu X_t dt + \sigma X_t dW_t, \quad dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$$

the geometric Brownian motion simply gives

$$dX_t/X_t = d\log X_t = \mu dt + \sigma dW_t,$$

so it is the just a Brownian motion with drift in percentage points (or log-points, to be exact). In the Ito process, the instantaneous drift and variance depend on the current value of X_t and is related to the version of Ito's Lemma that we will look at below.

Before we move along, note that both the Poisson process and Brownian motion are Markov processes, but while the Brownian motion has a continuous time path a.s., the Poisson process has a discontinuous time path a.s. Also, Poisson was not a martingale, but $dN_t - \lambda dt$ was.

It will be useful to know the quadratic variation of the Brownian motion: we will use a particular formulation that exploits the CLT in discrete time:

$$\begin{aligned}
\langle W \rangle_t &\equiv \mathbb{E}[W_t^2] = \lim_{n\to\infty} \sum_{i=1}^{2^n-1} [\Delta_i W_t]^2 \\
\Delta_i W_t &\equiv W_{t_{i+1}^n} - W_{t_i^n}
\end{aligned}$$

where $t_i^n \equiv it/2^n$. This makes the difference in adjacent W equal $t/2^n$, so

$$Z_i \equiv \sqrt{2^n}[\Delta_i W_t] \sim \mathcal{N}(0,t).$$

That is, all Z_i are normal with variance t. Since

$$\sum_{i=1}^{2^n-1} \Delta_i W_t = \sum_{t=1}^{2^n-1} Z_i^2/2^n,$$

the term converges to t a.s. by SLLN:

$$\langle W \rangle_t \equiv \mathbb{E}[W_t^2] = t.$$

Hence the quadratic variation of a Brownian motion is equal to t. This is an important notion that will help us understand Ito.

Conversely, suppose Z_i is a random walk s.t. ΔZ equals $\pm\Delta h$ with probability $(p, 1-p)$. That is, Z is Bernoulli. So $\mathbb{E}\Delta Z = \Delta h(2p-1)$ and $\mathbb{V}\Delta Z = 4p(1-p)(\Delta h)^2$. Now we repeat this process n times; this is a Bernoulli process and we can write

$$\begin{aligned}
\mathbb{E}[\Delta_n Z] &= n\Delta h(2p-1) &&= T\Delta h(2p-1)/\Delta t &&\equiv \mathbb{E}[\Delta_T Z] \\
\mathbb{V}[\Delta_n Z] &= n(\Delta h)^2 4p(1-p) &&= T(\Delta h)^2 4p(1-p)/\Delta t &&\equiv \mathbb{V}[\Delta_T Z],
\end{aligned}$$

where all we have done is to consider that the n trials happened in a time interval T with time $\Delta t = T/n$ increments. If we want this process to converge to a Wiener process as $\Delta t \to 0, n \to \infty$, we just choose Δh and p so that

$$\begin{aligned} &\Delta h(2p-1)/\Delta t = \mu, \quad (\Delta h)^2 4p(1-p)/\Delta t = \sigma^2, \\ \Rightarrow \quad &\begin{cases} \Delta h = \sigma\sqrt{\Delta t}\left[\sqrt{1+(\mu/\sigma)\cdot\Delta t}\right] = \sigma\sqrt{\Delta t} & \text{as } \Delta t \to 0 \\ p = \left[1 \pm (\mu/\sigma)\cdot\sqrt{\Delta t}\Big/\sqrt{1+(\mu/\sigma)\cdot\Delta t}\right]\Big/ 2 \end{cases} \end{aligned}$$

For the standard BM W_t, $\mu = 0$ and $\sigma = 1$, so $\Delta h = \sqrt{\Delta t}, p = 1/2$. So BM can be viewed as the limit of the sum of Bernoulli i.i.d. r.v.'s with $1/2$ probabilities ± 1:

$$W_T = \lim_{\Delta t \to 0}[\Delta_T Z] = \lim_{n\to\infty}[\Delta_n Z]$$

which converges to $\mathcal{N}(0, T)$ since $\sqrt{\Delta}t = \sqrt{T/n}$.

4. Stochastic Integral with BM and CPP

The Ito stochastic integral is defined on a semi-martingale (basically, a random walk plus a process with finite variation), where the underlying martingale component has finite quadratic variation. That is,

$$X_t = X_0 + B_t + M_t \quad \Rightarrow \quad dX_t = dB_t + dM_t,$$

where $\mathbb{E}[M_t|\mathcal{F}_t] = 0$, B_t is adapted to $\mathcal{F}_t$, and $\langle M \rangle_t < \infty$. For example, if M_t is BM, $\langle M \rangle_t = t$. Similarly if M_t is the compensated Poisson process $N_t - \lambda t$, $\langle M \rangle_t = N_t$.

THEOREM 1 (ITO'S LEMMA FOR CONTINUOUS MARTINGALES) *If X_t is continuous and f is a three-times continuously differentiable function, the stochastic integral of $f(X_t)$ is*

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t f'(X_s)dB_s + \int_0^t f'(X_s)dM_s + \frac{1}{2}\int_0^t f''(M_s)d\langle M \rangle_s \\ \text{or} \quad df(X_t) &= f'(X_t)dB_t + f'(X_t)dM_t + \frac{1}{2}f''(M_t)d\langle M \rangle_t. \end{aligned}$$

Note that this version of Ito does not apply to Poisson. No proof given, but the intuition is that Ito extends Riemann integrals to stochastic increments:

$$\int_0^t f(X_t)dX_t = \lim_{n\to\infty}\sum_{i=1}^n f(X_{t_{i-1}})\left[X_{t_i} - X_{t_{i-1}}\right]$$

where Π_n is an n-partition of $[0, t]$. It is important that the point of approximation for each interval is taken from the left. Also importantly, the stochastic integral itself is not a deterministic concept: it is the martingale such that its quadratic variation equals the expectation of the square of all realized paths integrated over $\langle X \rangle_t$.

Formally, note that any function of M_t is simply a process Y_t that is adapted to $\mathcal{F}_t$, the filtration of the martingale. Let M_t be square-integrable in the sense that

$$\langle M \rangle_t = \mathbb{E}M_t^2 < \infty.$$

Let $Y^{(k)}$ be simple process, that is for an infinitely fine partition $\{t_i\}_{i=0}^{\infty}$ on $[0, t]$, and countably infinite sequence of random variables $\{\zeta_i^{(k)}\}_{i=0}^{\infty}$,

$$Y_t^{(k)} = \zeta_0^{(k)} \mathbb{1}(0) + \sum_{i=1}^{\infty} \zeta_{i-1}^{(k)} \mathbb{1}(t_{i-1} - t_i].$$

One definition of the stochastic integral of Y_t over M_t is the (unique) square integrable martingale $I_t(Y)$ s.t.

Definition 3 (Heuristic Definition of Stochastic Integral) *For all sequences of simple processes* $\lim_{k\to\infty}[Y^{(k)} - Y] = 0$ *(in quadratic variation):*

$$\lim_{k\to\infty} \|Y^{(k)} - Y\|^2 = \lim_{k\to\infty} \mathbb{E}[Y^{(k)} - Y]^2$$

$I(Y)$ is the martingale s.t.

$$\lim_{k\to\infty} \|I(Y^{(k)}) - I(Y)\|^2 = \lim_{k\to\infty} \mathbb{E}\left[I(Y^{(k)}) - I(Y)\right]^2 = 0,$$

where for each $Y^{(n)}$,

$$I_t(Y^{(k)}) = \sum_{i=1}^{\infty} \zeta_{i-1}^{(k)} \left[M_{t_i} - M_{t_{i-1}}\right].$$

This is just a complicated way of saying the stochastic integral is a Riemann-Stieltjes integral where the measure of integration is stochastic (so need Lebegue). When and when it doesn't work, and why it's unique, we won't worry about. Perhaps the most important property of the stochastic integral defined as such is that

$$\mathbb{E}[I_t(Y)|\mathcal{F}_s] = I_s(Y) \quad \text{and} \quad \mathbb{E}[I_t(Y)^2] = \mathbb{E}\left[\int_0^t Y_s^2 d\langle M \rangle_s\right]$$

where the first part just means it is a martingale, and the second that *the square can be taken inside the integral*. The problem is when $M_{t_i} - M_{t_{i-1}}$ goes to ∞, but Ito tells us that as long as $\langle M \rangle_t$ is bounded we can define an integral.

[*Graphical Representation of Riemann-Stieltjes and Ito*]

Although the above doesn't apply to Poisson, we already know how to write the integral for CPP. Consider the general jump diffusion process (which is all we're going to deal with, really)

$$\begin{aligned} X_t - X_0 &= \mu(t, X_t) + \sigma(t, X_t) W_t + Y_t \\ \Rightarrow \quad dX_t &= \underbrace{\mu(t, X_t)dt}_{dB_t} + \underbrace{\sigma(t, X_t)dW_t}_{dM_t} + \underbrace{dY_t}_{\text{jumps}}, \end{aligned} \tag{2}$$

where W_t and Y_t are independent Wiener and CPP. Then since $\langle W \rangle_t = t$, we have

$$\begin{aligned} f(X_t) - f(X_0) = &\int_0^t \mu(s, X_s) f'(X_s) ds + \int_0^t \sigma(s, X_s) f'(X_s) dW_s + \frac{1}{2} \int_0^t \sigma^2(s, X_s) f''(X_s) ds \\ &+ \int_0^t [f(X_s) - f(X_{s^-})] dN_s \\ \text{or} \quad df(X_t) = &\mu(s, X_s) f'(X_t) dt + \sigma(s, X_s) f'(X_t) dW_t + \frac{1}{2} \sigma^2(s, X_s) f''(M_t) dt \\ &+ [f(X_t) - f(X_{t^-})] dN_t. \end{aligned}$$

Note of caution: you **cannot** just write $\lambda[f(X_t) - f(X_{t^-})]dt$ instead of $[f(X_t) - f(X_{t^-})dN_t]$ there, since $dN_t - \lambda dt$ is only 0 in expectation. Without proving (again!) we can also let λ vary with time and state, of which the non-homogeneous Poisson process satisfies

$$\mathbb{E}\left[N_t - \int_0^t \lambda(s, X_s) ds\right] = 0 = \mathbb{E}_t \left[dN_t - \lambda(t, X_t) dt\right].$$

Intuitively, we can always reset the underlying Poisson process following any hit until the next hit arrives, during which the process remains “homogeneous."

5. Stochastic HJB

Let x_t denote realizations from the stochastic process X_t that follows the jump diffusion process (2). Now consider the stochastic control problem

$$\begin{aligned} v(t, x_t, a_t) &= \max_{(c_s)} \left\{ \mathbb{E}_t \int_t^T U(s, X_s, a_s, c_s) ds \right\} \\ \text{s.t.} \quad da_s &= f(s, x_s, a_s, c_s) ds \end{aligned}$$

where T can be finite or infinite, and I have suppressed the scrap value $Z(T, a_T, X_T)$ which may or may not be there. Except for x_t, all that has changed from the deterministic control problem is that we added an expectation operator over the objective.

But let us generalize the jumps a bit. Let Y_t, be associated with a non-homogeneous Poission process with time- and state-dependent arrival rates $\lambda(t, x_t)$, and also have jumps Z_k that are drawn from a time- and state-dependent measure $G_z(z; t, x_t)$. For example, X_t can be a random wage or dividend process, or interest rate process (in which case things can be simplified, since it would multiply the state a_s; likewise if we wanted a stochastic discount rate in which case it would show up multiplicatively in U).

Without going into the details, we can use similar methods as in deterministic control (stochastic versions of Taylor expansion, verification theorem) to show that the following heuristic method works:

$$0 \geq U(t, a_t, c_t)dt + \mathbb{E}_t\left[dV(t, x_t, a_t)\right]$$

where by Ito we have

$$\begin{aligned} dV = &\left[V_t + V_a f(t, x_t, a_t, c_t) + \frac{V_{xx}}{2}\sigma^2(t, x_t, a_t)\right] dt + \text{crap} \\ &+ \lambda(t, x_t)\left[\mathbb{E}_t V(t, x_t + Z_k, a_t) - V(t, x_t, a_t)\right] dt \end{aligned}$$

where we have set $X_{t^-} = x_t$ a.s., since it is already realized, and have allowed μ, σ to also depend on a. The crap term is

$$\begin{aligned} \mathbb{E}_t[\text{crap}] &= 0 \\ &= \mathbb{E}_t[\mu(t, x_t, a_t)V_x dW_t] + \mathbb{E}_t\left\{[V(t, x_t + Z_k, a) - V(t, x_t, a)] \cdot [dN_t - \lambda(t, x_t)]\right\} \end{aligned}$$

the latter since Z_k is independent of N_t. So we get the HJB equation

$$\begin{aligned} -V_t(t, x, a) =& \mathcal{H}^*(t, x, a, V_a(t, x, a)) + \frac{V_{xx}}{2}\sigma^2(t, x, a) \\ &+ \lambda(t, x)\left[\int V(t, x + z, a)dG_z(z; t, x) - V(t, x, a)\right] \end{aligned}$$

so when $U = e^{-\rho t}u$, we can multiply the whole system by $e^{\rho t}$ and define $v(t, x, a) \equiv e^{\rho t}V(t, x, a)$ to obtain

$$\begin{aligned} \rho v(t, x, a) =& v_t(t, x, a) + \hat{\mathcal{H}}^*(t, x, a, v_a(t, x, a)) + \frac{v_{xx}}{2}\sigma^2(t, x, a) \\ &+ \lambda(t, x)\left[\int v(t, x + z)dG_z(z; t, x) - v(t, x, a)\right]. \end{aligned}$$

Note that the only time the expectation operator comes in is for the (possibly) stochastic r.v. Z_k; everything else is adapted to $\mathcal{F}_t$. That is, for the HJB, there is no longer any expectations taken over X_t; all of that is washed out in continuous time with martingales. Since the Hamiltonians are the same as in the deterministic case, it follows that the f.o.c. holds deterministically in continuous time, that is, $u_c(t, x, a, c) + v_a f(t, x, a, c) = 0$ (no expectations over v or v_a!)

6. Fokker-Planck (Kolmogorov Forward) Equation

The last tool that will be relevant for our purposes is the KFE. Given a solution to $v(t, x, a)$, we want to understand the evolution of $p(t, x, a)$, the population p.d.f. over (x, a) at time t. KFE gives us a (partial) differential equation that does exactly this. Formally, KFE tells you: suppose at time t, you know $P\{(x,a) \in B\}$ for all $B \in \mathcal{F}_t$. How is $P(B)$ evolving going forward in the filtration (for the same set B)?

For example, in the savings problem a solution to v admits optimal policy functions $c^*(t,x,a)$ and associated (change in) assets $\dot{a}^*(t,x,a)$. Now suppose at time t, the p.d.f. is represented as $p(t,x,a)$. KFE tells us how the distribution evolves *going forward*. (Conversely, Feynman-Kac, or the Kolmogorov Backward Equation, tells us how you would have got to $p(t,x,a)$ going backward; but we are not so interested in this). To compare with discrete time, it is as if we are simulating a distribution of individuals starting from some given initial distribution.

The following is a version of Fokker-Planck:

THEOREM 2 (FOKKER-PLANCK-KOLMOGOROV) *Let X_t be a stochastic process as in (2), where Y_t is a CPP with jumps $Z_k \sim G_z(t, X_t)$ associated with a non-homogeneous Poisson process with rate $\lambda(t, X_t)$. Let $p(t,x)$ denote the p.d.f. of x at time t. Then for all $x \in (x_{min}, x_{max})$ (the interior of possibly realizable states),*

$$\begin{aligned}\frac{\partial}{\partial t}p(t,x) = &-\frac{\partial}{\partial x}\left[\mu(t,x)p(t,x)\right] + \frac{\partial^2}{2\partial x^2}\left[\sigma^2(t,x)p(t,x)\right] \\ &- \lambda(t,x)p(t,x) + \int g_z(x-x';t,x')\lambda(t,x')p(t,x')dx'.\end{aligned}$$

Proof (heuristic): The trick is to use a function that is differentiable, so we can apply Ito. For any $x \in S$ (the range of X_t), approximate the probability of the event by the expectation of a smooth function:

$$P(X_t \leq x) = \int^x p(t,x')dx' = \mathbb{E}[\chi(X_t \leq x)] = \int \chi(x' \leq x)dx'.$$

With some abuse of notation, we will assume that the indicator is already smoothed (it does not matter how it is smoothed; convolving it with any mollifer will do). Using Ito, we obtain (the derivative is w.r.t. time):

$$\begin{aligned}dP(X_t \leq x) =&\mathbb{E}[\chi'(X_t \leq x)\mu(t,X_t) + \frac{1}{2}\sigma^2(t,X_t)]dt + \text{crap} \\ &+ \mathbb{E}\left\{\lambda(t,X_t)[\chi(X_{t^-} + Z_k \leq x) - \chi(X_{t^-} \leq x)]\right\}\end{aligned} \tag{3}$$

where crap is again 0 in expectations. Note that the derivatives of χ are w.r.t X_t, not x. First look at the Poisson part. To compute the expectation over X_t, we denote the variable of integration by $X_t = x'$, and remember that $X_t = X_{t^-} = x'$ a.s.:

$$\begin{aligned}\Rightarrow \quad &\int \lambda(t,x')\left[\int_z \chi(x'+z \leq x)dG(z;t,x') - \chi(x' \leq x)\right]p(t,x')dx' \\ = &\int \lambda(t,x')G_z(x-x';t,x')p(t,x')dx' - \int^x \lambda(t,x')p(t,x')dx'.\end{aligned}$$

Note that this is “as if" we were looking “backward," not “forward” like we did in the HJB. This is because we are looking at all the X_t when they are still random variables, not a realized point like in the HJB. (The density functions are deterministic in x, not $X_t = x'$).

For the diffusion part, we can integrate by parts:

$$\int \chi'(x' \leq x)\mu(t,x')p(t,x')dx' = B_1(x) - \int^x \frac{\partial}{\partial x'}[\mu(t,x')p(t,x')]dx',$$

where B_1 is a term determined by boundary conditions:

$$B_1(x) \equiv \chi(x_{\max} \leq x)\mu(t,x_{\max})p(t,x_{\max}) - \mu(t,x_{\min})p(t,x_{\min}).$$

And likewise

$$\begin{aligned}\int \chi''(x' \leq x)\sigma^2(t,x')p(t,x')dx' =& B_2(x) - \int \left\{\chi'(x' \leq x)\frac{\partial}{\partial x'}[\sigma^2(t,x')p(t,x')]\right\} dx' \\ =& B_2(x) - B_3(x) + \int^x \left\{\frac{\partial^2}{\partial x'^2}[\sigma^2(t,x')p(t,x')]\right\} dx'\end{aligned}$$

where B_j are determined by boundary conditions:

$$\begin{aligned}B_2(x) \equiv& \chi'(x_{\max} \leq x)\sigma^2(t,x_{\max})p(t,x_{\max}) - \chi'(x_{\min} \leq x)\sigma^2(t,x_{\min})p(t,x_{\min}) \\ B_3(x) \equiv& \chi(x_{\max} \leq x)\frac{\partial}{\partial x}\left[\sigma^2(t,x_{\max})p(t,x_{\max})\right] - \frac{\partial}{\partial x}\left[\sigma^2(t,x_{\min})p(t,x_{\min})\right].\end{aligned}$$

So we have obtained that (3) becomes

$$\begin{aligned}dP(t,x) = & \int^x \left[-\frac{\partial}{\partial x'}\mu(t,x') + \frac{\partial^2}{2\partial x'^2}\sigma^2(t,x') - \lambda(t,x')\right] p(t,x')dx' \\ & + \int \lambda(t,x')G_z(x-x';t,x')p(t,x')dx' + B_1(x) + B_2(x) - B_3(x)\end{aligned} \tag{4}$$

Note that $B_j'(x) = 0$ except at the boundaries. So taking the derivative w.r.t. x on both sides of (4) we obtain the formula in the theorem.

□

Of course in our economics problems, typically we want $p(t,x,a)$ not $p(t,x)$. But most problems will assume a law of motion s.t. x is subsumed in a, as we will see later.