Chapter 11

Examples of Markov Processes

Section 11.1 looks at the evolution of densities under the action of the logistic map; this shows how deterministic dynamical systems can be brought under the sway of the theory we've developed for Markov processes.

Section 11.2 finds the transition kernels for the Wiener process, as an example of how to manipulate such things.

Section 11.3 generalizes the Wiener process example to other processes with stationary and independent increments, and in doing so uncovers connections to limits of sums of IID random variables and to self-similarity.

11.1 Probability Densities in the Logistic Map

Let's revisit the first part of Exercise 5.3, from the point of view of what we now know about Markov processes. The exercise asks us to show that the density $\frac{1}{\pi\sqrt{x(1-x)}}$ is invariant under the action of the logistic map with $a = 4$.

Let's write the mapping as $F(x) = 4x(1-x)$. Solving a simple quadratic equation gives us the fact that $F^{-1}(x)$ is the set $\left\{\frac{1}{2}\left(1-\sqrt{1-x}\right), \frac{1}{2}\left(1+\sqrt{1-x}\right)\right\}$. Notice, for later use, that the two solutions add up to 1. Notice also that $F^{-1}([0,x]) = \left[0, \frac{1}{2}\left(1-\sqrt{1-x}\right)\right] \cup \left[\frac{1}{2}\left(1+\sqrt{1-x}\right), 1\right]$. Now we consider the cumulative distribution function of X_{n+1}, $\mathbb{P}(X_{n+1} \leq x)$.

$$\begin{aligned}
&\mathbb{P}(X_{n+1} \leq x) \\
&= \mathbb{P}(X_{n+1} \in [0, x]) && (11.1)\\
&= \mathbb{P}\left(X_n \in F^{-1}([0, x])\right) && (11.2)\\
&= \mathbb{P}\left(X_n \in \left[0, \frac{1}{2}\left(1 - \sqrt{1-x}\right)\right] \cup \left[\frac{1}{2}\left(1 + \sqrt{1-x}\right), 1\right]\right) && (11.3)\\
&= \int_0^{\frac{1}{2}\left(1-\sqrt{1-x}\right)} \rho_n(y)\, dy + \int_{\frac{1}{2}\left(1+\sqrt{1-x}\right)}^1 \rho_n(y)\, dy && (11.4)
\end{aligned}$$

where ρ_n is the density of X_n. So we have an integral equation for the evolution of the density,

$$\int_0^x \rho_{n+1}(y)\, dy = \int_0^{\frac{1}{2}\left(1-\sqrt{1-x}\right)} \rho_n(y)\, dy + \int_{\frac{1}{2}\left(1+\sqrt{1-x}\right)}^1 \rho_n(y)\, dy \quad (11.5)$$

This sort of integral equation is complicated to solve directly. Instead, take the derivative of both sides with respect to x; we can do this through the fundamental theorem of calculus. On the left hand side, this will just give $\rho_{n+1}(x)$, the density we want.

$$\begin{aligned}
&\rho_{n+1}(x) && (11.6)\\
&= \frac{d}{dx}\int_0^{\frac{1}{2}\left(1-\sqrt{1-x}\right)} \rho_n(y)\, dy + \frac{d}{dx}\int_{\frac{1}{2}\left(1+\sqrt{1-x}\right)}^1 \rho_n(y)\, dy \\
&= \rho_n\left(\frac{1}{2}\left(1-\sqrt{1-x}\right)\right)\frac{d}{dx}\left(\frac{1}{2}\left(1-\sqrt{1-x}\right)\right) && (11.7)\\
&\quad -\rho_n\left(\frac{1}{2}\left(1+\sqrt{1-x}\right)\right)\frac{d}{dx}\left(\frac{1}{2}\left(1+\sqrt{1-x}\right)\right) \\
&= \frac{1}{4\sqrt{1-x}}\left(\rho_n\left(\frac{1}{2}\left(1-\sqrt{1-x}\right)\right) + \rho_n\left(\frac{1}{2}\left(1+\sqrt{1-x}\right)\right)\right) && (11.8)
\end{aligned}$$

Notice that this defines a linear operator taking densities to densities. (You should verify the linearity.) In fact, this is a Markov operator, by the terms of Definition 117. Markov operators of this sort, derived from deterministic maps, are called *Perron-Frobenius* or *Frobenius-Perron* operators, and accordingly denoted by P. Thus an invariant density is a ρ^* such that $\rho^* = P\rho^*$. All the problem asks us to do is to verify that $\frac{1}{\pi\sqrt{x(1-x)}}$ is such a solution.

$$\begin{aligned}
\rho^*\left(\frac{1}{2}\left(1-\sqrt{1-x}\right)\right) && (11.9)\\
= \frac{1}{\pi}\left(\frac{1}{2}\left(1-\sqrt{1-x}\right)\left(1-\left(\frac{1}{2}\left(1-\sqrt{1-x}\right)\right)\right)\right)^{-1/2} \\
= \frac{1}{\pi}\left(\frac{1}{2}\left(1-\sqrt{1-x}\right)\frac{1}{2}\left(1+\sqrt{1-x}\right)\right)^{-1/2} && (11.10)\\
= \frac{2}{\pi\sqrt{x}} && (11.11)
\end{aligned}$$

Since $\rho^*(x) = \rho^*(1-x)$, it follows that

$$\begin{aligned}
P\rho^* &= 2\frac{1}{4\sqrt{1-x}}\rho^*\left(\frac{1}{2}\left(1-\sqrt{1-x}\right)\right) && (11.12)\\
&= \frac{1}{\pi\sqrt{x(1-x)}} && (11.13)\\
&= \rho^* && (11.14)
\end{aligned}$$

as desired.

By Lemma 135, for any distribution ρ, $\|P^n\rho - P^n\rho^*\|$ is a non-increasing function of n. However, $P^n\rho^* = \rho^*$, so the iterates of *any* distribution, under the map, approach the invariant distribution monotonically. It would be very handy if we could show that any initial distribution ρ eventually converged on ρ^*, i.e. that $\|P^n\rho - \rho^*\| \to 0$. When we come to ergodic theory, we will see conditions under which such distributional convergence holds, as it does for the logistic map, and learn how such convergence in distribution is connected to both pathwise convergence properties, and to the decay of correlations.

11.2 Transition Kernels and Evolution Operators for the Wiener Process

We have previously defined the Wiener process (Examples 38 and 79) as the real-valued process on $\mathbb{R}^+$ with the following properties:

1. $W(0) = 0$;
2. For any three times $t_1 \leq t_2 \leq t_3$, $W(t_3) - W(t_2) \perp\!\!\!\perp W(t_2) - W(t_1)$ (independent increments);
3. For any two times $t_1 \leq t_2$, $W(t_2) - W(t_1) \sim \mathcal{N}(0, t_2 - t_1)$ (Gaussian increments);
4. Continuous sample paths (in the sense of Definition 73).

In Exercise 4.1, you showed that a process satisfying points (1)–(3) above exists. To show that any such process has a version with continuous sample paths, invoke Theorem 97 (Exercise 11.3). Here, however, we will show that it is a homogeneous Markov process, and find both the transition kernels and the evolution operators. Markovianity follows from the independent increments property (2), while homogeneity and the form of the transition operators comes from the Gaussian assumption (3).

First, here's how the Gaussian increments property gives us the transition probabilities:

$$\begin{aligned}
\mathbb{P}\left(W(t_2) \in B | W(t_1) = w_1\right) &= \mathbb{P}\left(W(t_2) - W(t_1) \in B - w_1\right) && (11.15)\\
&= \int_{B-w_1} du \frac{1}{\sqrt{2\pi(t_2-t_1)}} e^{-\frac{u^2}{2(t_2-t_1)}} && (11.16)\\
&= \int_{B} dw_2 \frac{1}{\sqrt{2\pi(t_2-t_1)}} e^{-\frac{(w_2-w_1)^2}{2(t_2-t_1)}} && (11.17)\\
&\equiv \mu_{t_1,t_2}(w_1, B) && (11.18)
\end{aligned}$$

Since this evidently depends only on $t_2 - t_1$, and not the individual times, the process is homogeneous. Now, assuming homogeneity, we can find the time-evolution operators for well-behaved observables:

$$\begin{aligned}
K_t f(w) &= \mathbf{E}\left[f(W_t + s) | W_s = w\right] && (11.19)\\
&= \int f(u) \mu_t(w, du) && (11.20)\\
&= \int f(u) \frac{1}{\sqrt{2\pi t}} e^{-\frac{(u-w)^2}{2t}} du && (11.21)\\
&= \mathbf{E}\left[f(w + \sqrt{t}Z)\right] && (11.22)
\end{aligned}$$

where Z is a standard Gaussian random variable independent of W_s.

To show that $W(t)$ is a Markov process, we must show that, for any finite collection of times $t_1 \leq t_2 \leq \ldots \leq t_k$,

$$\mathcal{L}\left(W_{t_1}, W_{t_2}, \ldots W_{t_k}\right) = \mu_{t_1,t_2} \mu_{t_2,t_3} \ldots \mu_{t_{k-1},t_k} \qquad (11.23)$$

Let's just go through the $k = 3$ case, as the others are fundamentally similar, but with more notation. Notice that $W(t_3) - W(t_1) = (W(t_3) - W(t_2)) + (W(t_2) - W(t_1))$. Because increments are independent, then, $W(t_3) - W(t_1)$ is the sum of two independent random variables, $W(t_3) - W(t_2)$ and $W(t_2) - W(t_1)$. The distribution of $W(t_3) - W(t_1)$ is then the convolution of distributions of $W(t_3) - W(t_2)$ and $W(t_2) - W(t_1)$. Those are $\mathcal{N}(0, t_3 - t_2)$ and $\mathcal{N}(0, t_2 - t_1)$ respectively. The convolution of two Gaussian distributions is a third Gaussian, summing their parameters, so according to this argument, we must have $W(t_3) - W(t_1) \sim \mathcal{N}(0, t_3 - t_1)$. But this is precisely what we should have, by the Gaussian-increments property. Since the trick we used above to get the transition kernel from the increment distribution can be applied again, we conclude that $\mu_{t_1,t_2} \mu_{t_2,t_3} = \mu_{t_1,t_3}$. The same trick applies when $k > 3$. Therefore (Theorem 106), $W(t)$ is a Markov process, with respect to its natural filtration.

11.3 Lévy Processes and Limit Laws

Let's think a bit more about the trick we've just pulled with the Wiener process. We wrote the time evolution operators in terms of a random variable,

$$K_t f(w) = \mathbf{E}\left[f(w + \sqrt{t}Z)\right] \tag{11.24}$$

Notice that $\sqrt{t}Z \sim \mathcal{N}(0,t)$, i.e., it is the increment of the process over an interval of time t. This can be generalized to other processes of a similar sort. These are one natural way of generalizing the idea of a sum of IID random variables.

Definition 137 (Processes with Stationary and Independent Increments) *A stochastic process X is a* stationary, independent-increments process*, or* has stationary, independent increments *when*

1. *The increments are independent; for any collection of indices $t_1 \leq t_2 \leq \ldots \leq t_k$, with k finite, the increments $X(t_i) - X(t_{i-1})$ are all jointly independent.*
2. *The increments are stationary; for any $\tau > 0$,*

$$\begin{aligned} & \mathcal{L}\left(X(t_2) - X(t_1), X(t_3) - X(t_2), \ldots X(t_k) - X(t_{k-1})\right) \\ = \; & \mathcal{L}\left(X(t_2+\tau) - X(t_1+\tau), X(t_3+\tau) - X(t_2+\tau), \ldots X(t_k+\tau) - X(t_{k-1}+\tau)\right) \end{aligned} \tag{11.25}$$

Definition 138 (Lévy Processes) *A* Lévy process *is a process with stationary and independent increments and cadlag sample paths, which start at zero, $X(0) = 0$.*

Example 139 (Wiener Process is Lévy) *The Wiener process is a Lévy process (since it has not just cadlag but continuous sample paths).*

Example 140 (Poisson Counting Process) *The* Poisson counting process with intensity λ *is the integer-valued process N on $\mathbb{R}^+$ where $N(0) = 0$, $N(t) \sim$* Poisson(λt)*, and independent increments. It defines a Poisson point process (in the sense of Example 20) by assigning the interval $[0,t]$ the measure $N(t)$, the measure extending to other Borel sets in the usual way. Conversely, any Poisson process defines a counting process in this sense. N can be shown to be continuous in probability, and so, via Theorem 96, to have a cadlag modification. (Exercise 11.5.) This is a Lévy process.*

It is not uncommon to see people writing just "processes with independent increments" when they mean "stationary and independent increments".

Theorem 141 (Processes with Stationary Independent Increments are Markovian) *Any process with stationary, independent increments is Markovian with respect to its natural filtration, and homogeneous in time.*

PROOF: The joint distribution for any finite collection of times factors into a product of incremental distributions:

$$\begin{aligned}&\mathcal{L}(X(t_1), X(t_2), \ldots X(t_k)) &(11.26)\\ &= \mathcal{L}(X(t_1))\,\mathcal{L}(X(t_2) - X(t_1)) \ldots \mathcal{L}(X(t_k) - X(t_{k-1}))\end{aligned}$$

Define $\mu_t(x, B) = \mathbb{P}(X(t_1 + t) - X(t_1) + x \in B)$. This is a probability kernel. Moreover, it satisfies the semi-group property, since $X(t_3) - X(t_1) = X(t_3) - X(t_2) + X(t_2) + X(t_1)$ (and similarly for more indices). Thus the μ_t, so defined, are transition probability kernels (Definition 105). By Theorem 106, X is Markovian with respect to its natural filtration. □

Remark: The assumption that the increments are stationary is not necessary to prove Markovianity. What property of X does it then deliver?

Theorem 142 (Time-Evolution Operators of Processes with Stationary, Independent Increments) *If X has stationary and independent increments, then its time-evolution operators K_t are given by*

$$K_t f(x) = \mathbf{E}[f(x + Z_t)] \tag{11.27}$$

where $\mathcal{L}(Z_t) = \mathcal{L}(X(t) - X(0))$, the increment from 0 to t.

PROOF: Exactly parallel to the Wiener process case.

$$\begin{aligned}K_t f(x) &= \mathbf{E}[f(X_t)|\mathcal{F}_0] &(11.28)\\ &= \mathbf{E}[f(X_t)|X_0 = x] &(11.29)\\ &= \int f(y)\mu_t(x, dy) &(11.30)\end{aligned}$$

Since $\mu_t(x, B) = \mathbb{P}(X(t_1 + t) - X(t_1) + x \in B)$, the theorem follows. □

Notice that so far everything has applied equally to discrete or to continuous time. For discrete time, we can chose the distribution of increments over a single time-step, $\mathcal{L}(X(1) - X(0))$, to be essentially whatever we like; stationarity and independence then determine the distribution of all other increments. For continuous time, however, our choices are more constrained, and in an interesting way.

By the semi-group property of the time-evolution operators, we must have $K_s K_t = K_{t+s}$ for all times s and t. Applying Theorem 142, it must be the case that

$$\begin{aligned}K_{t+s} f(x) = \mathbf{E}[f(x + Z_{t+s})] & &(11.31)\\ &= K_s \mathbf{E}[f(x + Z_t)] &(11.32)\\ &= \mathbf{E}[f(x + Z_s + Z_t)] &(11.33)\end{aligned}$$

where $Z_s \perp\!\!\!\perp Z_t$. That is, the distribution of Z_{t+s} must be the convolution of the distributions of Z_t and Z_s. Let us in particular pick $s = t$; this implies $\mathcal{L}(Z_{2t}) = \mathcal{L}(Z_t) \star \mathcal{L}(Z_t)$, where $\star$ indicates convolution. Writing ν_t for $\mathcal{L}(Z_t)$, another way to phrase our conclusion is that $\nu_t = \nu_{t/2} \star \nu_{t/2}$. Clearly, this argument could be generalized to relate ν_t to $\nu_{t/n}$ for any n:

$$\nu_t = \nu_{t/n}^{\star n} \tag{11.34}$$

where the superscript $\star n$ indicates the n-fold convolution of the distribution with itself. Equivalently, for each t, and for all n,

$$Z_t \stackrel{d}{=} \sum_{i=1}^{n} D_i \tag{11.35}$$

where the IID $D_i \sim \nu_{t/n}$. This is a very curious-looking property with a name.

Definition 143 (Infinitely-Divisible Distributions and Random Variables) *A probability distribution ν is* infinitely divisible *when, for each n, there exists a ν_n such that $\nu = \nu_n^{\star n}$. A random variable Z is* infinitely divisible *when, for each n, there are n IID random variables $D_i^{(n)}$ such that $Z \stackrel{d}{=} \sum_i D_i^{(n)}$, i.e., when its distribution is infinitely divisible.*

Clearly, if ν is an infinitely divisible distribution, then it can be obtained as the limiting distribution of a sum of IID random variables $D_i^{(n)}$. (To converge, the individual terms must be going to zero in probability as n grows.) More remarkably, the converse is also true:

Proposition 144 (Limiting Distributions Are Infinitely Divisible) *If ν is the limiting distribution of a sequence of IID sums, then it is an infinitely-divisible distribution.*

PROOF: See, for instance, Kallenberg (2002, Theorem 15.12). □

Remark: This should *not* be obvious. If we take larger and larger sums of (centered, standardized) IID Bernoulli variables, we obtain a Gaussian as the limiting distribution, but at no time is the Gaussian the distribution of any of the finite sums. That is, the IID sums in the definition of infinite divisibility are not necessarily at all the same as those in the proposition.

Theorem 145 (Infinitely Divisible Distributions and Stationary Independent Increments) *If X is a process with stationary and independent increments, then for all t, $X(t) - X(0)$ is infinitely divisible.*

PROOF: See the remarks before the definition of infinite divisibility. □

Corollary 146 (Infinitely Divisible Distributions and Lévy Processes) *If ν is an infinitely divisible distribution, there exists a Lévy process where $X(1) \sim \nu$, and this determines the other distributions.*

PROOF: If $X(1) \sim \nu$, then by Eq. 11.34, $X(n) \sim \nu^{*n}$ for all integer n. Conversely, the distribution of $X(1/n)$ is also fixed, since $\nu = \mathcal{L}(X(1/n))^{*n}$, which means that the characteristic function of ν is equal to the n^{th} power of that of $X(1/n)$; inverting the latter then gives the desired distribution. Because increments are independent, $X(n+1/m) \stackrel{d}{=} X(n)+X(1/m)$, hence $\mathcal{L}(X(1))$ fixes the increment distribution for all rational time-intervals. By continuity, however, this also fixes it for time intervals of any length. Since the distribution of increments and the condition $X(0) = 0$ fix the finite-dimensional distributions, it remains only to show that the process is cadlag, which can be done by observing that it is continuous in probability, and then using Theorem 96. □

We have established a correspondence between the limiting distributions of IID sums, and processes with stationary independent increments. There are many possible applications of this correspondence; one is to reduce problems about limits of IID sums to problems about certain sorts of Markov process, and vice versa. Another is to relate discrete and continuous time processes. For $0 \leq t \leq 1$, set

$$X_t^{(n)} = \frac{1}{\sqrt{n}} \sum_{i=0}^{\lfloor nt \rfloor} Y_i \tag{11.36}$$

where $Y_0 = 0$ but otherwise the Y_i are IID with mean 0 and variance 1. Then each $X^{(n)}$ is a Markov process in discrete time, with stationary, independent increments, and its time-evolution operator is

$$K_n f(x) = \mathbf{E}\left[f(x + \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} Y_i) \right] \tag{11.37}$$

As n grows, the normalized sum approaches a standard Gaussian random variable, which suggests that in some sense $X^{(n)}$ should be approaching the Wiener process W. This is in fact true, but we will have to go deeper into the structure of the operators concerned before we can make it precise, as a first form of the *functional central limit theorem*. Just as the infinitely divisible distributions are the limits of IID sums, processes with stationary and independent increments are the limits of the corresponding random walks.

There is a yet further sense which the Gaussian distribution is a special kind of limiting distribution, which is reflected in another curious property of the Wiener process. Gaussian random variables are not only infinitely divisible, but they can be additively decomposed into more Gaussians. Distributions which can be infinitely divided into others of the same kind will be called "stable" (under convolution or averaging). "Of the same kind" isn't very mathematical; here is a more precise expression of the idea.

Definition 147 (Self-similarity) *A process is* self-similar *if, for all t, there exists a measurable $h(t)$, the* scaling function*, such that $h(t)X(1) \stackrel{d}{=} X(t)$.*

Definition 148 (Stable Distributions) *A distribution ν is* stable *if, for any Lévy process X where $X(1) \sim \nu$, X is self-similar.*

Theorem 149 (Scaling in Stable Lévy Processes) *In a self-similar Lévy process, the scaling function $h(t)$ must take the form t^α for some real α, the* index of stability*.*

PROOF: Exercise 11.7. □

It turns out that there are analogs to the functional central limit theorem for a broad range of self-similar processes, especially ones with stable increment distributions; Embrechts and Maejima (2002) is a good short introduction to this topic, and some of its statistical implications.

11.4 Exercises

Exercise 11.1 (Wiener Process with Constant Drift) *Consider a process $X(0)$ which, like the Wiener process, has $X(0) = 0$ and independent increments, but where $X(t_2) - X(t_1) \sim \mathcal{N}(a(t_2 - t_1), \sigma^2(t_2 - t_1))$. a is called the* drift rate *and σ^2 the* diffusion constant*. Show that $X(t)$ is a Markov process, following the argument for the standard Wiener process ($a = 0$, $\sigma^2 = 1$) above. Do such processes have continuous modifications for all (finite) choices of a and σ^2? If so, prove it; if not, give at least one counter-example.*

Exercise 11.2 (Perron-Frobenius Operators) *Verify that P defined in the section on the logistic map above is a Markov operator.*

Exercise 11.3 (Continuity of the Wiener Process) *Show that the Wiener process has continuous sample paths, using its finite-dimensional distributions and Theorem 97.*

Exercise 11.4 (Independent Increments with Respect to a Filtration) *Let X be adapted to a filtration $\{\mathcal{F}\}_t$. Then X has* independent increments with respect to $\{\mathcal{F}\}_t$ *when $X_t - X_s$ is independent of $\mathcal{F}_s$ for all $s \leq t$. Show that X is Markovian with respect to $\{\mathcal{F}\}_t$, by analogy with Theorem 141.*

Exercise 11.5 (Poisson Counting Process) *Consider the Poisson counting process N of Example 140.*

1. *Prove that N is continuous in probability.*
2. *Prove that N has a cadlag modification.* Hint*: Theorem 96 is one way, but there are others.*

Exercise 11.6 (Poisson Distribution is Infinitely Divisible) *Prove, directly from the definition, that every Poisson distribution is infinitely divisible. What is the corresponding stationary, independent-increment process?*

Exercise 11.7 (Self-Similarity in Lévy Processes) *Prove Theorem 149.*

Exercise 11.8 (Gaussian Stability) *Find the index of stability a standard Gaussian distribution.*

Exercise 11.9 (Poissonian Stability?) *Is the standard Poisson distribution stable? If so, prove it, and find the index of stability. If not, prove that it is not.*

Exercise 11.10 (Lamperti Transformation) *Prove* Lamperti's Theorem*: If Y is a strictly stationary process and $\alpha > 0$, then $X(t) = t^{\alpha} Y(\log t)$ is self-similar, and if X is self-similar with index of stability α, then $Y(t) = e^{-\alpha t} X(e^t)$ is strictly stationary. These operations are called the* Lamperti transformations*, after Lamperti (1962).*