Chapter 16

Convergence of Random Walks

This lecture examines the convergence of random walks to the Wiener process. This is very important both physically and statistically, and illustrates the utility of the theory of Feller processes.

Section 16.1 finds the semi-group of the Wiener process, shows it satisfies the Feller properties, and finds its generator.

Section 16.2 turns random walks into cadlag processes, and gives a fairly easy proof that they converge on the Wiener process.

16.1 The Wiener Process is Feller

Recall that the Wiener process $W(t)$ is defined by starting at the origin, by independent increments over non-overlapping intervals, by the Gaussian distribution of increments, and by continuity of sample paths (Examples 38 and 79). The process is homogeneous, and the transition kernels are (Section 11.2)

$$\mu_t(w_1, B) = \int_B dw_2 \frac{1}{\sqrt{2\pi t}} e^{-\frac{(w_2-w_1)^2}{2t}} \tag{16.1}$$

$$\frac{d\mu_t(w_1, w_2)}{d\lambda} = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(w_2-w_1)^2}{2t}} \tag{16.2}$$

where the second line gives the density of the transition kernel with respect to Lebesgue measure.

Since the kernels are known, we can write down the corresponding evolution operators:

$$K_t f(w_1) = \int dw_2 f(w_2) \frac{1}{\sqrt{2\pi t}} e^{-\frac{(w_2-w_1)^2}{2t}} \tag{16.3}$$

We saw in Section 11.2 that the kernels have the semi-group property, so (Lemma 121) the evolution operators do too.

Let's check that $\{K_t\}, t \geq 0$ is a Feller semi-group. The first Feller property is easier to check in its probabilistic form, that, for all t, $y \rightarrow x$ implies $W_y(t) \stackrel{d}{\rightarrow} W_x(t)$. The distribution of $W_x(t)$ is just $\mathcal{N}(x,t)$, and it is indeed true that $y \rightarrow x$ implies $\mathcal{N}(y,t) \rightarrow \mathcal{N}(x,t)$. The second Feller property can be checked in its semi-group form: as $t \rightarrow 0$, $\mu_t(w_1, B)$ approaches $\delta(w-w_1)$, so $\lim_{t\rightarrow 0} K_t f(x) = f(x)$. Thus, the Wiener process is a Feller process. This implies that it has cadlag sample paths (Theorem 193), but we already knew that, since we know it's continuous. What we did not know was that the Wiener process is not just Markov but strong Markov, which follows from Theorem 194.

To find the generator of $\{K_t\}, t \geq 0$, it will help to re-write it in an equivalent form, as

$$K_t f(w) = \mathbf{E}\left[f(w + Z\sqrt{t})\right] \tag{16.4}$$

where Z is an independent $\mathcal{N}(0,1)$ random variable. (We saw that this was equivalent in Section 11.2.) Now let's pick an $f \in C_0$ which is also twice continuously differentiable, i.e., $f \in C_0 \cap C^2$. Look at $K_t f(w) - f(w)$, and apply Taylor's theorem, expanding around w:

$$\begin{aligned} K_t f(w) - f(w) &= \mathbf{E}\left[f(w + Z\sqrt{t})\right] - f(w) &(16.5)\\ &= \mathbf{E}\left[f(w + Z\sqrt{t}) - f(w)\right] &(16.6)\\ &= \mathbf{E}\left[Z\sqrt{t}f'(w) + \frac{1}{2}tZ^2 f''(w) + R(Z\sqrt{t})\right] &(16.7)\\ &= \sqrt{t}f'(w)\mathbf{E}\left[Z\right] + t\frac{f''(w)}{2}\mathbf{E}\left[Z^2\right] + \mathbf{E}\left[R(Z\sqrt{t})\right] &(16.8) \end{aligned}$$

Recalling that $\mathbf{E}\left[Z\right] = 0$, $\mathbf{E}\left[Z^2\right] = 1$,

$$\lim_{t\downarrow 0} \frac{K_t f(w) - f(w)}{t} = \frac{1}{2}f''(w) + \lim_{t\downarrow 0} \frac{\mathbf{E}\left[R(Z\sqrt{t})\right]}{t} \tag{16.9}$$

So, we need to investigate the behavior of the remainder term $R(Z\sqrt{t})$.

We know from Taylor's theorem that

$$R(Z\sqrt{t}) = \frac{tZ^2}{2}\int_0^1 du\ f''(w + uZ\sqrt{t}) - f''(w) \tag{16.10}$$

$$\tag{16.11}$$

Since $f \in C_0 \cap C^2$, we know that $f'' \in C_0$. Therefore, f'' is uniformly continuous, and has a modulus of continuity,

$$m(f'', h) = \sup_{x,y:\ |x-y|\leq h} |f''(x) - f''(y)| \tag{16.12}$$

which goes to 0 as $h \downarrow 0$. Thus

$$
\begin{aligned}
\left|R(Z\sqrt{t})\right| &\leq \frac{tZ^2}{2} m(f'', Z\sqrt{t}) && (16.13)\\
\lim_{t\to 0} \frac{\left|R(Z\sqrt{t})\right|}{t} &\leq \lim_{t\to 0} \frac{Z^2 m(f'', Z\sqrt{t})}{2} && (16.14)\\
&= 0 && (16.15)
\end{aligned}
$$

Plugging back in to Equation 16.9,

$$
\begin{aligned}
Gf(w) &= \frac{1}{2} f''(w) + \lim_{t\downarrow 0} \frac{\mathbf{E}\left[R(Z\sqrt{t})\right]}{t} && (16.16)\\
&= \frac{1}{2} f''(w) && (16.17)
\end{aligned}
$$

That is, $G = \frac{1}{2}\frac{d^2}{dw^2}$, one half of the Laplacian. We have shown this only for $C_0 \cap C^2$, but this is clearly a linear subspace of C_0, and, since C^2 is dense in C, it is dense in C_0, i.e., this is a core for the generator. Hence the generator is really the extension of $\frac{1}{2}\frac{d^2}{dw^2}$ to the whole of C_0, but this is too cumbersome to repeat all the time, so we just say it's the Laplacian.

16.2 Convergence of Random Walks

Let $X_1, X_2, \ldots$ be a sequence of IID variables with mean 0 and variance 1. The random walk process S_n is then just $\sum_{i=1}^{n} X_i$. It is a discrete-time Markov process, and consequently also a strong Markov process. Imagine each step of the walk takes some time h, and imagine this time interval becoming smaller and smaller. Then, between any two times t_1 and t_2, the number of steps of the random walk will be about $\frac{t_2 - t_1}{h}$, which will go to infinity. The increment of the random walk from t_1 to t_2 will then be a sum of an increasingly large number of IID random variables, and by the central limit theorem will approach a Gaussian distribution. Moreover, if we look at the interval of time from t_2 to t_3, we will see another Gaussian, but all of the random-walk steps going into it will be independent of those going into our first interval. So, we expect that the random walk will in some sense come to look like the Wiener process, no matter what the exact distribution of the X_1. (We saw some of this in Section 11.3.) Let's consider this in more detail.

Definition 209 (Continuous-Time Random Walk (Cadlag)) *Let $X_1, X_2, \ldots$ be an IID sequence of real-valued random variables with mean 0 and variance 1. Define $S(m)$ as $X_1 + X_2 \ldots + X_m$, and $X_0 = 0$. The corresponding* continuous-time random walks *(CTRWs) are the processes*

$$
\begin{aligned}
Y_n(t) &\equiv \frac{1}{n^{1/2}} \sum_{i=0}^{\lfloor nt \rfloor} X_i && (16.18)\\
&= n^{-1/2} S(\lfloor nt \rfloor) && (16.19)
\end{aligned}
$$

Remark 1: You should verify (Exercise 16.3) that continuous-time random walks are inhomogeneous Markov processes with cadlag sample paths.

Remark 2: We will later see CTRWs with continuous sample paths, obtained by linear interpolation, but these, with their piece-wise constant sample paths, will do for now.

As has been hinted at repeatedly, random walks converge in distribution on the Wiener process. There are, broadly speaking, two ways to show this. One is to use the Feller process machinery of Chapter 15, and apply Corollary 206. The other is to directly manipulate the criteria for convergence of cadlag processes. Both lead to the same conclusion.

16.2.1 Approach Through Feller Processes

The processes Y_n are not homogeneously Markovian, though the discrete-time processes $n^{-1/2}S(m)$ are. Nonetheless, we can find the equivalent of their evolution operators, and show that they converge on the evolution operators of the Wiener process. First, let's establish a nice property of the increments of Y_n.

Lemma 210 (Increments of Random Walks) *For a continuous-time random walk, for all n,*

$$Y_n(t+h) - Y_n(t) = n^{-1/2}S'(\lfloor n(t+h)\rfloor - \lfloor nt\rfloor) \tag{16.20}$$

where $S'(m) \equiv \sum_{i=0}^{m} X'_i$ and the IID sequence X' is an independent copy of X.

Proof: By explicit calculation. For any n, for any t and any $h > 0$,

$$\begin{aligned} Y_n(t+h) &= \frac{1}{n^{1/2}} \sum_{i=0}^{\lfloor n(t+h)\rfloor} X_i && (16.21)\\ &= Y_n(t) + n^{-1/2} \sum_{i=\lfloor nt\rfloor+1}^{\lfloor n(t+h)\rfloor} X_i && (16.22)\\ &= Y_n(t) + n^{-1/2} \sum_{i=0}^{\lfloor n(t+h)\rfloor - \lfloor nt\rfloor} X'_i && (16.23) \end{aligned}$$

using, in the last line, the fact that the X_i are IID. Eq. 16.20 follows from the definition of $S(m)$. □

Lemma 211 (Continuous-Time Random Walks are Pseudo-Feller) *Every continuous-time random walk has the Feller properties (but is not a homogeneous Markov process).*

Proof: This is easiest using the process/probabilistic forms (Definition 177) of the Feller properties.

To see that the first property (Eq. 14.3) holds, we need the distribution of $Y_{n,y}(t)$, that is, the distribution of the state of Y_n at time t, when started from state y (rather than 0). By Lemma 210,

$$Y_{n,y}(t) \stackrel{d}{=} y + n^{-1/2}S'(\lfloor nt \rfloor) \tag{16.24}$$

Clearly, as $y \to x$, $y + n^{-1/2}S'(\lfloor nt \rfloor) \xrightarrow{d} x + n^{-1/2}S'(\lfloor nt \rfloor)$, so the first Feller property holds.

To see that the second Feller property holds, observe that, for each n, for all t, and for all ω, $Y_n(t+h,\omega) = Y_n(t,\omega)$ if $0 \geq h < 1/n$. This sure convergence implies almost-sure convergence, which implies convergence in probability. □

Lemma 212 (Evolution Operators of Random Walks) *The "evolution" (i.e., conditional expectation) operators of the random walk Y_n, $K_{n,t}$, are given by*

$$K_{n,t}f(y) = \mathbf{E}\left[f(y + Y'_n(t))\right] \tag{16.25}$$

where Y'_n is an independent copy of Y_n.

Proof: Substitute Lemma 210 into the definition of the evolution operator.

$$\begin{aligned} K_{n,h}f(y) &\equiv \mathbf{E}\left[f(Y_n(t+h))|Y_n(t) = y\right] &(16.26)\\ &= \mathbf{E}\left[f\left(Y_n(t+h) + Y_n(t) - Y_n(t)\right)|Y_n(t) = y\right] &(16.27)\\ &= \mathbf{E}\left[f\left(n^{-1/2}\S'(\lfloor nt \rfloor)\right) + Y_n(t)|Y_n(t) = y\right] &(16.28)\\ &= \mathbf{E}\left[f(y + n^{-1/2}S'(\lfloor nt \rfloor))\right] &(16.29)\end{aligned}$$

In words, the transition operator simply takes the expectation over the increments of the random walk, just as with the Wiener process. Finally, substitution of $Y'_n(t)$ for $n^{-1/2}S'(\lfloor nt \rfloor)$ is licensed by Eq. 16.19. □

Theorem 213 (Functional Central Limit Theorem (I)) $Y_n \xrightarrow{d} W$ *in* $\mathbf{D}$.

Proof: Apply Theorem 206. Clause (4) of the theorem says that if any of the other three clauses are satisfied, and $Y_n(0) \xrightarrow{d} W(0)$ in $\mathbb{R}$, then $Y_n \xrightarrow{d} W$ in $\mathbf{D}$. Clause (2) is that $K_{n,t} \to K_t$ for all $t > 0$. That is, for any $t > 0$, and $f \in C_0$, $K_{n,t}f \to K_t f$ as $n \to \infty$. Pick any such t and f and consider $K_{n,t}f$. By Lemma 212,

$$K_{n,t}f(y) = \mathbf{E}\left[f(y + n^{-1/2}S'(\lfloor nt \rfloor))\right] \tag{16.30}$$

As $n \to \infty$, $n^{-1/2} \to t^{1/2}\lfloor nt \rfloor^{-1/2}$. Since the X_i (and so the X'_i) have variance 1, the central limit theorem applies, and $\lfloor nt \rfloor^{-1/2}S'(\lfloor nt \rfloor) \xrightarrow{d} \mathcal{N}(0,1)$, say Z. Consequently $n^{-1/2}S'(\lfloor nt \rfloor \xrightarrow{d} \sqrt{t}Z$. Since $f \in C_0$, it is continuous and bounded, hence, by the definition of convergence in distribution,

$$\mathbf{E}\left[f\left(y + n^{-1/2}S'(\lfloor nt \rfloor)\right)\right] \to \mathbf{E}\left[f\left(y + \sqrt{t}Z\right)\right] \tag{16.31}$$

But $\mathbf{E}\left[f(y+\sqrt{t}Z)\right] = K_t f(y)$, the time-evolution operator of the Wiener process applied to f at y. Since the evolution operators of the random walks converge on those of the Wiener process, and since their initial conditions match, by the theorem $Y_n \xrightarrow{d} W$ in $\mathbf{D}$. □

16.2.2 Direct Approach

The alternate approach to the convergence of random walks works directly with the distributions, avoiding the Feller properties. It is not quite so slick, but provides a comparatively tractable example of how general results about convergence of stochastic processes go.

We want to find the limiting distribution of Y_n as $n \to \infty$. First of all, we should convince ourselves that a limit distribution exists. But this is not too hard. For any fixed t, $Y_n(t)$ approaches a Gaussian distribution by the central limit theorem. For any fixed finite collection of times $t_1 \leq t_2 \ldots \leq t_k$, $Y_n(t_1), Y_n(t_2), \ldots Y_n(t_k)$ approaches a limiting distribution if $Y_n(t_1), Y_n(t_2) - Y_n(t_1), \ldots Y_n(t_k) - Y_n(t_{k-1})$ does, but that again will be true by the (multivariate) central limit theorem. Since the limiting finite-dimensional distributions exist, some limiting distribution exists (via Theorem 23). It remains to convince ourselves that this limit is in $\mathbf{D}$, and to identify it.

Lemma 214 (Convergence of Random Walks in Finite-Dimensional Distribution) $Y_n \xrightarrow{fd} W$.

PROOF: For all n, $Y_n(0) = 0 = W(0)$. For any $t_2 > t_1$,

$$\begin{aligned}
\mathcal{L}\left(Y_n(t_2) - Y_n(t_1)\right) &= \mathcal{L}\left(\frac{1}{\sqrt{n}} \sum_{i=\lfloor nt_1 \rfloor}^{\lfloor nt_2 \rfloor} X_i\right) && (16.32)\\
&\xrightarrow{d} \mathcal{N}(0, t_2 - t_1) && (16.33)\\
&= \mathcal{L}\left(W(t_2) - W(t_1)\right) && (16.34)
\end{aligned}$$

Finally, for any three times $t_1 < t_2 < t_3$, $Y_n(t_3) - Y_n(t_2)$ and $Y_n(t_2) - Y_n(t_1)$ are independent for sufficiently large n, because they become sums of disjoint collections of independent random variables. The same applies to large groups of times. Thus, the limiting distribution of Y_n starts at the origin and has independent Gaussian increments. Since these properties determine the finite-dimensional distributions of the Wiener process, $Y_n \xrightarrow{fd} W$. □

Theorem 215 (Functional Central Limit Theorem (II)) $Y_n \xrightarrow{d} W$.

PROOF: By Proposition 200, it is enough to show that $Y_n \xrightarrow{fd} W$, and that any of the properties in Proposition 201 hold. The lemma took care of the finite-dimensional convergence, so we can turn to the second part. A sufficient condition is property (1) in the latter theorem, that $|Y_n(\tau_n + h_n) - Y_n(\tau_n)| \xrightarrow{P} 0$ for all finite optional times τ_n and any sequence of positive constants $h_n \to 0$.

$$\begin{aligned}
|Y_n(\tau_n + h_n) - Y_n(\tau_n)| &= n^{-1/2} \left|S(\lfloor n\tau_n + nh_n \rfloor) - S(\lfloor n\tau_n \rfloor)\right| && (16.35)\\
&\stackrel{d}{=} n^{-1/2} \left|S(\lfloor nh_n \rfloor) - S(0)\right| && (16.36)\\
&= n^{-1/2} \left|S(\lfloor nh_n \rfloor)\right| && (16.37)\\
&= n^{-1/2} \left|\sum_{i=0}^{\lfloor nh_n \rfloor} X_i\right| && (16.38)
\end{aligned}$$

To see that this converges in probability to zero, we will appeal to Chebyshev's inequality: if Z_i have common mean 0 and variance σ^2, then, for every positive ϵ,

$$\mathbb{P}\left(\left|\sum_{i=1}^{m} Z_i\right| > \epsilon\right) \leq \frac{m\sigma^2}{\epsilon^2} \tag{16.39}$$

Here we have $Z_i = X_i/\sqrt{n}$, so $\sigma^2 = 1/n$, and $m = \lfloor nh_n \rfloor$. Thus

$$\mathbb{P}\left(n^{-1/2} \left|S(\lfloor nh_n \rfloor)\right| > \epsilon\right) \leq \frac{\lfloor nh_n \rfloor}{n\epsilon^2} \tag{16.40}$$

As $0 \leq \lfloor nh_n \rfloor / n \leq h_n$, and $h_n \to 0$, the bounding probability must go to zero for every fixed ϵ. Hence $n^{-1/2} \left|S(\lfloor nh_n \rfloor)\right| \xrightarrow{P} 0$. □

16.2.3 Consequences of the Functional Central Limit Theorem

Corollary 216 (The Invariance Principle) *Let $X_1, X_2, \ldots$ be IID random variables with mean μ and variance σ^2. Then*

$$Y_n(t) \equiv \frac{1}{\sqrt{n}} \sum_{i=0}^{\lfloor nt \rfloor} \frac{X_i - \mu}{\sigma} \quad \xrightarrow{d} \quad W(t) \tag{16.41}$$

PROOF: $(X_i - \mu)/\sigma$ has mean 0 and variance 1, so Theorem 215 applies. □

This result is called "the invariance principle", because it says that the limiting distribution of the sequences of sums depends only on the mean and variance of the individual terms, and is consequently *invariant* under changes which leave those alone. Both this result and the previous one are known as the "functional central limit theorem", because convergence in distribution is the same as convergence of all bounded continuous *functionals* of the sample path. Another name is "Donsker's Theorem", which is sometimes associated however with the following corollary of Theorem 215.

Corollary 217 (Donsker's Theorem) *Let $Y_n(t)$ and $W(t)$ be as before, but restrict the index set T to the unit interval $[0, 1]$. Let f be any function from $\mathbf{D}([0, 1])$ to $\mathbb{R}$ which is measurable and a.s. continuous at W. Then $f(Y_n) \xrightarrow{d} f(W)$.*

PROOF: Exercise. □

This version is especially important for statistical purposes, as we'll see a bit later.

16.3 Exercises

Exercise 16.1 (Example 167 Revisited) *Go through all the details of Example 167.*

1. *Show that* $\mathcal{F}_t^X \subseteq \mathcal{F}_t^W$ *for all* t, *and that* $\left\{\mathcal{F}_t^X\right\} \subset \left\{\mathcal{F}_t^W\right\}$.
2. *Show that* $\tau = \inf_t X(t) = (0,0)$ *is a* $\left\{\mathcal{F}_t^X\right\}$*-optional time, and that it is finite with probability 1.*
3. *Show that* X *is Markov with respect to both its natural filtration and the natural filtration of the driving Wiener process.*
4. *Show that* X *is not strongly Markov at* τ.
5. *Which, if either, of the Feller properties does* X *have?*

Exercise 16.2 (Generator of the d-dimensional Wiener Process) *Consider a* d*-dimensional Wiener process, i.e., an* $\mathbb{R}^d$*-valued process where each coordinate is an independent Wiener process. Find the generator.*

Exercise 16.3 (Continuous-time random walks are Markovian) *Show that every continuous-time random walk (as per Definition 209) is an* inhomogeneous *Markov process, with cadlag sample paths.*

Exercise 16.4 (Donsker's Theorem) *Prove Donsker's Theorem (Corollary 217).*

Exercise 16.5 (Diffusion equation) *The partial differential equation*

$$\frac{1}{2}\frac{\partial^2 f(x,t)}{\partial x^2} = \frac{\partial f(x,t)}{\partial t}$$

is called the diffusion equation. *From our discussion of initial value problems in Chapter 12 (Corollary 159 and related material), it is clear that the function* $f(x,t)$ *solves the diffusion equation with initial condition* $f(x,0)$ *if and only if* $f(x,t) = K_t f(x,0)$, *where* K_t *is the evolution operator of the Wiener process.*

1. *Take* $f(x,0) = \left(2\pi 10^{-4}\right)^{-1/2} e^{-\frac{x^2}{2\cdot 10^{-4}}}$. $f(x,t)$ *can be found analytically; do so.*

2. *Estimate* $f(x, 10)$ *over the interval* $[-5, 5]$ *stochastically. Use the fact that* $K_t f(x) = \mathbf{E}[f(W(t))|W(0) = x]$, *and that random walks converge on the Wiener process. (Be careful that you scale your random walks the right way!) Give an indication of the error in this estimate.*

3. *Can you find an analytical form for* $f(x, t)$ *if* $f(x, 0) = \mathbf{1}_{[-0.5,0.5]}(x)$*?*

4. *Find* $f(x, 10)$*, with the new initial conditions, by numerical integration on the domain* $[-10, 10]$*, and compare it to a stochastic estimate.*

Exercise 16.6 (Functional CLT for Dependent Variables) *Let* X_i*,* $i = 1, 2, \ldots$*, be a weakly stationary but dependent sequence of real-valued random variables, with mean 0 and standard deviation 1. (Note that any weakly-stationary sequence with finite variance can be normalized into this form.) Let* Y_n *be the corresponding continuous-time random walk, i.e.,*

$$Y_n(t) = \frac{1}{n^{1/2}} \sum_{i=0}^{\lfloor nt \rfloor} X_i$$

Suppose that, despite their interdependence, the X_i *still obey the central limit theorem,*

$$\frac{1}{n^{1/2}} \sum_{i=1}^{n} X_i \xrightarrow{d} \mathcal{N}(0, 1)$$

Are these conditions enough to prove a functional central limit theorem, that $Y_n \xrightarrow{d} W$*? If so, prove it. If not, explain what the problem is, and suggest an additional sufficient condition on the* X_i*.*