Plan Fun with White Noise Part III

1. Gambler's Ruin and Martingales
2. The Poisson Process
3. Brownian Motion

Next Time: Stationary Processes, … then Markov Processes

Reading: –
Homework 2 due Thursday

Example 1. The Gambler's Ruin.

Let S_n be a simple random walk with probabilities p and $q = 1 - p$ of moving up and down respectively. Suppose that S_n represent the total wealth of a gambler at time n, so S_0 is her initial wealth (recall that S_0 is independent of the Xs). The gambler needs \$$m$ to pay a debt and has decided to gamble in subsequent best of \$1 each until her wealth reaches m (success) or 0 (ruin). Let T be the time at which either occurs and let R be the indicator that the gambler is ruined.

Find $r_k = \mathsf{E}(R \mid S_0 = k)$ and $t_k = \mathsf{E}(T \mid S_0 = k)$.

Method 1. Direct Computation.

Condition on X_1 and solve recurrence relation directly.

$$\begin{aligned}
r_k &= \mathsf{E}(R \mid S_0 = k) && (1)\\
&= \mathsf{E}\mathsf{E}(R \mid X_1, S_0 = k) && (2)\\
&= p\mathsf{E}(R \mid X_1 = 1, S_0 = k) + q\mathsf{E}(R \mid X_1 = -1, S_0 = k) && (3)\\
&= \begin{cases} pr_{k+1} + qr_{k-1} & \text{if } 1 \le k \le m-1 \\ 1 & \text{if } k = 0 \\ 0 & \text{if } k \ge m \text{ or } k < 0, \end{cases} && (4)
\end{aligned}$$

and by the same reasonin,

$$\begin{aligned}
t_k &= \mathsf{E}(T \mid S_0 = k) && (5)\\
&= \begin{cases} 1 + pt_{k+1} + qt_{k-1} & \text{if } 1 \le k \le m-1 \\ 0 & \text{if } k \ge m \text{ or } k \le 0. \end{cases} && (6)
\end{aligned}$$

Hence, for $1 \le k \le m-1$,

$$\Delta_{k+1} \equiv r_{k+1} - r_k = \frac{q}{p}(r_k - r_{k-1}) \equiv \frac{q}{p}\Delta_k. \tag{7}$$

So, $\Delta_k = (q/p)^{k-1}\Delta_1$, for $1 \le k \le m$, which telescopes to

$$r_k = 1 + \sum_{1 \le j \le k} \Delta_j. \tag{8}$$

Solving for r_k gives us:

$$r_k = \begin{cases} \frac{\left(\frac{q}{p}\right)^k - \left(\frac{q}{p}\right)^m}{1 - \left(\frac{q}{p}\right)^m} & \text{if } p \ne q \\ 1 - \frac{k}{m} & \text{if } p = q. \end{cases} \tag{9}$$

Similarly, for the t_ks, let $\Delta_k = p(t_k - t_{k-1})$ for $1 \le k \le m$, so $pt_k = \sum_{1\le j\le k} \Delta_j$. Then,

$$\Delta_{k+1} = \frac{q}{p}\Delta_k - 1 = \left(\frac{q}{p}\right)^k pt_1 - \left(1 + \frac{q}{p} + \cdots + \left(\frac{q}{p}\right)^{k-1}\right). \tag{10}$$

For $p = q$, $\Delta_k = t_1/2 - (k-1)$, and using $t_m = 0$, yields $t_k = k(m-k)$. For $p \neq q$,

$$\Delta_k = \left(\frac{q}{p}\right)^{k-1} pt_1 - \frac{\left(\frac{q}{p}\right)^{k-1} - 1}{\frac{q}{p} - 1} \tag{11}$$

$$= \left(\frac{q}{p}\right)^{k-1} pt_1 + \frac{p}{q-p} - \frac{p}{q-p}\left(\frac{q}{p}\right)^{k-1}, \tag{12}$$

and thus

$$t_k = \frac{p}{q-p}\left(\left(\frac{q}{p}\right)^k - 1\right)t_1 + \frac{k}{q-p} - \frac{p}{(q-p)^2}\left(\left(\frac{q}{p}\right)^k - 1\right). \tag{13}$$

Using $t_m = 0$, we can solve this for t_k.

Method 2. Generating Functions.

Consider a more general recurrence of the form:

$$cg_{k+1} + bg_k + ag_{k-1} = u_k \qquad \text{if } 1 \le k \le m-1 \tag{14}$$

$$g_0 = u_0 \tag{15}$$

$$g_m = u_m \tag{16}$$

Form the generating functions

$$G(z) = \sum_{k=1}^{m-1} g_k z^k \qquad U(z) = \sum_{k=1}^{m-1} u_k z_k. \tag{17}$$

Now multiply the recurrence above by z^{k+1} and sum:

$$zU(z) = c\sum_{k=1}^{m-1} g_{k+1}z^{k+1} + bzG(z) + az^2 \sum_{k=1}^{m-1} g_{k-1}z^{k-1} \tag{18}$$

$$= c\sum_{k=2}^{m} g_k z^k + bzG(z) + az^2 \sum_{k=0}^{m-2} g_k z^k \tag{19}$$

$$= c(G(z) + u_m z^m - g_1 z) + bzG(z) + az^2(G(z) + u_0 - g_{m-1}z^{m-1}) \tag{20}$$

$$= G(z)(az^2 + bz + c) - z(ag_{m-1}z^m - cu_m z^{m-1} - au_0 z + cg_1). \tag{21}$$

Hence,

$$G(z)(az^2 + bz + c) = zP(z; g_1, g_{m-1}), \tag{22}$$

for a polynomial

$$P(z; g_1, g_{m-1}) = U(z) + ag_{m-1}z^m - cu_m z^{m-1} - au_0 z + cg_1 \tag{23}$$

$$= cg_1 + (u_1 - au_0)z + u_2 z^2 + \cdots u_{m-2} z^{m-2} + (u_{m-1} - cu_m) z^{m-1} + ag_{m-1} z^m. \quad (24)$$

We have only two unknowns.

Suppose the polynomial $az^2 + bz + c$ has two distinct, non-zero roots $1/\rho_1$ and $1/\rho_2$. Then we proceed in two steps. First, the unknowns g_1 and g_{m-1} satisfy the two equations

$$P(1/\rho_1; g_1, g_{m-1}) = 0 \quad (25)$$

$$P(1/\rho_2; g_1, g_{m-1}) = 0, \quad (26)$$

which are easily solvable. Second, $az^2+bz+c = a(z-1/\rho_2)(z-1/\rho_1) = a/(\rho_1\rho_2)(1-\rho_1 z)(1-\rho_2 z) = c(1-\rho_1 z)(1-\rho_2 z)$. We have

$$\frac{1}{(1-\rho_1 z)(1-\rho_2 z)} = \frac{A}{1-\rho_1 z} - \frac{B}{1-\rho_2 z}, \quad (27)$$

where $A = \rho_1/(\rho_1 - \rho_2)$ and $B = \rho_2/(\rho_1 - \rho_2)$. Then,

$$G(z) = \frac{1}{c(\rho_1 - \rho_2)} P(z) \left(\frac{\rho_1 z}{1-\rho_1 z} - \frac{\rho_2 z}{1-\rho_2 z} \right) \quad (28)$$

$$= \frac{1}{c(\rho_1 - \rho_2)} P(z) \sum_{k=1}^{\infty} (\rho_1^k - \rho_2^k) z^k. \quad (29)$$

Note that $|c(\rho_1 - \rho_2)| = \sqrt{b^2 - 4ac}$. We can then just read the coefficients off by convolution. (What to do in the double or zero root case is left for homework.)

Now let's apply this to the Gambler's ruin. We have $a = -q$, $b = 1$, and $c = -p$, so $\rho_1 = 1$, $\rho_2 = q/p$. For the r_ks, $u_0 = 1$ and $u_k = 0$ for $k > 0$. For the t_ks, $u_0 = u_m = 0$ and $u_k = 1$ for $0 < k < m$. A bit messier in both cases, but we get the same answers as above.

Method 3. Martingales.

Consider now the symmetric ($p = q = 1/2$) case.

Let $\mathcal{F}_n = \sigma(S_0, X_1, \ldots, X_n)$ be the *history* of the process up to time n.

Notice that for every n.

$$\mathsf{E}(S_{n+1} \mid \mathcal{F}_n) = \mathsf{E}(S_n + X_{n+1} \mid \mathcal{F}_n) \quad (30)$$

$$= S_n + \mathsf{E}(X_n) \quad (31)$$

$$= S_n \quad (32)$$

A process (S_n) with this property is called a *martingale*. In particular, note that $\mathsf{E}(S_n) = \mathsf{E}(S_0)$ for all n.

For each n, $1\{T = n\}$ is determined by $S_0, X_1, \ldots, X_n$ and hence is $\mathcal{F}_n$ measurable. Such a random variable is called a *stopping time* of the process.

When S_n is a martingale and T is a stopping time of the process, we will see that

$$\mathsf{E}(S_T \mid S_0) = S_0.$$

In the Gambler's ruin problem:

$$k = \mathsf{E}(S_T \mid S_0 = k) \tag{33}$$
$$= m(1 - r_k) + 0 \cdot r_k. \tag{34}$$

Hence, $r_k = (m - k)/m$.

Consider now $M_n = S_n^2 - n$:

$$\mathsf{E}(M_{n+1} \mid \mathcal{F}_n) = \mathsf{E}((S_n + X_{n+1})^2 - (n + 1) \mid \mathcal{F}_n) \tag{35}$$
$$= S_n^2 - n + \mathsf{E}X_{n+1}^2 - 1 + S_n\mathsf{E}(X_{n+1} \mid \mathcal{F}_n) \tag{36}$$
$$= M_n. \tag{37}$$

So, M_n is also a martingale.

Thus,

$$k^2 = \mathsf{E}(M_T \mid S_0 = k) \tag{38}$$
$$= (m^2 - t_k)(1 - r_k) - t_k r_k \tag{39}$$
$$= m^2(1 - r_k) - t_k \tag{40}$$

so,

$$t_k = m^2 - k^2 + -m^2 r_k \tag{41}$$
$$= (m + k)(m - k) - m(m - k) \tag{42}$$
$$= k(m - k). \tag{43}$$

Example 2. Consider a random walk with $S_0 = 0$ and X_i IID Exponential$\langle\lambda\rangle$. (Note: If $W_i = X_i - \lambda$, $S_n = \sum_i W_i + n\lambda$, a sum of white noise plus drift.)

Then, S_n has a Gamma$\langle n, \lambda\rangle$ distribution.

Define $N_t = \max\{n \geq 1 \text{ such that } S_n \leq t\}$. We have $N_0 = 0$ and the following.

$$\mathsf{P}\{N_t = k\} = \mathsf{P}\{S_k \leq t, S_{k+1} > t\} \tag{44}$$
$$= \mathsf{E}\mathsf{P}\{S_k \leq t, S_{k+1} > t \mid S_k\} \tag{45}$$
$$= \int_0^t \mathsf{P}\{X_{k+1} > t - u\}\, \mathsf{P}\{S_k \text{ near } u\} \tag{46}$$
$$= \int_0^t e^{-\lambda(t-u)} \frac{\lambda^k u^{k-1}}{\Gamma(k)} e^{-\lambda u}\, du \tag{47}$$
$$= e^{-\lambda t} \frac{\lambda^k}{k!}\, k \int_0^t u^{k-1}\, du \tag{48}$$
$$= e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \tag{49}$$

so N_t has a Poisson$\langle\lambda t\rangle$ distribution.

Now consider $s < t < s' < t'$, two disjoint time intervals. What is the distribution of $N_t - N_s$?

$$\begin{aligned}
&\mathsf{P}\{N_s \ge j, N_t - N_s \ge k\} \\
&\qquad = \mathsf{P}\{S_j \le s, S_{k+N_s} \le t\} && (50)\\
&\qquad = \mathsf{P}\{S_j \le s, S_{k+N_s} - S_{N_s} \le t - s\} && (51)\\
&\qquad = \mathsf{E}\,\mathsf{P}\{S_j \le s, S_{k+N_s} - S_{N_s} \le t - s \mid N_s\} && (52)\\
&\qquad = \sum_{n=0}^{\infty} \mathsf{P}\{S_j \le s, S_{k+n} - S_n \le t - s \mid N_s = n\}\,\mathsf{P}\{N_s = n\} && (53)\\
&\qquad = \sum_{n=j}^{\infty} \mathsf{P}\{S_{k+n} - S_n \le t - s \mid N_s = n\}\,\mathsf{P}\{N_s = n\} && (54)\\
&\qquad = \sum_{n=j}^{\infty} \mathsf{P}\{S_{k+n} - S_n \le t - s \mid S_n \le s, X_{n+1} > s - S_n\}\,\mathsf{P}\{N_s = n\} && (55)\\
&\qquad = \sum_{n=j}^{\infty} \mathsf{P}\{S_{k+n} - S_n \le t - s \mid X_{n+1} > s - S_n\}\,\mathsf{P}\{N_s = n\} && (56)\\
&\qquad = \mathsf{P}\{N_s \ge j\}\,\mathsf{P}\{N_t - N_s \ge k\}\,. && (57)
\end{aligned}$$

That last step requires an argument. It follows from the relation $\mathsf{P}\{X_i > t + u \mid X_i > u\} = \mathsf{P}\{X_i > t\}$ for an Exponential distribution.

Or we could do it more messily:

$$\begin{aligned}
&\mathsf{P}\{N_t - N_s = k, N_s = j\} \\
&\quad = \mathsf{P}\{S_j \le s, S_{j+1} > s, S_{j+k} \le t, S_{j+k+1} > t\} && (58)\\
&\quad = \mathsf{P}\left\{S_j \le s, X_{j+1} > s - S_j, \sum_{i=j+1}^{j+k} X_i \le t - S_j, \sum_{i=j+1}^{j+k+1} X_i > t - S_j\right\} && (59)\\
&\quad = \mathsf{E}\,\mathsf{P}\left\{S_j \le s, X_{j+1} > s - S_j, \sum_{i=j+1}^{j+k} X_i \le t - S_j, \sum_{i=j+1}^{j+k+1} X_i > t - S_j \;\middle|\; S_j\right\} && (60)\\
&\quad = \mathsf{E}\,\mathsf{P}\left\{S_j \le s, X_{j+1} > s - S_j, \sum_{i=j+1}^{j+k} X_i \le t - s + s - S_j, \sum_{i=j+1}^{j+k+1} X_i > t - s + s - S_j \;\middle|\; S_j\right\} && (61)\\
&\quad = \int_0^s du\, f_{S_j}(u)\,\mathsf{P}\left\{X_{j+1} > s - u, \sum_{i=j+1}^{j+k} X_i \le t - s + s - u, \sum_{i=j+1}^{j+k+1} X_i > t - s + s - u\right\} && (62)\\
&\quad = \int_0^s du\, f_{S_j}(u)\,\cdot && (63)\\
&\qquad\quad \mathsf{P}\left\{X_{j+1} > s - u, X_{j+1} - (s-u) + \sum_{i=j+2}^{j+k} X_i \le t - s, X_{j+1} - (s-u) + \sum_{i=j+2}^{j+k+1} X_i > t - s\right\} \\
&\quad = \int_0^s du\, f_{S_j}(u)\,\mathsf{P}\{X_{j+1} > s - u\}\,\cdot && (64)
\end{aligned}$$

$$\mathsf{P}\left\{ X_{j+1} - (s-u) + \sum_{i=j+2}^{j+k} X_i \leq t-s, X_{j+1} - (s-u) + \sum_{i=j+2}^{j+k+1} X_i > t-s \,\middle|\, X_{j+1} > s-u \right\}$$

$$= \mathsf{P}\{S_j \leq s, S_{j+1} > s\}\,\mathsf{P}\{\dots\} \tag{65}$$

$$= \mathsf{P}\{N_s = j\}\,\mathsf{P}\{N_{t-s} = k\}\,. \tag{66}$$

The third- and second-to-last inequalities here use the same fact about the Exponential distribution.

By the same trick, we can see that $N_t - N_s$ and N_{t-s} have the same distribution. But using independence directly and the fact that

$$G(z) = \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} z^n = e^{\lambda(z-1)}, \tag{67}$$

we can show this directly. We have that $N_t = N_s + (N_t - N_s)$ where the latter two terms are independent. Hence,

$$G_{N_t}(z) = G_{N_s}(z) G_{N_t - N_s}(z) \tag{68}$$

$$e^{\lambda t(z-1)} = e^{\lambda s(z-1)} G_{N_t - N_s}(z) \tag{69}$$

and thus,

$$G_{N_t - N_s}(z) = e^{\lambda(t-s)(z-1)}. \tag{70}$$

Hence, $N_t - N_s$ has a Poisson$\langle \lambda(t-s) \rangle$ distribution and is independent of $N_s - N_0$. The same argument above shows that $N_t - N_s$ and $N_{t'} - N_{s'}$ are independent and similarly for any finite collection $0 \leq s_1 < t_1 < s_2 < t_2 < \cdots < s_m < t_m$ Thus, from a white noise random walk, we get the following process.

Process 3. Let $(N_t)_{t \geq 0}$ be a process with the following properties:

1. $N_0 = 0$
2. For $0 \leq s < t$, $N_t - N_s$ has a Poisson$\langle \lambda(t-s) \rangle$ distribution.
3. For $0 \leq s_1 < t_1 < s_2 < t_2 < \cdots < s_m < t_m$, the random variables $N_{t_i} - N_{s_i}$ are independent.

This is called a (homogeneous) *Poisson process* with rate λ.

Definition 4. Let $L^2(0,1)$ be (up to some formal details) the set of functions g on $(0,1)$ such that $\int_0^1 g^2 < \infty$.

A complete, orthonormal basis (ψ_n) for $L^2(0,1)$ is a countable subset of $L^2(0,1)$ such that $\int \psi_j \psi_k = \delta_{jk}$ and for any $g \in \mathcal{L}^2(0,1)$, we can find $c_n = \int g \psi_n$ such that

$$\int \left(g - \sum_{k=0}^{n} c_k \psi_k \right)^2 \to 0, \tag{71}$$

as $n \to \infty$.

Example 5. Define a process $(\xi_t)_{t\geq 0}$ as follows. Let $(A_n)_{n\geq 0}$ be a standard normal white noise Process, i.e., A_n are IID Normal$\langle 0,1\rangle$. Define

$$\xi_t(\omega) = \sum_{n=0}^{\infty} A_n(\omega)\psi_n(t), \tag{72}$$

for a particular complete orthonormal basis $(\psi_)$.

Taking some liberties (we'll see a more formal derivation another time), we will think of ξ_t as a continuous white noise process meaning that in some formal sense we should have $\mathsf{E}\xi_t = 0$ and $\mathsf{E}\xi_s\xi_t = \delta(s-t)$. Arguing loosely, this works:

$$\mathsf{E}\xi_t = \sum_n \mathsf{E}A_n\psi_n(t) = 0 \tag{73}$$

$$\mathsf{E}\xi_s\xi_t = \sum_{n,m} \mathsf{E}A_nA_m\psi_n(t)\psi_m(s) \tag{74}$$

$$= \sum_n \psi_n(t)\psi_n(s), \tag{75}$$

which "makes sense" when $s \neq t$ because

$$\int dt\, \mathsf{E}\xi_s\xi_t = \sum_n \psi_n(s) \int dt\, \psi_n(t) = 0. \tag{76}$$

But don't take that last calculation too seriously.

Now, just as we got a random walk process by taking cumulative sums of a discrete white noise process, we can see what we get when we take cumulative *integrals* of a *continuous* white noise process.

Define

$$W_t = \int_0^t \xi_s\, ds = \sum_{n=0}^{\infty} A_n \int_0^t \psi_n(s)\, ds, \tag{77}$$

where we choose a specific basis ψ_n.

Definition 6. Define $H = 1_{(0,1/2]} - 1_{(1/2,1]}$. Then let

$$H_{jk}(t) = 2^{j/2}H(2^jt-k). \tag{78}$$

Then, $F = 1_{(0,1]}$ and H_{jk} for $j \geq 0$, $k = 0,\ldots,2^j-1$ form a complete orthonormal basis for $L^2(0,1)$. It is called the *Haar basis*.

To see this note that $\int F^2 = 1$, $\int H_{jk}H_{j'k'} = \delta_{jj'}\delta_{kk'}$, and $\int H_{jk} = 0$. And we have that $\alpha F + \sum_{j=0}^{J-1}\sum_{k=0}^{2^j-1}\beta_{jk}H_{jk}$ gives a representation for all piecewise constant functions on dyadic intervals of length 2^{-J}. (We'll have a cool martingale proof of this another time.)

Example cont'd 7. Now order the Haar functions $(F, H_{00}, H_{10}, H_{11}, H_{20}, H_{21}, H_{22}, H_{23}, \ldots)$ and label these as ψ_n for $n \geq 0$. (For $2^j \leq n < 2^{j+1}$ and $j \in \mathbb{Z}_+$, take $k = n - 2^j$ and let $H_n \equiv H_{jk}$.)

For $n \geq 1$,

$$\int_0^t H_n(s)\, ds \equiv S_n(t), \tag{79}$$

called the Schauder function.

It follows that

$$W_t = \sum_{n \geq 0} A_n S_n(t). \tag{80}$$

Lemma 1. If a sequence (a_k) satisfies $|a_k| = O(k^\gamma)$ for $0 \leq \gamma < 1/2$, we have that $\sum_{n\geq 0} a_n S_n(t)$ converges uniformly on $(0, 1)$.

Lemma 2. A standard normal white noise sequence A_n satisfies $|A_n| = O(\sqrt{\log n})$ with probability 1.

It follows that W_t exists as a random function with probability 1.

Moreover, *Lemma 3*: If $0 \leq s, t \leq 1$,

$$\sum_{n \geq 0} S_n(s) S_n(t) = \min(s, t). \tag{81}$$

To see this, let $\phi_s = 1_{[0,s]}$. Then, if $s \leq t$,

$$s = \int_0^1 \phi_t \phi_s = \sum_{n \geq 0} a_n b_n \tag{82}$$

where

$$a_n = \int_0^1 \phi_t H_n = S_n(t) \tag{83}$$

$$b_n = \int_0^1 \phi_s H_n = S_n(s). \tag{84}$$

So, we get the following:

$$\mathsf{E} W_t = \sum_{n \geq 0} \mathsf{E} A_n S_n(t) = 0 \tag{85}$$

$$\mathsf{E} W_t^2 = \sum_{n \geq 0} \mathsf{E} A_n^2 S_n(t) = t \tag{86}$$

$$\mathsf{E} W_s W_t = \sum_{n,m \geq 0} \mathsf{E} A_n A_m S_n(t) S_m(s) = \min(s, t) \tag{87}$$

and for $u < s < t$,

$$\mathsf{E}(W_t - W_s) W_u = \sum_{n,m \geq 0} \mathsf{E} A_n A_m (S_n(t) - S_n(s)) S_m(u) = \min(t, u) - \min(s, u) = 0. \tag{88}$$

Moreover, using the characteristic generating functions with $s \leq t$,

$$\mathsf{E} e^{i\lambda(W_t - W_s)} = \mathsf{E} e^{i\lambda \sum_n A_n (S_n(t) - S_n(s))} \tag{89}$$

$$= \prod_{n=0}^{\infty} \mathsf{E} e^{i\lambda A_n (S_n(t) - S_n(s))} \tag{90}$$

$$= \prod_{n=0}^{\infty} e^{-\frac{\lambda^2}{2}(S_n(t)-S_n(s))^2} \tag{91}$$

$$= e^{-\frac{\lambda^2}{2}\sum_n (S_n(t)-S_n(s))^2} \tag{92}$$

$$= e^{-\frac{\lambda^2}{2}(t-2s+s)} \tag{93}$$

$$= e^{-\frac{\lambda^2}{2}(t-s)}, \tag{94}$$

using normality of the A_ns. Hence, $W_t - W_s$ has a Normal$\langle 0, t-s\rangle$ distribution. We get the following process.

Process 8. Thus, derived from a white noise process, we get a process $(W_t)_{t\geq 0}$ with the following properties:

1. $W_0 = 0$
2. For $0 \leq s < t$, $W_t - W_s$ has a Normal$\langle 0, t-s\rangle$ distribution.
3. For $0 \leq s_1 < t_1 < s_2 < t_2 < \cdots$, the random variables $W_{t_i} - W_{s_i}$ are independent.

We call this process a *Weiner process* or equivalently, a *Brownian Motion*.