Applied Stochastic Processes

Lecture notes for Math 607

MATH 607 • Applied Stochastic Processes
Fall 2018 — Dr Peter Ralph

Printed on 21 October 2018.

Department of Mathematics
Fenton Hall
University of Oregon
Eugene, OR 97403-1222 USA

Abstract. These are notes.

2010 *Mathematics Subject Classification.* 60G05.

Key words and phrases. stochastic calculus.

Part I

Applied Stochastic Processes

Contents

Fill Staley

⋆ ⋆ ⋆ ⋆ ⋆ ⋆ ⋆

Lecture 1
24 September 2018

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Fill Staley

⋆ ⋆ ⋆ ⋆ ⋆ ⋆ ⋆

Lecture 2
26 September 2018

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lecture 3
28 September 2018

⋆ ⋆ ⋆ ⋆ ⋆ ⋆ ⋆

Joe

§ 1 | A Gaussian Process on $L^2(\mathbb{R})$

We will construct a process on $L^2(\mathbb{R})$ by taking a limit of a simpler process on the space of sequences

$$\ell^2(\mathbb{R}) \coloneqq \left\{ (a_k)_{k\in\mathbb{Z}} \colon \sum_{k\in\mathbb{Z}} |a_k|^2 < \infty,\, a_k \in \mathbb{R} \right\} .$$

Given a set of independent random variables $\{X_k\}_{k\in\mathbb{Z}}$ satisfying $X_k \sim N(0,1)$ (think of each X_k as "noise" at the integer k), define $Z(a) \coloneqq \sum_{k\in\mathbb{Z}} a_k X_k$ for $a = (a_k) \in \ell^2(\mathbb{R})$. The collection $\{Z(a)\}_{a\in\ell^2(\mathbb{R})}$ is a Gaussian process that is centered (ie, $\mathbb{E}[Z(a)] = 0$) and satisfies

$$\operatorname{var}[Z(a)] = \operatorname{cov}\left[\sum_{k\in\mathbb{Z}} a_k X_k, \sum_{l\in\mathbb{Z}} a_l X_l \right] = \sum_{k,l\in\mathbb{Z}} a_k a_l \operatorname{cov}[X_k, X_l] = \sum_{k\in\mathbb{Z}} a_k^2$$

for each $a \in \ell^2(\mathbb{R})$.

Now recall $L^2(\mathbb{R}) \coloneqq \{f \colon \mathbb{R} \to \mathbb{R} \colon \int_{-\infty}^{\infty} |f(x)|^2 \, dx < \infty\}$. Given $f \in L^2(\mathbb{R})$, define a sequence of random variables $(Z^{(N)}(f))_{N\in\mathbb{N}}$ by

$$Z^{(N)}(f) \coloneqq \frac{1}{\sqrt{N}} \sum_{k\in\mathbb{Z}} f\left(\frac{k}{N}\right) X_k .$$

It can be shown that $Z^{(N)}(f)$ converges in distribution to a random variable which we will call $Z(f)$. It can be also shown that the collection $\{Z(f)\}_{f\in L^2(\mathbb{R})}$ is a centered Gaussian process satisfying $\operatorname{cov}[Z(f), Z(g)] = \int_{-\infty}^{\infty} f(x)g(x)\, dx$. It turns out that these properties (centered Gaussian and satisfying the above covariance formula) characterize $\{Z(f)\}_{f\in L^2(\mathbb{R})}$, so we might as well define it this way:

> Definition. The collection $\{Z(f)\}_{f\in L^2(\mathbb{R})}$ is **the centered Gaussian process on** $L^2(\mathbb{R})$ satisfying
>
> $$\operatorname{cov}[Z(f), Z(g)] \coloneqq \int_{-\infty}^{\infty} f(x)g(x)\, dx .$$

Note that $Z(af + bg) = aZ(f) + bZ(g)$ for all $a, b \in \mathbb{R}$ and all $f, g \in L^2(\mathbb{R})$, so $f \mapsto Z(f)$ defines a linear map from $L^2(\mathbb{R})$ to the space of random variables. The random variable $Z(f)$ is called the *stochastic integral* of f, denoted

$$Z(f) \eqqcolon \int_{-\infty}^{\infty} f(t)\, dW_t ,$$

and the integrator dW_t is interpreted as "white noise" that weights the value of $f(t)$. Assuming it is known that $\operatorname{cov}[dW_s, dW_t] = \delta_s(t)\, ds\, dt$, where δ_s is the Dirac δ-functional centered at s, we can recover the covariance formula using stochastic integral notation:

$$\begin{aligned}\operatorname{cov}[Z(f), Z(g)] &= \operatorname{cov}\left[\int_{-\infty}^{\infty} f(s)\, dW_s, \int_{-\infty}^{\infty} g(t)\, dW_t\right] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s)g(t)\ \operatorname{cov}[dW_s, dW_t] = \int_{-\infty}^{\infty} f(t)g(t)\, dt.\end{aligned}$$

The following examples show that some stochastic processes can be defined as stochastic integrals.

1.1 Example (*Brownian motion*) If we define $B_t := Z(\mathbf{1}_{[0,t)}) = \int_0^t dW_s$, then the process $\{B_t\}_{t\in[0,\infty)}$ is a Brownian motion. Indeed,

(i) B_t is Gaussian,

(ii) $\operatorname{cov}[B_s, B_t] = \int_{-\infty}^{\infty} \mathbf{1}_{[0,s)}(u)\mathbf{1}_{[0,t)}(u)\, du = \int_0^{\min\{s,t\}} du = \min\{s,t\}$ (and hence $\operatorname{var}[B_t] = t$),

(iii) if $u < v \le s < t$ then

$$\operatorname{cov}[B_t - B_s, B_v - B_u] = \int_{-\infty}^{\infty} \mathbf{1}_{[s,t)}(x)\mathbf{1}_{[u,v)}(x)\, dx = \int_{-\infty}^{\infty} \mathbf{1}_{[s,t)\cap[u,v)}(x)\, dx = 0\,.$$

•

1.2 Example (*Energy*) Define a process on $\{Y_t\}_{t\in\mathbb{R}}$ by

$$Y_t := Z\left(e^{-(t-s)}\mathbf{1}_{(-\infty,t]}(s)\right) = \int_{-\infty}^{\infty} e^{-(t-s)}\, dW_s\,.$$

Then

(i) Y_t is Gaussian,

(ii) for $s \le t$ we have

$$\operatorname{cov}[Y_s, Y_t] = \int_{-\infty}^{\infty} e^{-(s-u)}\mathbf{1}_{(-\infty,s]}(u)^2 e^{-(t-u)}\mathbf{1}_{(-\infty,t]}(u)^2\, du = e^{-s-t}\int_{-\infty}^{s} e^{2u}\, du = \frac{e^{-(t-s)}}{2}\,,$$

and hence $\operatorname{var}[Y_t] = \frac{1}{2}$.

•

Lecture 4
01 October 2018

⋆ ⋆ ⋆ ⋆ ⋆ ⋆ ⋆

Joe

§ 2 | A Different Construction of $Z(f)$

Definition. Define the *Haar basis* $\{h_{n,k}\}_{n,k\in\mathbb{N}}$ for $L^2([0,1])$ by $h_{0,0}\equiv 1$ and

$$h_{n,k}(t) := \begin{cases} 2^{(n-1)/2} & \text{if } 2^{-n}(2^k) \le t < 2^{-n}(2k+1)\,, \\ -2^{(n-1)/2} & \text{if } 2^{-n}(2k+1) \le t < 2^{-n}(2k+2)\,, \\ 0 & \text{otherwise} \end{cases}$$

if $n \ge 1$ and $0 \le k \le 2^{n-1}$ (otherwise set $h_{n,k}\equiv 0$).

It is easily verified that $\{h_{n,k}\}$ is an orthonormal set in $L^2([0,1])$, ie,

$$\int_0^1 h_{n,k}(t)h_{m,\ell}(t)\,dt = \begin{cases} 1 & \text{if } n=m \text{ and } k=\ell, \\ 0 & \text{otherwise} \end{cases}$$

Now suppose $\{Z_{n,k}\}$ is a collection of independent $N(0,1)$ variables and $f \in L^2([0,1])$ expands as

$$f(t) = \sum_{n=0}^{\infty}\sum_{k=0}^{2^{n-1}} a_{n,k}h_{n,k}(t)\,, \quad \text{where} \quad a_{n,k} = \int_0^1 f(t)h_{n,k}(t)\,dt\,.$$

The random variable $Z(f)$ defined by

$$Z(f) := \sum_{n=0}^{\infty}\sum_{k=0}^{2^{n-1}} a_{n,k}Z_{n,k}$$

gives a Gaussian process $\{Z(f)\}$ for which

$$\operatorname{cov}[Z(f),Z(g)] = \sum_{n=0}^{\infty}\sum_{k=0}^{2^{n-1}}\sum_{m=0}^{\infty}\sum_{\ell=0}^{2^{m-1}} a_{n,k}b_{m,\ell}\operatorname{cov}[Z_{n,k},Z_{m,k}] = \sum_{n=0}^{\infty}\sum_{k=0}^{2^{n-1}} a_{n,k}b_{n,k}\,.$$

In particular we have

$$\operatorname{var}[Z(f)] = \sum_{n=0}^{\infty}\sum_{k=0}^{2^{n-1}} a_{n,k}^2 = \int_0^1 |f(t)|^2\,dt\,,$$

where the latter equality is given by Parseval's identity (see page 85 of Rudin's *Real & Complex Analysis*). In the following example we see that Brownian motion can again be constructed by applying Z to a particular collection of functions.

2.1 Example (*Lévy's Construction*) Setting $B_t = Z(\mathbf{1}_{[0,t)})$ for $t\in[0,1]$, we have

$$a_{n,k} = \int_0^1 mathbf1_{[0,t)}h_{n,k}(s)\,ds = \int_0^t h_{n,k}(s)\,ds$$

and

$$B_t = \sum_{n=0}^{\infty}\sum_{k=0}^{2^{n-1}} Z_{n,k}\int_0^t h_{n,k}(s)\,ds.$$

Since Z is linear (with respect to f), for $s\le t$ we have

$$\operatorname{cov}[B_t - B_s] = \operatorname{cov}[Z(\mathbf{1}_{[0,t)}) - Z(\mathbf{1}_{[0,s)})] = \operatorname{cov}[Z(\mathbf{1}_{[s,t)})] = \|\mathbf{1}_{[s,t)}\|_2^2 = t-s\,.$$

•

§ 3 | The Brownian Bridge

Let $U_t = (B_t \mid B_1 = 0)$ for $0 \leq t \leq 1$. Since $B_1 = Z_{0,0}$, we have

$$U_t = (B_t \mid Z_{0,0} = 0) = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} Z_{n,k} \int_0^t h_{n,k}(s)\, ds = B_t - tB_t.$$

Chandler

⋆ ⋆ ⋆ ⋆ ⋆ ⋆ ⋆

Lecture 5
03 October 2018

§ 4 | An Application

We operate a cufflink machine:

(A cufflink)

The weights of the two ends, M_1 and M_2 are *i.i.d* $N(5, 10)$. We want to only keep cufflinks that are balanced. That means $|M_1 - M_2|$ is small.

A. It's easy to weigh both ends, $M_1 + M_2$. Does this help identify the bad cufflinks?

No. $M_1 + M_2$ and $M_1 - M_2$ are independent. This is because:

$$\operatorname{cov}[M_1 + M_2, M_1 - M_2] = \operatorname{var}[M_1] - \operatorname{var}[M_2] = 0\,.$$

B. What is the distribution of M_1 among cufflinks with $M_1 + M_2 \approx 12$?

Let $A = M_1 + M_2$ and $B = M_1 - M_2$, then $M_1 = \frac{A+B}{2}$. From Part A we know A and B are independent.

$$(M_1 | A = 12) \stackrel{d}{=} \frac{12 + B}{2} \sim N(6, 2)$$

This result follows from the fact that:

$$B \sim N(0, 2)$$

since $B = M_1 - M_2$.

§ 5 | Conditional Distributions

5.1 Lemma (Conditional Distributions) Let $\underset{n\times n}{X}$ and $\underset{m\times m}{Y}$ be jointly Gaussian with mean zero and covariance:

$$\begin{bmatrix} \underset{n\times n}{\Sigma_{XX}} & \underset{m\times n}{\Sigma_{XY}} \\ \underset{n\times m}{\Sigma_{XY}^{\mathrm{T}}} & \underset{m\times m}{\Sigma_{YY}} \end{bmatrix}$$

Note :

Then,

$$\mathbb{E}[X|Y] = \Sigma_{XY}\Sigma_{YY}^{-1}Y$$

and

$$(X|Y=y) \sim N(\Sigma_{XY}\Sigma_{YY}^{-1}, \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{XY}^{\mathrm{T}})$$

Note that if Σ_{YY} is not invertible, we may use the Moore-Penrose inverse. That is, the above equations then remain true if the inverse used is the Moore-Penrose generalized inverse. Further, note that $\Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{XY}^{\mathrm{T}}$ is the Schur complement of our covariance matrix.

Proof. Let $A = \Sigma_{XY}\Sigma_{YY}^{-1}Y$ and $X = A + B$. Claim: B is independent of Y.

Is this line rect?

$$\mathrm{cov}[X - \Sigma_{XX}\Sigma_{YY}^{-1}Y, Y] = \mathrm{cov}[X, Y] = \mathrm{cov}[X, Y] - \Sigma_{XY}\Sigma_{YY}^{-1}\,\mathrm{cov}[Y, Y] = 0\,.$$

Now we can say:

$$(X|Y=y) \overset{d}{=} \Sigma_{XY}\Sigma_{YY}^{-1}y + B$$

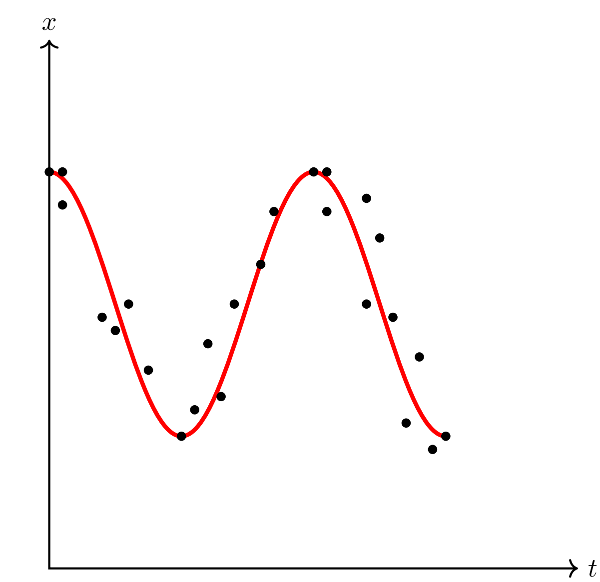
To confirm that this is the same distribution as before we need to calculate:

$$\begin{aligned} \mathrm{cov}[B, B] &= \mathrm{cov}[X - \Sigma_{XY}\Sigma_{YY}^{-1}Y, X - \Sigma_{XY}\Sigma_{YY}^{-1}Y] \\ &= \mathrm{cov}[X, X] - \mathrm{cov}[X, Y]\Sigma_{YY}^{-1}\Sigma_{XY}^{\mathrm{T}} - \Sigma_{XY}\Sigma_{YY}^{-1}\,\mathrm{cov}[Y, X] + \Sigma_{XY}\Sigma_{YY}^{-1}\,\mathrm{cov}[Y, Y]\Sigma_{YY}^{-1}\Sigma_{XY}^{\mathrm{T}} \\ &= \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{XY}^{\mathrm{T}} \end{aligned}$$

Thus, $(X|Y=y) \sim N(\Sigma_{XY}\Sigma_{YY}^{-1}, \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{XY}^{\mathrm{T}})$. □

§ 6 | Gaussian Process Regression

Suppose we have a random curve that is drawn from a Gaussian process. We observe the value of this process at a few noisy points.



Using this information we can find out more about this curve, including confidence intervals about points.

But, first we'll look more at the Brownian Bridge. Recall (for $0 \le t < 1$),

$$U_t = (B_t | B_1 = 0) \stackrel{d}{=} B_t - tB_1$$

Check this,

$$\mathrm{cov}[B_1, B_t - tB_1] = t - t \cdot 1 = 0$$

Recall,

$$U_t = B_t - Z_{0,0} \int_0^t h_0(s)ds = Z(\mathbf{1}_{[0,t)})) - tZ_{0,0} = \sum_{n \geq 1} \sum_{k=1}^{2^n} Z_{n,k} \int_0^t h_{n,k}(s)ds$$

What is the distribution of $U_t - U_s$? (for $s < t$)

$$U_t - U_s = \sum_{n \geq 1} \sum_{k=1}^{2^n} Z_{n,k} \int_s^t h_{n,k}(s)ds = Z(\mathbf{1}_{[s,t)}) - (t-s)Z_{0,0}$$

So,

$$\mathrm{var}[U_t - U_s] = \mathrm{var}[Z(\mathbf{1}_{[s,t)})] - 2(t-s) \cdot \mathrm{cov}[Z(\mathbf{1}_{[s,t)}), Z(\mathbf{1}_{[0,1)})] + (t-s)^2 \cdot \mathrm{var}[Z(\mathbf{1}_{[0,1)})]$$

Note that,

$$\mathrm{var}[Z(\mathbf{1}_{[s,t)})] = \int_0^1 \mathbf{1}_{[s,t)}(u) \cdot \mathbf{1}_{[s,t)}(u)du = t - s$$

So, we'll have

$$\mathrm{var}[U_t - U_s] = (t-s) - 2(t-s)^2 + (t-s)^2 = (t-s)[1-(t-s)]$$

Thus,

$$U_t - U_s \sim N\Big(0, (t-s)[1-(t-s)]\Big)$$

Lecture 6
05 October 2018

⋆ ⋆ ⋆ ⋆ ⋆ ⋆ ⋆

Chandler

§ 7 | Another Application

Sediment decomposition: Let $S_k =$ (amount of sediment deposited in year k) for $0 \leq K \leq N = 10^4$ and we model:

$$S_{k+1} = \mu + (1-a) \cdot [S_k - \mu] + \eta_{k+1}$$

we could also write this as:

$$[S_{k+1} | S_k \sim N\big((1-a) \cdot [S_k - \mu], \sigma^2\big)]$$

Here $\eta_{k+1} \sim i.i.d N(0, \sigma^2)$ where $\sigma^2 = 10^{-4}$, and $a = 10^{-3}$ So, S_k with $\mu = 0$ would look something like this:

process.png

As shown this is a process that stays distributed about a mean.

Suppose we have noisy observations of S_t for a few years Let Y_i be our measurements

$$Y_i = S_{k_i} + \varepsilon_i$$

For $1 \leq i \leq N$ and $\varepsilon_i \sim i.i.d N(0, \sigma_\varepsilon^2)$ Our Goal: Estimate the total amount of sediment deposited

$$S_{\text{total}} = \sum_{k=1}^{N} S_k$$

Let $D_k = S_k - \mu$ we can rewrite this as:

$$D_k = \eta_k + (1-a)D_{k-1} = \eta_k + (1-a)\eta_{k-1} + (1-a)^2 D_{k-2} = \sum_{j \geq 0} (1-a)^j \eta_{k-j}$$

Thus, we know:

$$D_k \sim N\left(0, \sigma^2 \cdot \sum_{j \geq 0} (1-a)^2 j\right) = N\left(0, \sigma^2 \frac{1}{(2-a)a}\right)$$

And since $\sigma^2 = 10^{-4}$ and $a = 10^{-3}$ we can simplify this further, $D_k \sim N(0, \sim 0.05)$. Let's say that $(W_t)_{t \in \mathbb{R}}$ is a Brownian motion and $\eta_k = W_{\frac{k+1}{N}} - W_{\frac{k}{N}} \sim N\big(0, \frac{1}{N} = \sigma^2\big)$ Then,

$$D_k = \sum_{j \geq 0} (1-a)^j \left(W_{\frac{k-j+1}{N}} - W_{\frac{k-j}{N}}\right) \approx \sum_{j \geq 0} e^{-aj} \left(W_{\frac{k-j+1}{N}} - W_{\frac{k-j}{N}}\right) = \int_{-\infty}^{t} e^{-aN(t-s)} dW_s$$

Where, $t = \frac{k}{N}$ and $0 \leq t \leq 1$ Further simplifying we get,

$$D_k = \int_{-\infty}^{t} e^{-aN(t-s)} dW_s \sim N(0, \int_{-\infty}^{t} (e^{-aN(t-s)})^2 ds = \frac{1}{2aN}) \approx \frac{\sigma^2}{2a} + \mathcal{O}(a^2 \sigma^2)$$

Let $U_t = \int_{-\infty}^{t} e^{-aN(t-s)} dW_s$ Also, $S_{\text{total}} \simeq N\mu + \sum_{k=1}^{N} U_{\frac{k}{N}} \sim N \cdot (\mu + \int_0^1 U_t dt)$

New question: Let $T = \int_0^1 U_t dt$ and $X_i = U_{t_i} + \varepsilon_i$ where $\varepsilon_i \sim i.i.d.N(0, \sigma_\varepsilon^2)$ What is the conditional distribution of $(T|X = X_1, \dots, X_N = X_N)$?

Now, we are trying to estimate the total integral, or the purple area below.

area2.png

The points on the blue curve are values of U_t, and the purple area is: $\int_0^1 U_t dt = T$ To do this we need everything inside to be jointly Gaussian. i.e. we need to use the same white noise.

Using the lemma from last lecture, we only need their covariance to calculate this area. Recall the lemma, if (X, Y) is $N\left(0, \begin{bmatrix} \Sigma'_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix}\right)$ then, $(X|Y = y) \sim N(\Sigma_{XY}\Sigma_{YY}^{-1}y, \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{XY}^{\intercal})$
To apply this we first need to expand T

$$T = \int_0^1 U_t dt = \int_0^1 \int_{-\infty}^{t} e^{-aN(t-s)} dW_s dt = \int_{-\infty}^{1} \int_{\max(s,0)}^{t} e^{-aN(t-s)} dt dW_s$$
$$= \int_{-\infty}^{1} \frac{1}{aN} e^{aN\min(0,s)} \big(1 - e^{-aN}\big) dW_s := \int_{-\infty}^{1} \phi(s) dW_s$$

Now, we know the following:

$$\text{var}[T] = \int_{-\infty}^{1} \phi^2(s) ds$$

$$\text{var}[X_i] = \text{var}[U_{t_i}] + \sigma_\varepsilon = \frac{1}{2aN} + \sigma_\varepsilon^2$$

$$\text{cov}[X_i, X_j] = \int_{-\infty}^{t_i} e^{-aN(t_i - s)} \cdot e^{-aN(t_j - s)} ds = \frac{1}{2aN}[e^{-aN(t_j - t_i)}]$$

for $i \neq j$ and $t_i < t_j$ Thus,

$$\text{cov}[X_i, T] = \int_{-\infty}^{t} e^{-aN(t_i - s)} \phi(s) ds$$

where,

$$\Sigma = \begin{bmatrix} ||\phi^2|| & 0 & \dots & \dots & 0 \\ 0 & \frac{1}{2aN} + \varepsilon^2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \frac{1}{2aN} + \varepsilon^2 \end{bmatrix}$$

Figure 1. A realization of the basic noise used to construct a Gaussian process.

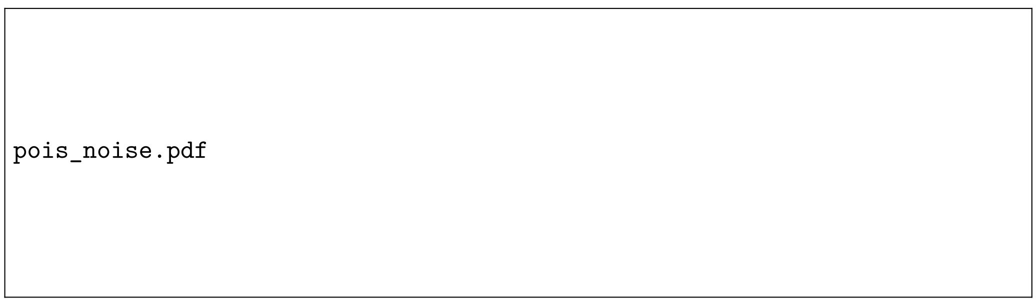


Figure 2. A realization of the basic noise used to construct a Poisson process.

Where the columns are associated with T, X_1, X_2, and so forth. And the rows are similarly associated with T, X_1, X_2, and so forth. (This means the diagonal elements are the variances of T, X_1, X_2, $\ldots$, X_N).

Matt ⋆ ⋆ ⋆ ⋆ ⋆ ⋆ ⋆ Lecture 7
01 October 2018

Recall a realization of the basic noise for Gaussian processes looked like that in Figure 1. Now, arrows are either muted or (rarely) point up. See Figure 2.

§ 8 | Motivation

Suppose in some space X we lay down a large number of LED lights, each with their own battery, with density given by a σ-finite measure μ. We do this in a way so that, for each region $A \subset X$, we put down about $M\mu(A)$ lights in that region, where M is some large number. Independently we turn on each light with probability M^{-1}, and leave off otherwise.

We would like to answer the following question: how many lights in A are on? To that end, let $N(A)$ denote the number of lights on in A and compute

$$\mathbb{E}\left[N(A)\right] = \mathbb{E}\left[\sum_{\text{lights in } A} \mathbb{1}_{\{\text{light on}\}}\right] = \sum_{\text{lights in } A} \mathbb{P}\left\{\text{light is on}\right\} = M\mu(A)\left(\frac{1}{M}\right) = \mu(A). \tag{1}$$

Thus μ gives the expected density for the set of lights that are on in A. By construction, we know $N(A) \sim \text{Binom}\left(M\mu(A), M^{-1}\right)$, and hence the distribution of $N(A)$ is approximately $\text{Pois}(\mu(A))$. To see this, put $L = M\mu(A)$ and observe,

$$\mathbb{P}\left\{N(A) = n\right\} = \binom{L}{n}\left(\frac{1}{M}\right)^n\left(1 - \frac{1}{M}\right)^{L-n} \tag{2}$$

$$= \frac{L(L-1)\cdots(L-n+1)}{n!M^n}\left(1 - \frac{1}{M}\right)^{L-n} \tag{3}$$

$$\simeq \frac{1}{n!}\left(\frac{L}{M}\right)^n \exp\left(-\frac{L}{M}\right) + \mathcal{O}\left(\frac{1}{M}\right) \tag{4}$$

$$\simeq \frac{\mu(A)^n}{n!}e^{-\mu(A)} \tag{5}$$

This motivates the following definition.

Definition. Let μ be a σ-finite measure on some space X. A *Poisson Point Process* (PPP) on X with *mean measure* (or, *intensity*) μ is a random point measure N such that:

(a) For any Borel set $A \subset X$, we have $N(A) \in \mathbb{Z}_{\geq 0}$ and $N(A) \sim \text{Pois}(\mu(A))$, i.e.

$$\mathbb{P}\left\{N(A) = n\right\} = \frac{\mu(A)^n}{n!}e^{-\mu(A)}. \tag{6}$$

(b) If A and B are disjoint Borel subsets of X, then $N(A)$ and $N(B)$ are independent random variables.

Recall a point measure is just a measure whose mass is atomic. That is, if $\{x_i\} \subset X$ then a point measure is of the form

$$\mu = \sum_i a_i \delta_{x_i} \tag{7}$$

where δ_x is the unit point mass at x.

§ 9 | PPP Properties

It is sometimes useful to think of a PPP as a random collection of points. With this in mind, we list some important properties of $N \sim \mathrm{PPP}(\mu)$ on some space X:

- *Enumeration:* It is always possible to enumerate the points of N, i.e. there is a random collection of points $\{x_i\} \subset X$ such that

$$N = \sum_i \delta_{x_i}. \tag{8}$$

- *Mean measure:* If $f\colon X \to \mathbb{R}$ then

$$\mathbb{E}\left[\int f(x)\, dN(x)\right] = \int f(x) d\mu(x). \tag{9}$$

  Note: This is a more general property of point processes, as any point process has a mean measure. To see (9) holds without needing N to be a *Poisson* point process, let f be a simple function, i.e.

$$f(x) = \sum_{i=1}^{n} f_i \mathbb{1}_{A_i}(x), \qquad \text{where} \qquad X = \bigcup_i A_i, \quad A_i \cap A_j = \varnothing \text{ for } i \neq j. \tag{10}$$

  Then we compute

$$\mathbb{E}\left[\int_X f(x)\, dN(x)\right] = \mathbb{E}\left[\sum_i f_i N(A_i)\right] = \sum_i f_i \mathbb{E}\left[N(A_i)\right] = \sum_i f_i \mu(A_i) = \int_X f(x)\, d\mu(x). \tag{11}$$

  This can then be extended to arbitrary measurable functions through the standard limiting procedure.

- *Thinning:* Independently discard each point of N with probability $1 - p(x)$ for a point at $x \in X$. The result is a $\mathrm{PPP}(\nu)$, where

$$\nu(A) = \int_A p(x)\, d\mu(x). \tag{12}$$

  In other words, if $N = \sum_i \delta_{x_i}$ and $A_i = 1$ with probability $p(x_i)$ and $A_i = 0$ otherwise, then

$$\widetilde{N} = \sum_i A_i \delta_{x_i} \sim \mathrm{PPP}(\nu). \tag{13}$$

- *Additivity:* If $N_1 \sim \mathrm{PPP}(\mu_1)$ and $N_2 \sim \mathrm{PPP}(\mu_2)$ are independent on X, then $N_1 + N_2 \sim \mathrm{PPP}(\mu_1 + \mu_2)$. In particular, if $\mathbb{P}\{X = n\} = \frac{\lambda^n}{n!} e^{-\lambda}$ and $\mathbb{P}\{Y = n\} = \frac{\nu^n}{n!} e^{-\nu}$ are independent, then

$$\mathbb{P}\{X + Y = n\} = \frac{(\lambda + \nu)^n}{n!} e^{-(\lambda+\nu)}. \tag{14}$$

- *Labeling:* For each point in a PPP, associate an independent label from a space Y according to some probability distribution ν. Let $N = \sum_i \delta_{x_i}$ for $\{x_i\} \subset X$ and let $G_1, G_2, \ldots \in Y$ be iid with density ν. Then

$$\overline{N} := \sum_i \delta_{(x_i, G_i)} \sim \mathrm{PPP}(\mu \times \nu) \tag{15}$$

  on $X \times Y$.

§ 10 | Examples

Henceforth, let λ denote Lebesgue measure.

10.1 Example Let $N \sim \mathrm{PPP}(\lambda)$ on $\mathbb{R}_{\geq 0}$, where λ is Lebesgue measure. As before, we think of the points of N as 'lights', here positioned on the positive reals.

(a) How far until the first light?
(b) Suppose each light is independently either red or green with probability $\frac{1}{2}$. How far until the first red light?

Solution. Let $N = \sum_i \delta_{x_i}$ and put $T = \min\{x_i\}$. Using (6) we compute

$$\mathbb{P}\{T > t\} = \mathbb{P}\{N([0,t]) = 0\} = e^{-t}. \tag{16}$$

This solves part (a). For the colorblind readers, this also solves part (b).

Now let $\{\tilde{x}_i\} \subset \{x_i\}$ be the (random) set of red lights and define $\widetilde{N} = \sum_i \delta_{\tilde{x}_i}$, the point process for the red lights from N. By the thinning property (12), $\widetilde{N} \sim \mathrm{PPP}\left(\frac{1}{2}\lambda\right)$. Similarly define $\widetilde{T} = \min\{\tilde{x}_i\}$ and observe

$$\mathbb{P}\left\{\widetilde{T} > t\right\} = \mathbb{P}\left\{\widetilde{N}([0,t]) = 0\right\} = e^{-t/2}, \tag{17}$$

thus (b) is solved. ♦

•

10.2 Example Rain falls for 10 minutes on a large patio at a rate of $\nu = 5000$ drops per minute per square meter. Each drop splatters to a random radius R that has an Exponential distribution, with mean 1cm, independently of the other drops. Assume the drops are 1mm thick and the set of locations of the raindrops is a PPP.

(a) What is the mean and variance of the total amount of water falling on a square with area 1m^2?
(b) A very small ant is running around the patio. See Figure **??**. What is the chance the ant gets hit?

Solution. Let $N = \sum_i \delta_{(x_i,y_i)}$ where (x_i, y_i) is the center of the ith drop. Take $N \sim \mathrm{PPP}(\nu\lambda)$ and let M denote the number of drops in $[0,1]^2$, so that $M = N([0,1]^2) \sim \mathrm{Pois}(\nu)$. Then the total volume V is

$$V = \sum_{i=1}^{M} \frac{\pi}{10^3} R_i^2 \tag{18}$$

where R_i is the radius of the ith drop. Note this is a sum of random variables where the number of terms is also a random variable. Thus we use Wald's equation (28) to obtain

$$\mathbb{E}[V] = \frac{\pi}{10^3}\mathbb{E}[M]\mathbb{E}[R_1^2] = \frac{\pi}{10^3} \cdot \nu \cdot \frac{2}{100^2} = \frac{2\pi}{10^7}\nu \tag{19}$$

The second step in (19) was obtained from the fact that an exponentially distributed random variable X with mean β^{-1} has higher moments given by

$$\mathbb{E}[X^n] = \frac{n!}{\beta^n}. \tag{20}$$

This is proved by an iterated application of integration by parts, and the result gives rise to

$$\mathrm{var}[X^n] = \mathbb{E}\left[X^{2n}\right] - \mathbb{E}[X^n]^2 = \frac{(2n)! - (n!)^2}{\beta^{2n}}. \tag{21}$$

The $n = 2$ case will turn out to be useful when computing the variance of V.

Indeed, to compute the variance we utilize the variance decomposition formula. Observe,

$$\begin{aligned}
\operatorname{var}[V] &= \mathbb{E}\left[\operatorname{var}[V \mid M]\right] + \operatorname{var}\left[\mathbb{E}[V \mid M]\right] \qquad (22)\\
&= \mathbb{E}\left[M\left(\frac{\pi}{10^3}\right)^2 \operatorname{var}(R^2)\right] + \operatorname{var}\left[M\left(\frac{\pi}{10^3}\right)\mathbb{E}[R^2]\right] \qquad (23)\\
&= \nu\left(\frac{\pi}{10^3}\right)^2\left(\frac{20}{100^4}\right) + \nu\left(\frac{\pi}{10^3}\right)^2\left(\frac{2}{100^2}\right)^2 \qquad (24)\\
&= \left(\frac{\pi}{10^3}\right)^2\left(\frac{24}{100^4}\right)\nu. \qquad (25)
\end{aligned}$$

This solves part (a).

Now, part (b) can be solved by way of the labeling property. Here, we use the radius R_i of the ith drop to label the point (x_i, y_i). Recall the density of an Exponential random variable with mean 0.01 is $100\exp(-100r)\,dr$. So we define a measure μ on $X := \mathbb{R}^2 \times [0, \infty)$ by

$$\mu(A) = \int_A 100\nu \exp(-100r)\, dxdydr. \qquad (26)$$

We think of X as the (closed) upper half plane in $\mathbb{R}^3$ where the third coordinate is a realization of R. By the labeling property (15), $\overline{N} := \sum_i \delta_{(x_i,y_i,R_i)} \sim \mathrm{PPP}(\mu)$ on X. For the ant to remain dry, any drop with radius r must land outside the circle of radius r centered at the ant. Viewed from the space X, we want to integrate over the cone with its tip at the ant, whose horizontal cross-section at height r is a circle of radius r. From this we compute

(27)

$$\mathbb{P}\{\text{ant is dry}\} = \mathbb{P}\left\{\overline{N}(A) = 0\right\} = \exp(-\mu(A)) = \exp\left(-100\pi\nu\int_0^\infty r^2 e^{-100r}\,dr\right) = \exp\left(-\frac{\pi\nu}{5000}\right).$$

Plugging in the given value for ν yields $\mathbb{P}\{\text{ant is dry}\} = \exp(-\pi) \approx 0.0432$. The ant had better grab an umbrella! ♦

•

10.1 Wald's Equation

The following is the statement of Wald's equation, taken from Wikipedia †.

10.3 Theorem (Wald's Equation) Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of real-valued, independent and identically distributed random variables and let N be a nonnegative integer-valued random variable that is independent of the sequence $(X_n)_{n\in\mathbb{N}}$. Suppose that N and the X_n have finite expectations. Then

$$\mathbb{E}\left[\sum_{i=1}^{N} X_i\right] = \mathbb{E}[N]\,\mathbb{E}[X_1]. \qquad (28)$$

† The proof is also on Wikipedia.

Figure 3. A realization of the ant from Example 10.2. Looks like he had an umbrella after all.

rain_ant.jpg