

La pregunta

Como dije antes, el Cálculo Malliavin no se relaciona fácilmente con cosas cotidianas. La mayoría de los autores que consulté no detallan demasiado por qué un objeto como este existe, o por qué existe la fórmula de Clark-Ocone. En cambio, les interesa profundizar sobre la condición de Hormander (?).

Dijimos antes que el Cálculo Malliavin es la extensión del cálculo de variaciones a los procesos estocásticos. Ha llegado el momento de explorar el tópico de procesos estocásticos y martingalas.

Procesos estocásticos

Podemos pensar los procesos estocásticos como una sucesión de variables aleatorias indexadas por tiempo. Comenzaré con un caso muy simple, el paseo aleatorio unidimensional (también conocido como la caminata de borracho):

$$\mathbb{P}(\{X_i = +1\}) = 0.5$$

$$\mathbb{P}(\{X_i = -1\}) = 0.5$$

$$X_0 = 0$$

$$Z_n = \sum_{i=1}^n X_i$$

Esto es muy simple: en el momento i , después de que haya transcurrido un tiempo Δt , el borracho da un paso X_i , que puede ser hacia arriba o hacia abajo en la calle con la misma probabilidad. La posición Z en t es simplemente la ubicación final, luego de todos los pasos que el borracho dio. Con esta forma de plantear el problema, cada movimiento “hacia arriba” puede considerarse un “éxito”, por lo que esta suma sigue una distribución binomial centrada en 0.

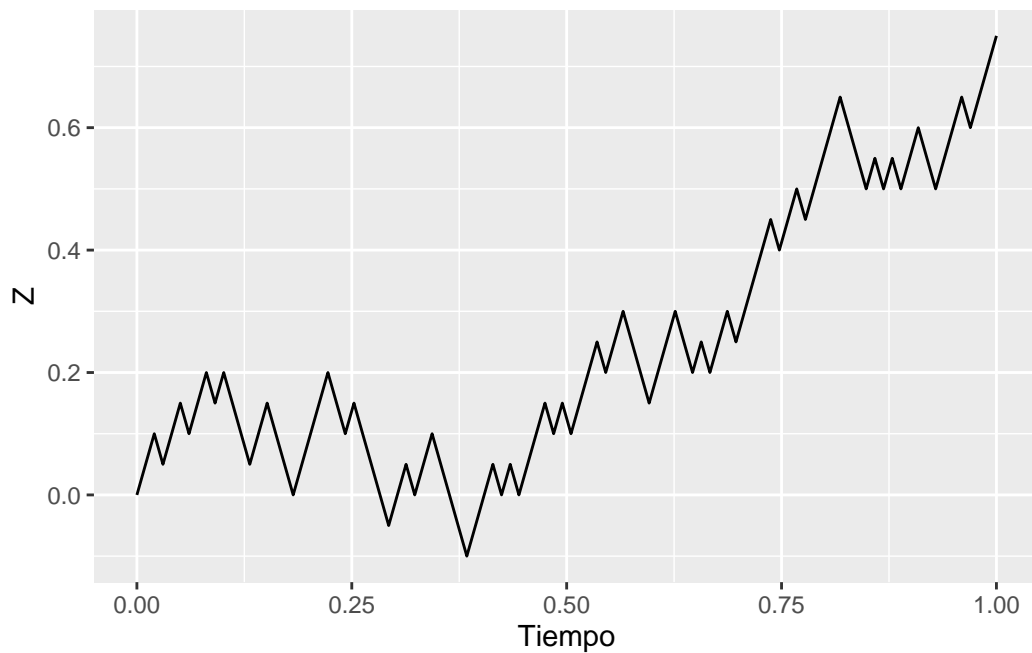
Pensemos otro caso: dejemos que el recorrido aleatorio dé pasos más frecuentes y más cortos. Usando los símbolos de arriba, Δt y X_i se hacen más pequeños. A medida que los pasos se vuelven infinitesimales, el proceso se vuelve continuo. Así es como se vería una caminata cuando $\Delta t = 0.1, 0.01, 0.001$:

```

library(ggplot2)
library(rgl)

# Left
h <- 0.01
steps = 1 / h
xl <- seq(from = 0, to = 1, length.out=steps)
yl <- (rbinom(steps-1,1,0.5)*2-1)*sqrt(h)/2
zl <- c(0,cumsum(yl))
ggplot(mapping = aes(x = xl), legend=TRUE) +
  geom_line(mapping = aes(y=zl, fill="f(X)")) +
  xlab('Tiempo') +
  ylab('Z')

```

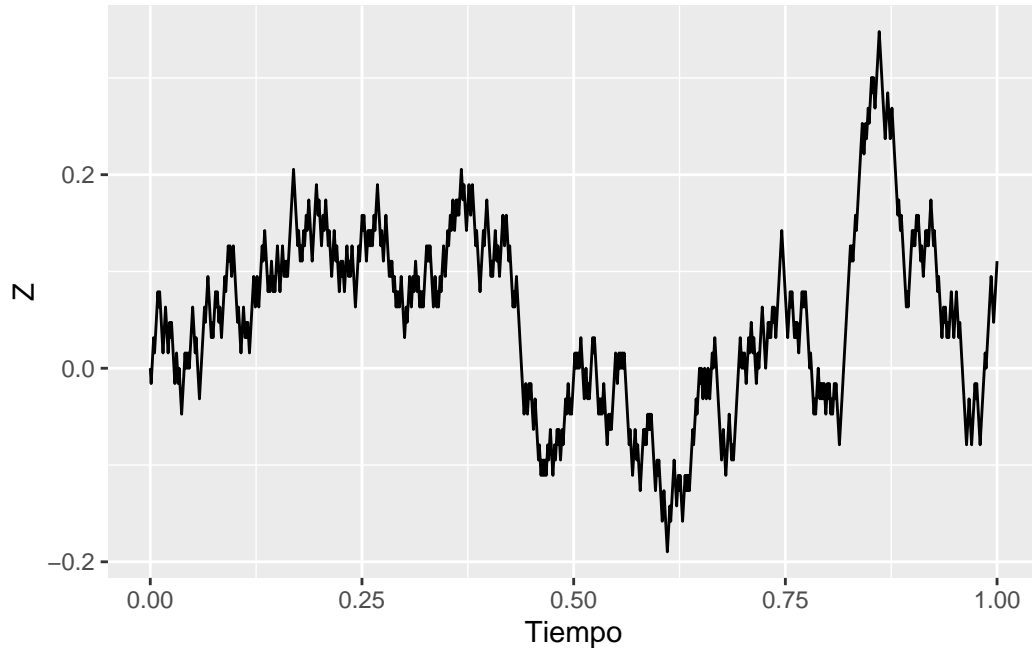


```

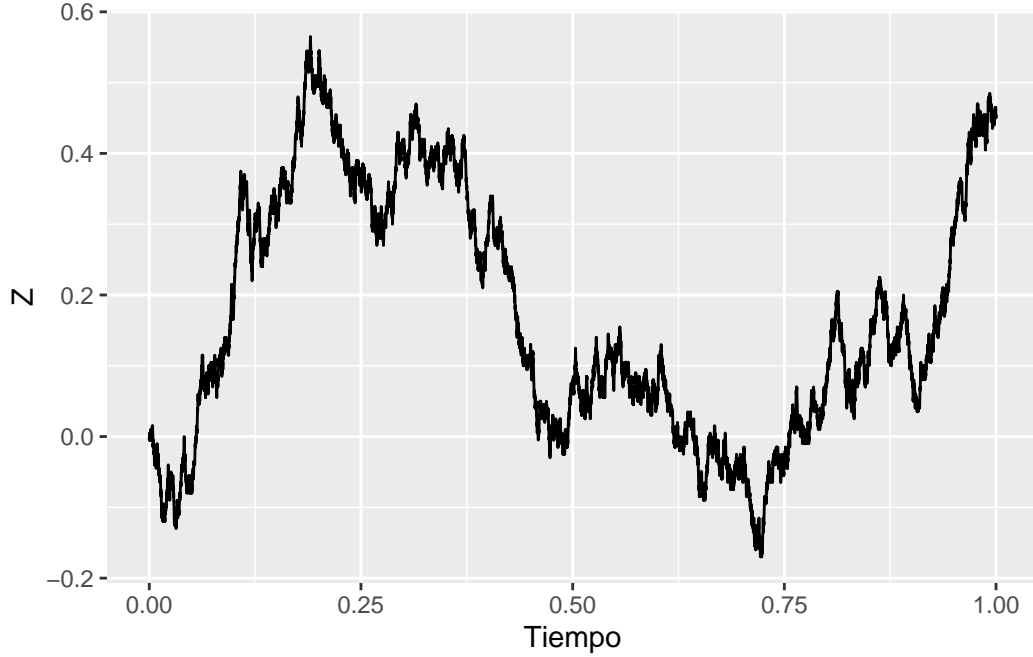
# Middle
h <- 0.001
steps = 1 / h
xl <- seq(from = 0, to = 1, length.out=steps)
yl <- (rbinom(steps-1,1,0.5)*2-1)*sqrt(h)/2
zl <- c(0,cumsum(yl))
ggplot(mapping = aes(x = xl), legend=TRUE) +

```

```
geom_line(mapping = aes(y=z1, fill="f(X)")) +
xlab('Tiempo') +
ylab('Z')
```



```
# Right
h <- 0.0001
steps = 1 / h
x1 <- seq(from = 0, to = 1, length.out=steps)
y1 <- (rbinom(steps-1,1,0.5)*2-1)*sqrt(h)/2
z1 <- c(0,cumsum(y1))
ggplot(mapping = aes(x = x1), legend=TRUE) +
  geom_line(mapping = aes(y=z1, fill="f(X)")) +
  xlab('Tiempo') +
  ylab('Z')
```



Una caminata aleatoria unidimensional de 0 a t , de manera similar al Teorema del Límite Central, converge a una variable normal:

$$Z_t \sim \mathcal{N}(\mu = 0, \sigma^2 = t)$$

Estos procesos estocásticos de tiempo continuo se denominan movimientos brownianos o procesos de Wiener. Se denotan como B_t o W_t y cumplen un conjunto de condiciones:

- $B_0 = 0$
- Los incrementos son independientes, es decir, $(B_{t+s} - B_t)$ no depende de valores anteriores B_r , $r < t$
- Los incrementos siguen una distribución normal, es decir, $(B_{t+s} - B_t) \sim \mathcal{N}(\mu = 0, \sigma^2 = s)$
- B_t es continua en t . Un poco circular, ya sé.

Sin embargo, hay un problema. Aunque son continuas, estas funciones no son diferenciables. Se puede ver en los gráficos de trayectorias de más arriba, con lo irregulares que se vuelven a medida que los índices (o pasos) se achican. También tiene sentido si pensás en la forma en que construimos el proceso: cuando estás en un punto en el tiempo t , no sabés de dónde viniste aleatoriamente. ¿Llegaste desde un valor más alto o más bajo en $t - \Delta t$? No tiene sentido hablar de $\frac{\partial Z_t}{\partial t}$.

Cálculo de Ito

La caminata de borracho es un caso muy simple de un proceso estocástico. ¿Qué pasa con variables aleatorias más interesantes, o con funciones variables aleatorias?

Empezamos con una sustitución sencilla: no podemos hacer $\frac{\partial Z_t}{\partial t}$ pero “podemos” pasar el dt multiplicando, como hacemos con las derivadas totales, y referirnos solo a dZ_t . En general, nos gustaría expresar un cambio minúsculo en la variable aleatoria como una suma entre un cambio asociado al tiempo y un cambio asociado a un “ruido”. En resumen, algo como esto:

$$dZ_t = a(\cdot)dt + b(\cdot)dB_t$$

Para empezar, supondremos que las funciones a y b solamente dependen del tiempo. Entonces, integramos a lo largo del tiempo y llegamos a la definición de un **proceso de Ito**:

$$\int_0^t dZ_t = Z_t - Z_0 = \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s$$

Esta representación es buena porque muestra directamente que la variable aleatoria Z_t tiene una media o tendencia ligada al tiempo, y una varianza o dispersión ligada al ruido. Dicho esto, necesitamos poder trabajar más fácilmente con dB_t . Para ello, vamos a suponer algo simple: que en lugar de un continuo de dB_t , hay una lista de t_i , con incrementos ΔB_{t_i} . La expresión de arriba nos queda ahora:

$$\sigma_{t_i} * (B_{t_i} - B_{t_{i-1}})$$

Podemos calcular esta expresión porque una diferencia de movimientos brownianos se puede reemplazar con una distribución normal que podemos manipular. Para demostrar cómo esto nos ayuda, podemos hacer el siguiente cálculo y confirmar que el valor esperado de Z sólo depende del primer término:

$$\mathbb{E} \left[\int_0^t \sigma_s dB_s \right] = \sum_{i=1}^n \sigma_{t_i} \mathbb{E} [B_{t_i} - B_{t_{i-1}}] = \sum_{i=1}^n \sigma_{t_i} \mathbb{E} [\mathcal{N}(0, t_i - t_{i-1})] = 0$$

Usando este enfoque, podemos demostrar que los procesos de Ito satisfacen:

$$\begin{aligned} \mathbb{E} [X_t] &= \int_0^t \mu_s ds \\ \text{Var} [X_t] &= \int_0^t \sigma_s^2 ds \end{aligned}$$

Otra forma de manipular estos B_t es con la **isometría de Ito**, que permite intercambiar el dB_t por un dt más amigable de integrar:

$$\mathbb{E} \left[\left(\int_0^t X_s dB_s \right)^2 \right] = \mathbb{E} \left[\int_0^t X_s^2 ds \right]$$

Con estas herramientas, ya podemos calcular una media y una varianza para cada punto en el tiempo de un proceso. Pero hay una herramienta aún más poderosa en nuestro arsenal: el **lema de Ito** o **fórmula de Ito**. Comenzamos su presentación con un proceso de Ito para una variable aleatoria, digamos X_t :

$$\begin{aligned} dX_t &= \mu_t dt + \sigma_t dB_t \\ X_t &= X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s \end{aligned}$$

Ahora, queremos plantear algo más complejo. Algo como:

$$Y_t = f(t, X_t)$$

$$dY_t = a(t, X_t) dt + b(t, X_t) dB_t$$

with

$$dX_t = \mu_t dt + \sigma_t dB_s$$

Nuestro proceso ya no depende solo del tiempo, sino de una variable aleatoria que sigue un proceso de Ito. De nuevo, no podemos simplemente hacer la regla de la cadena típica del cálculo. Es decir, todavía no podemos hacer algo como:

$$\frac{\partial Y_t}{\partial t} = \frac{\partial Y_t}{\partial X_t} \frac{\partial X_t}{\partial t}$$

Pero se puede demostrar que existe una regla de la cadena alternativa que podemos aplicar: el **lema de Ito**:

$$\begin{aligned} df(t, X_t) &= \frac{\partial f}{\partial t} dt + \mu_t \frac{\partial f}{\partial x} dt + \sigma_t \frac{\partial f}{\partial x} dB_t + \frac{\sigma_t^2}{2} \frac{\partial^2 f}{\partial x^2} dt \\ df(t, X_t) &= \underbrace{\left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{\sigma_t^2}{2} \frac{\partial^2 f}{\partial x^2} \right)}_{a(t, X_t)} dt + \underbrace{\left(\sigma_t \frac{\partial f}{\partial x} \right)}_{b(t, X_t)} dB_t \end{aligned}$$

La primera ecuación casi parece una expansión en serie de Taylor hasta la 2ª derivada, y de hecho es lo que sucede en demostraciones informales. La segunda ecuación muestra las expresiones explícitas para las funciones $a(t, X_t)$ y $b(t, X_t)$ que mencionamos antes. Esto es genial porque todavía podemos separar el componente de “tendencia” y el componente de “ruido” para procesos aleatorios más complicados. Esto también significa que la “tendencia” se ve afectada por la intensidad del “ruido”. Por último, se puede verificar que una función no aleatoria con $\sigma_t = 0$ se convierte en una derivada total clásica, con la regla de la cadena que ya conocemos.

Martingalas y nociones finales

Algunos procesos estocásticos presentan propiedades adicionales. Uno de ellos son las martingalas. Las martingalas son procesos de tiempo discretos o continuo que satisfacen:

- $\mathbb{E}[|X_t|] < \infty, \forall t \geq 0$. Esto significa que el proceso siempre tiene un valor finito.
- $\mathbb{E}[X_{t+s}|X_t] = X_t, \forall t \leq s$. Esto significa que la esperanza matemática para futuras realizaciones del proceso es el valor que tenga en el momento presente.

La caminata de borracho es un ejemplo de martingala: en cualquier momento, es igualmente probable que suba o baje por la calle, por lo que la esperanza es dondequiera que esté en ese momento. Los movimientos brownianos también son martingalas: los incrementos siguen una distribución normal, por lo que la esperanza del incremento es 0. Otros ejemplos de martingalas típicas son apostar al negro o al rojo en la ruleta, o el precio de una acción en el corto plazo.

Existe un lema llamado el Teorema de Representación en Martingala, que establece que cualquier variable aleatoria X puede escribirse en términos de otro proceso C , con valores conocidos de antemano, de la siguiente manera:

$$X = \mathbb{E}[X] + \int_0^\infty C_s dB_s$$

Esto es valioso para la gente de finanzas porque significa que cualquier estrategia de inversión (que es un proceso aleatorio) puede replicarse con una estrategia de inversión diferente (otro proceso), pero con menor volatilidad (con un costo inicial más alto). Este teorema es menos valioso porque no proporciona ninguna forma de calcularla: no podemos aplicar una derivada y hacer cambios de variable para extraer ese dB_t

Una segunda forma de expresar martingalas, de Rogers y Williams (2000), es como integrales de Ito:

$$M_t = M_0 + \int_0^t \alpha_s dB_s$$

En un sentido acotado, α_t puede considerarse como una especie de $\frac{\partial M_t}{\partial t}$ porque si integramos hasta t recuperamos el punto final, pero no es exactamente lo que estamos buscando.

Todo esto es frustrante. Estamos tratando de encontrar una forma de calcular derivadas en estos procesos aleatorios, movimientos brownianos y martingalas todo el tiempo, pero la respuesta nos esquiva. Y el caso es que sabemos que es posible. De hecho, volvamos a nuestro primer proceso estocástico continuo, el proceso de Wiener unidimensional o caminata de borracho. Habíamos llegado al punto en el que podíamos decir que W_t tenía una distribución normal. Entonces, nada nos impide tomar una derivada sobre eso, ¿verdad?

$$W_t \sim \mathcal{N}(0, t)$$

$$W(x, t) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/(2t)}$$

$$\frac{\partial W}{\partial t} = \frac{e^{-x^2/(2t)} t^{-5/2} (x^2 - t)}{2\sqrt{2\pi}}$$

Y ahí la tenemos. No sabemos si es útil, y la expresión se ve fea, pero podemos tener una derivada de la densidad de una variable aleatoria con respecto al tiempo. Dibujemos tanto la densidad como la derivada en 3D.

```
options(rgl.useNULL=TRUE)
rgl::setupKnitr(autoprint = TRUE)
h = 200
Wt <- function(x, t) {
  if (t == 0) {
    return(0)
  } else {
    return(exp(-x^2/2/t)/sqrt(2*pi*t))
  }
}
dWt <- function(x, t) {
  if (t == 0) {
    return(0)
  } else {
    return(exp(-x^2/2/t)*t^(-5/2)*(x^2-t)/sqrt(8*pi))
  }
}
Wt <- Vectorize(Wt)
dWt <- Vectorize(dWt)

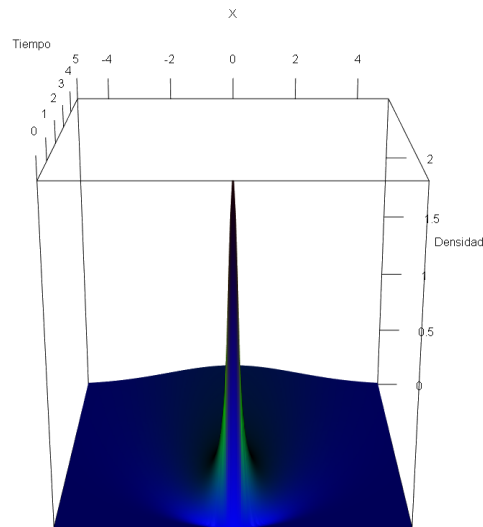
times <- seq(from=0, to=5, length.out=h)
xes <- seq(from=-5, to=5, length.out=h)
```



```

W <- outer(xes, times, Wt)
# persp(xes, times, W, col='white', shade=.1, theta = 150, phi = 15, ticktype='detailed',
# open3d(windowRect=c(50,50,800,800))
palette <- colorRampPalette(c("blue", "green", "yellow", "red"))
col.table <- palette(256)
col.ind <- cut(W, 256)
persp3d(x=xes, y=times, z=W, col=col.table[col.ind],
        xlab="X", ylab="Tiempo", zlab="Densidad")

```

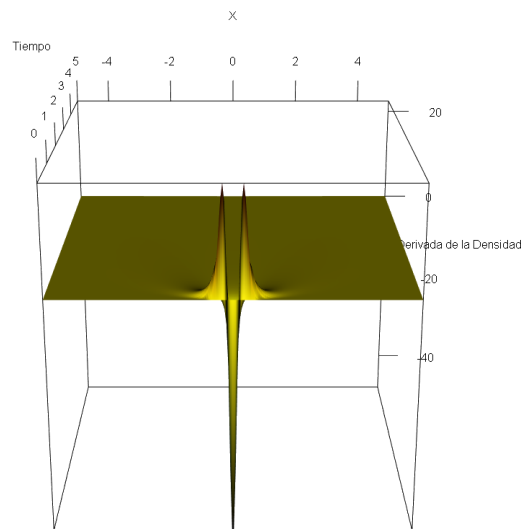


```

# rglwidget()

dW <- outer(xes, times, dWt)
# persp(xes, times, dW, col='white', shade=.1, theta = 150, phi = 15, ticktype='detailed',
col.ind <- cut(dW, 256)
persp3d(x=xes, y=times, z=dW, col=col.table[col.ind],
        xlab="X", ylab="Tiempo", zlab="Derivada de la Densidad")

```



```
# rglwidget()
```

Conclusión

A estas alturas, ya hemos visto suficiente. Sabemos lo que queremos: una forma de calcular derivadas sobre variables aleatorias y procesos de Wiener que, además, sirva para algo. Eso es lo que el Cálculo de Malliavin pretende.

Rogers, L. C. G., y David Williams. 2000. *Diffusions, Markov Processes and Martingales*. 2.^a ed. Vol. 2. Cambridge Mathematical Library. Cambridge University Press. <https://doi.org/10.1017/CBO9780511805141>.