

# Simulación de variables, vectores y procesos aleatorios<sup>1</sup>

30 de agosto de 2017

<sup>1</sup>Notas de clase del curso “Simulación en procesos estocásticos” (2017) preparadas por E. Mordecki

# Índice general

<b>1. Simulación de variables aleatorias</b>	<b>3</b>
1.1. Espacio de Probabilidad. Axiomas de Kolmogorov . . . . .	3
1.2. Variables aleatorias y distribución de probabilidad . . . . .	4
1.3. Función cuantil y distribución inversa . . . . .	5
1.4. Ejercicios . . . . .	7
<b>2. Simulación de variables y vectores aleatorios</b>	<b>9</b>
2.1. Simulación de variables aleatorias discretas . . . . .	9
2.1.1. Inversión . . . . .	9
2.1.2. Variables geométricas . . . . .	9
2.2. El método de aceptación y rechazo . . . . .	10
2.3. Distribución uniforme multivariada . . . . .	13
2.4. Ejercicios . . . . .	13
<b>3. Vectores gaussianos</b>	<b>15</b>
3.1. Vectores gaussianos . . . . .	15
3.2. Simulación de vectores gaussianos . . . . .	16
3.2.1. Simulación de variables gaussianas: el método de Box- Muller . . . . .	16
3.2.2. Simulation of bi-dimensional Gaussian vectors . . . . .	17
3.2.3. Simulación de vectores gaussianos: el método de Cho- lesky . . . . .	17
<b>4. Procesos gaussianos</b>	<b>20</b>
4.1. Procesos estocásticos y su simulación . . . . .	20
4.2. Proceso de Wiener . . . . .	22
4.2.1. Simulation of Brownian motion . . . . .	23
4.3. Ejercicios . . . . .	24

<b>5. Otros procesos gaussianos</b>	<b>25</b>
5.1. Sobre las trayectorias del movimiento Browniano . . . . .	25
5.2. Geometric Brownian Motion . . . . .	28
5.3. Brownian bridge . . . . .	28
5.3.1. Kolmogorov test . . . . .	29
5.3.2. Application . . . . .	29
5.4. Stationary gaussian processes . . . . .	30
5.5. Proceso de Slepian . . . . .	30
5.6. Ornstein-Uhlenbeck process . . . . .	31
5.7. Ejercicios . . . . .	32

# Capítulo 1

## Simulación de variables aleatorias

### 1.1. Espacio de Probabilidad. Axiomas de Kolmogorov

Consideremos un cierto conjunto  $\Omega$  no vacío, que llamamos *espacio de sucesos elementales*. A sus elementos, que llamamos *sucesos elementales* o *puntos*, los designamos con la letra  $\omega$ , con índices o sin ellos.

Sea  $\mathcal{A}$  un cierto conjunto no vacío de subconjuntos de  $\Omega$ , que cumple las siguientes propiedades: (1) si  $\mathbf{A} \in \mathcal{A}$  entonces  $\Omega \setminus \mathbf{A} \in \mathcal{A}$ ; (2) si  $\mathbf{A}_1, \mathbf{A}_2, \dots$  es un conjunto finito o numerable de subconjuntos pertenecientes a  $\mathcal{A}$ , entonces  $\bigcup_n \mathbf{A}_n \in \mathcal{A}$ . El conjunto  $\mathcal{A}$  se llama  *$\sigma$ -álgebra de sucesos*, o *campo boreliano de sucesos*, y sus elementos se llaman *sucesos*.

Observemos que el conjunto de todos los subconjuntos de un espacio  $\Omega$  es una  $\sigma$ -álgebra de sucesos, pero no toda  $\sigma$ -álgebra de sucesos es el conjunto de todos los subconjuntos de algún espacio  $\Omega$ .

Si  $\mathcal{A}$  es una  $\sigma$ -álgebra de sucesos, tenemos  $\Omega \in \mathcal{A}$  en vista de la igualdad  $\mathbf{A} \cup (\Omega \setminus \mathbf{A}) = \Omega$ , válida para todo conjunto  $\mathbf{A}$ ; además, el conjunto vacío  $\emptyset$  (llamado *suceso imposible*) también pertenece a  $\mathcal{A}$ . Si  $\mathbf{A}_1, \mathbf{A}_2, \dots$  es un conjunto finito o numerable de sucesos pertenecientes a la  $\sigma$ -álgebra  $\mathcal{A}$ , entonces  $\bigcap_n \mathbf{A}_n \in \mathcal{A}$  en vista de la igualdad  $(\bigcap_n \mathbf{A}_n)^c = \bigcup_n \mathbf{A}_n^c$ , donde  $\mathbf{B}^c = \Omega \setminus \mathbf{B}$  (complemento del conjunto  $\mathbf{B}$ ) para cualquier conjunto  $\mathbf{B}$ .

En resumen, toda  $\sigma$ -álgebra de sucesos es un conjunto de subconjuntos (no necesariamente todos) de un espacio de sucesos elementales  $\Omega$ , que contiene, junto con cada uno de sus elementos a su complemento, y junto con cualquier conjunto finito o numerable de sus elementos a su unión y a su in-

tersección; además, el propio espacio de sucesos elementales  $\Omega$  y el conjunto vacío  $\emptyset$  pertenecen a toda  $\sigma$ -álgebra de sucesos.

El surgimiento de la teoría de la probabilidad es muy anterior a la creación de la teoría de conjuntos. Por ésto, desde su mismo inicio, en teoría de la probabilidad se utilizó (y continúa utilizándose) una terminología específica, diferente de la terminología utilizada en teoría de conjuntos. Las letras  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$ , con índices o sin ellos, designan a los sucesos, es decir, a los elementos de una  $\sigma$ -álgebra de sucesos  $\mathcal{A}$ , relacionada con algún espacio de sucesos elementales  $\Omega$ .

## 1.2. Variables aleatorias y distribución de probabilidad

Consideremos un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbf{P})$ . Llamamos *variable aleatoria* a una función  $X = X(\omega)$  que toma valores reales, definida en el espacio de sucesos elementales  $\Omega$ , y que verifica la condición

$$\{\omega \in \Omega: X(\omega) \leq x\} \in \mathcal{A} \quad (1.1)$$

para todo  $x$  real.

En la terminología del análisis real, una función  $X(\omega)$  que cumple la condición (1.1) para todo  $x$ , se denomina *medible*. De esta forma, una variable aleatoria es una función real y medible de los sucesos elementales. Se puede verificar que la condición (1.1) para todo  $x$ , es equivalente a la condición

$$\{\omega \in \Omega: X(\omega) \in B\} \in \mathcal{A} \quad (1.2)$$

para cualquier conjunto boreliano<sup>1</sup>  $B$  de puntos de la recta real  $\mathbf{R}$ . En el caso particular en el que  $B$  es el intervalo  $(-\infty, x]$ , la condición (1.2) se convierte en la condición (1.1).

Como el conjunto  $\{\omega \in \Omega: X(\omega) \leq x\}$  es un suceso (es decir, un conjunto de la  $\sigma$ -álgebra de sucesos  $\mathcal{A}$ ), está definida la probabilidad  $\mathbf{P}(\{\omega: X(\omega) \leq x\})$  para todo  $x \in \mathbf{R}$ ; esta probabilidad será designada por brevedad  $\mathbf{P}(X \leq x)$  (se lee: la probabilidad de que la variable aleatoria  $X$  tome un valor menor o igual que  $x$ ). Se denomina *función de distribución* de la variable aleatoria  $X$ , a la función  $F(x)$ , definida para todos los valores  $x$  reales, mediante la fórmula

$$F(x) = \mathbf{P}(X \leq x). \quad (1.3)$$

---

<sup>1</sup>La clase de los conjuntos borelianos en la recta es la mínima  $\sigma$ -álgebra de conjuntos, que contiene a todos los intervalos.

Observemos que si  $X$  es una variable aleatoria y  $B$  un conjunto boreliano de puntos de la recta real  $\mathbf{R}$ , está definida la probabilidad  $\mathbf{P}(\{\omega: X(\omega) \in B\})$ , que será designada por  $\mathbf{P}(X \in B)$ , también por brevedad. La función  $\mathbf{P}_X(B) = \mathbf{P}(X \in B)$ , definida para todos los conjuntos borelianos  $B$  de puntos de la recta real, se llama *función de probabilidad* de la variable aleatoria  $X$ . Es claro que  $\mathbf{P}_X((-\infty, x]) = \mathbf{P}(X \leq x) = F(x)$  para cualquier  $x$ , donde  $F(x)$  es la función de distribución de la variable aleatoria  $X$ . Llamamos *distribución de probabilidad* (o más sencillamente *distribución*) de la variable aleatoria  $X$ , indistintamente, a la función de distribución  $F(x)$  de la variable aleatoria  $X$ , o a la función de probabilidad  $\mathbf{P}_X(B)$  de esta variable aleatoria.

*Ejemplo 1.* [Distribución uniforme] Consideremos el espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbf{P}) = ([0, 1], \mathcal{B}, \lambda)$ , es decir, el intervalo  $[0, 1]$  con su respectiva  $\sigma$ -álgebra Borel y  $\lambda$  la medida de Lebesgue. Para  $x \in [0, 1]$  consideramos la variable aleatoria  $U(x) = x$ , es decir, la identidad. Es claro que si  $\mathbf{B} \in \mathcal{B}$  tenemos

$$\mathbf{P}(\mathbf{B}) = \lambda(B),$$

y  $F_U(x) = x$ , si  $x \in [0, 1]$ , con  $F_U(x) = 0$  para  $x < 0$  y  $F_U(x) = 1$  para  $x > 1$ . La variable  $U$  tiene *distribución uniforme*.  $\square$

### 1.3. Función cuantil y distribución inversa

Si la distribución  $F$  de una variable aleatoria es estrictamente creciente y continua, su inversa  $F^{-1}(p)$ ,  $p \in [0, 1]$  es el único real  $x$  que verifica  $F(x) = p$ . En tal caso,  $F^{-1}$  se define como la *distribución inversa* o *función cuantil*.

Cuando la distribución no tiene estas propiedades (no es biyectiva, como la distribución uniforme) definimos, para  $p \in [0, 1]$ , la *distribución inversa generalizada*

$$F^{-1}(p) = \inf\{x \in \mathbf{R} : F(x) \geq p\}.$$

Por ejemplo,  $F^{-1}(0,5)$  define la mediana de la distribución  $X$ , y  $F^{-1}(0,95)$  el percentil del 95 %.

Veamos algunas propiedades.

#### Proposición 1 (Propiedades de la función cuantil)

- (a)  $F^{-1}$  es no decreciente.
- (b)  $F^{-1}(F(x)) \leq x$  para todo  $x \in \mathbf{R}$ .
- (c)  $p \leq F(F^{-1}(p))$  para todo  $p \in [0, 1]$ .

(d)  $F^{-1}(p) \leq x$  si y sólo si  $p \leq F(x)$

(e)  $F^{-1}$  es continua por izquierda.

(f) Si  $U$  tiene distribución  $U[0, 1]$ , entonces  $F^{-1}(U)$  tiene distribución  $F$ . Esta propiedad es la base del método de simulación de la transformación inversa.

*Demostración.* Veamos la demostración de la propiedad (e). Sea  $p_n \nearrow p$ . Por monotonía, sabemos que  $F^{-1}(p_n)$  es no decreciente, por lo que existe  $\ell = \lim_n F^{-1}(p_n)$ . Si

$$\ell < F^{-1}(p) = \inf\{x: F(x) \geq p\}$$

resulta que  $F(\ell) < p$ . Aplicando la propiedad (d), tenemos

$$p_n \leq F(F^{-1}(p_n)) \leq F(\ell) < p.$$

Esto es una contradicción, porque  $p_n \nearrow p$ . Luego  $\ell = F^{-1}(p)$ , probando la propiedad.  $\square$

*Ejemplo 2.* (Exponential distribution).

An exponential random variable  $X$  with parameter  $\alpha$  has density and distribution:

$$f(x) = \begin{cases} \alpha e^{-\alpha x}, & x \geq 0, \\ 0 & x < 0, \end{cases} \quad F(x) = \begin{cases} 1 - e^{-\alpha x}, & x \geq 0, \\ 0 & x < 0, \end{cases}$$

The inverse is simple to compute:

$$F^{-1}(x) = -\frac{1}{\alpha} \log(1 - x), \quad 0 \leq x \leq 1.$$

So  $-\frac{1}{\alpha} \log(1 - U)$  is exponential, as  $U$  is symmetric, also

$$X = -\frac{1}{\alpha} \log U \text{ is exponential with parameter } \alpha.$$

$\square$

*Ejemplo 3.* (Triangular distribution) If  $U_1, U_2$  are independent uniform in  $[0, 1]$ , then  $X = U_1 + U_2$  has triangular distribution. Density and distribution of  $X$  are

$$f(x) = \begin{cases} x, & 0 \leq x \leq 1, \\ 2 - x, & 1 \leq x \leq 2, \end{cases} \quad F(x) = \begin{cases} x^2/2, & 0 \leq x \leq 1, \\ 1 - (2 - x)^2/2, & 1 \leq x \leq 2. \end{cases}$$

We then have

$$F^{-1}(x) = \begin{cases} \sqrt{2x}, & 0 \leq x \leq 1/2, \\ 2 - \sqrt{2(1-x)}, & 1/2 \leq x \leq 1. \end{cases}$$

So  $F^{-1}(U)$  has triangular distribution if  $U$  is uniform in  $[0, 1]$   $\square$

## 1.4. Ejercicios

**Ejercicio 1.** (a) Simular  $n = 10^6$  variables aleatorias independientes con parámetro 2 mediante el método de la distribución inversa.

(b) Dibujar un histograma con los datos obtenidos.

(c) Utilizar el comando `density` para realizar una estimación de la densidad exponencial de parámetro 2, y compararlo en el mismo gráfico con esta densidad.

**Ejercicio 2.** (a) Simular  $n = 10^6$  variables aleatorias independientes con distribución triangular mediante el método de la distribución inversa.

(b) Dibujar un histograma con los datos obtenidos.

(c) Utilizar el comando `density` para realizar una estimación de la densidad triangular y compararlo en el mismo gráfico con esta densidad.

(d) ¿Qué método le resulta mas eficiente: el anterior o la simulación mediante sumas de uniformes? Justifique.

**Ejercicio 3.** *Pareto distribution.* A random variable  $X$  has Pareto distribution with parameter  $\alpha > 0$ , when it has the density and distribution given, respectively, by

$$f(x) = \frac{\alpha}{x^{\alpha+1}}, \quad F(x) = 1 - x^{-\alpha}, \quad (x \geq 1).$$

(a) Determine the existence of the expectation and the variance of  $X$  in terms of the parameter  $\alpha > 0$ .

(b) Find the inverse  $F^{-1}$  and simulate a vector of Pareto distributed random variables for  $\alpha = 1/2, 3/2, 5/2$ . Choose the length of the vectors.

(c) In three pictures, plot the corresponding histograms, and to each one add the plot of the respective density.



**Ejercicio 4.** *Simulation of Cauchy random variables* A random variable  $X$  has Cauchy distribution when it has a density

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathbf{R}.$$

- (a) Compute the corresponding cumulative distribution function and its inverse.
- (b) Simulate a vector of size  $n$  of Cauchy distributed variables, plot a histogram and compare with the plot of its density.

## Capítulo 2

# Simulación de variables y vectores aleatorios

### 2.1. Simulación de variables aleatorias discretas

#### 2.1.1. Inversión

En el caso finito, el método de inversión da resultados en forma directa. Supongamos entonces que tenemos una variable aleatoria  $X$  tal que

$$\mathbf{P}(X = n) = p_n,$$

con  $p_n > 0$  y  $\sum_{n=1}^{\infty} p_n = 1$ . El siguiente algoritmo genera una variable con la distribución de  $X$ . Precisamos una variable de control  $S$  que acumula las sumas parciales de la serie de los  $p_n$ .

1. Simular  $U$  uniforme y poner  $X = 0$  y  $S = 0$ .
2. **While**  $S < U$   
 $X \leftarrow X + 1$   
 $S \leftarrow S + p_X$
3. Devolver  $X$

#### 2.1.2. Variables geométricas

En algunos casos la variable discreta puede obtenerse tomando la parte entera de una variable continua. Sea entonces  $X \sim \text{exp}(\lambda)$ , y definamos

$T = \lfloor X \rfloor$  (parte entera de  $X$ ). Tenemos

$$\begin{aligned}\mathbf{P}(T = n) &= \mathbf{P}(n \leq X \leq n + 1) = e^{-n\lambda} - e^{-(n+1)\lambda} \\ &= e^{-n\lambda}(1 - e^{-\lambda}) = (1 - p)p^n.\end{aligned}$$

Es decir,  $T$  es una variable con distribución geométrica con parámetro  $p = e^{-\lambda}$ , que toma los valores  $0, 1, \dots$

Si queremos una geométrica que tome los valores  $1, 2, \dots$ , consideramos  $U = T + 1$ .

## 2.2. El método de aceptación y rechazo

Queremos simular una variable aleatoria  $X$  con densidad  $f(x)$ , y sabemos simular una variable  $Y$  con densidad  $g(x)$ , tales que se verifica

$$\sup_{x \in \mathbf{R}} \frac{f(x)}{g(x)} \leq M.$$

(Observar que  $M > 1$ ). Entonces, simulamos vectores  $(Y_n, U_n)$  con coordenadas independientes, la segunda uniforme,<sup>1</sup> hasta que se verifique

$$U_n \leq \frac{f(Y_n)}{Mg(Y_n)},$$

y ponemos  $Z = Y_n$ .

El algoritmo entonces es:

1. Genero  $Y$  con densidad  $g$
2. Genero  $U$  uniforme
3. Si  $U \leq \frac{f(Y)}{Mg(Y)}$  ponemos  $Z = Y$  (*acepto*), en caso contrario vuelvo al paso 1 (*rechazo*).

### Proposición 2

- (a) *El algoritmo anterior se detiene con probabilidad uno.*
- (b) *La cantidad esperada de pasos que requiere el algoritmo es  $M$ .*
- (c) *La variable aleatoria  $Z$  que genera el algoritmo tiene densidad  $f(x)$ .*

---

<sup>1</sup>Cuando decimos “uniforme” nos referimos a uniforme en el intervalo  $[0, 1]$ .

*Demostración.* Sea  $T_n$  la variable que indica el primer  $n$  tal que  $f(Y_n)/Mg(Y_n) \geq U_n$ , es decir

$$T = \inf \left\{ n = 1, 2, \dots : U_n \leq \frac{f(Y_n)}{Mg(Y_n)} \right\}.$$

La variable  $T$  tiene distribución geométrica, porque la sucesión  $(Y_n, U_n)$  está formada por vectores independientes (es el primer éxito de una sucesión de ensayos independientes con dos resultados posibles), por lo tanto es finita. Eso prueba (a).

Para ver (b), calculamos

$$\begin{aligned} \mathbf{P}(T = 1) &= \mathbf{P} \left( U_1 \leq \frac{f(Y_1)}{Mg(Y_1)} \right) = \int_{\mathbf{R}} g(t) dt \int_0^{\frac{f(t)}{Mg(t)}} du \\ &= \int_{\mathbf{R}} g(t) \frac{f(t)}{Mg(t)} dt = \frac{1}{M}. \end{aligned}$$

Como  $T$  es geométrica, la cantidad media de pasos es  $\mathbf{E}T = M$ .

Para ver (c), distinguimos según el momento en el que se dé el primer éxito. Tenemos

$$\mathbf{P}(Z \leq x) = \sum_{n=1}^{\infty} \mathbf{P}(Z \leq x, T = n).$$

Calculamos, por ejemplo

$$\begin{aligned} \mathbf{P}(Z \leq x, T = 1) &= \mathbf{P} \left( Y_1 \leq x, U_1 \leq \frac{f(Y_1)}{Mg(Y_1)} \right) = \int_{-\infty}^x g(t) dt \int_0^{\frac{f(t)}{Mg(t)}} du \\ &= \int_{-\infty}^x g(t) \frac{f(t)}{Mg(t)} dt = \frac{1}{M} \int_{-\infty}^x f(t) dt = \frac{1}{M} \mathbf{P}(X \leq x). \end{aligned}$$

$$\begin{aligned} \mathbf{P}(Z \leq x, T = 2) &= \mathbf{P} \left( U_1 \geq \frac{f(Y_1)}{Mg(Y_1)} \right) \mathbf{P} \left( Y_2 \leq x, U_2 \leq \frac{f(Y_2)}{Mg(Y_2)} \right) \\ &= \left( 1 - \frac{1}{M} \right) \frac{1}{M} \mathbf{P}(X \leq x). \end{aligned}$$

Más en general

$$\begin{aligned} \mathbf{P}(Z \leq x, T = n) &= \mathbf{P} \left( U_1 \geq \frac{f(Y_1)}{Mg(Y_1)} \right)^{n-1} \mathbf{P} \left( Y_n \leq x, U_n \leq \frac{f(Y_n)}{Mg(Y_n)} \right) \\ &= \left( 1 - \frac{1}{M} \right)^{n-1} \frac{1}{M} \mathbf{P}(X \leq x). \end{aligned}$$

Si sumamos ahora

$$\begin{aligned} \mathbf{P}(Z \leq x) &= \sum_{n=1}^{\infty} \mathbf{P}(Z \leq x, T = n) \\ &= \left[ 1 + \left(1 - \frac{1}{M}\right) + \cdots + \left(1 - \frac{1}{M}\right)^{n-1} + \cdots \right] \frac{1}{M} \mathbf{P}(X \leq x) \\ &= \mathbf{P}(X \leq x), \end{aligned}$$

probando la proposición.  $\square$

*Ejemplo 4. Densidades acotadas con soporte compacto.* Si la densidad  $f(x)$  está acotada por  $M$  y tiene soporte compacto  $D$ , entonces podemos considerar la variable auxiliar  $Y$  con distribución uniforme en  $D$ . Ese es el caso de la distribución triangular.  $\square$

*Ejemplo 5.* We simulate a standard normal variable  $X$  departing from an exponential. We first simulate  $|X|$ , and after we multiply by an independent symmetric Bernoulli random variable  $S$ :

$$\mathbf{P}(S = 1) = P(S = -1) = 1/2.$$

The variable  $|X|$  has density  $f(x) = \frac{2}{\sqrt{2\pi}} e^{-x^2/2}$ . Our auxiliar variable  $Y$  has exponential density with parameter 1  $g(x) = e^{-x}$ . A calculus computation gives that

$$\max_x \frac{f(x)}{g(x)} = \sqrt{\frac{2e}{\pi}} = M,$$

so

$$\frac{f(x)}{Mg(x)} = e^{-\frac{1}{2}(x-1)^2},$$

and we are ready to simulate.

The algorithm is then

1. Let  $Y = -\log U$ , exponential with parameter 1
2. Generate a uniform random variable  $U$
3. If  $U \leq e^{-\frac{1}{2}(Y-1)^2}$  set  $|X| = Y$  (*accept*) otherwise go to step 1 (*reject*).
4. Simulate  $S$  and set  $X = S|X|$

Note:  $S$  can be simulated from an uniform  $U$  as

$$S = \begin{cases} 1, & \text{if } U \leq 1/2, \\ -1, & \text{if } U > 1/2. \end{cases}$$

□

### 2.3. Distribución uniforme multivariada

A vector  $X$  has uniform distribution on a set  $D \subset \mathbf{R}^d$  ( $Y \sim U[D]$ ) when it has a constant density in  $D$  (so  $D$  must have finite volumen). When  $D$  is bounded, the *rejection method* gives a way of simulating  $X$ .

Let us explain the method in  $\mathbf{R}^2$ . Consider  $B = [a, b] \times [c, d] \supset D$ . Then we simulate a pair of independent uniform random variables in  $[a, b]$ ,  $[c, d]$  respectively, denote them  $X = (X_1, X_2)$ .

It can be proved that if we sample uniform variables in  $B$ , the first that drops on  $D$  is uniform on  $D$ .

*Ejemplo 6. Uniform distribution in the circle.*

We want to sample a random vector with uniform distribution in a circle  $C$  of radius one and center  $(0, 0)$ . We take a square  $Q = [-1, 1] \times [-1, 1] \supset C$ . We need to sample an uniform vector in  $Q$ , and this is achieved simulating  $U_1, U_2$ , two independent uniforms, and defining

$$(V_1, V_2) = (2U_1 - 1, 2U_2 - 1).$$

The algorithm is then

1. Simulate  $U_1, U_2$  independent uniforms
2. Compute  $(V_1, V_2) = (2U_1 - 1, 2U_2 - 1)$ .
3. If  $V_1^2 + V_2^2 \leq 1$  set  $V = (V_1, V_2)$  otherwise go to step 1.

□

### 2.4. Ejercicios

**Ejercicio 5.** Implementar un algoritmo para simular una variable aleatoria  $X$  con distribución de Poisson de parámetro 1, es decir tal que

$$\mathbf{P}(X = n) = e^{-1} \frac{1}{n!}, \quad n = 1, 2, \dots$$

**Ejercicio 6.** Sea  $U$  uniforme. Calcular la distribución de  $X = \lfloor 1/U \rfloor$ .

**Ejercicio 7.** (a) Simular  $n = 10^6$  variables aleatorias independientes con distribución triangular mediante el método de aceptación-rechazo.

(b) Dibujar un histograma con los datos obtenidos.

(c) Utilizar el comando `density` para realizar una estimación de la densidad triangular y compararlo en el mismo gráfico con esta densidad.

(d) ¿Qué método de todos los empleados le resulta mas eficiente?

**Ejercicio 8.** Implement the method of acceptance and rejection described to sample a normal random variable when simulating exponential random variables. Plot a histogram to check your results.

**Ejercicio 9.** *Simulation of uniform random vectors over irregular domains.* Consider the domain  $D$  in  $\mathbf{R}^2$  defined by

$$D = \{(x, y) \in \mathbf{R}^2: 0 \leq x \leq 1, x^2 \leq y \leq \sqrt{x}\}.$$

(a) Use the rejection method to sample two-dimensional vector  $(X, Y)$  with uniform distribution on  $D$

(b) Once you have your sample, plot it to recover the approximate shape of  $D$  in  $\mathbf{R}^2$ .

(c) Compute the marginal density of  $X$ , and compare your result with the corresponding histogram.

## Capítulo 3

# Vectores gaussianos

### 3.1. Vectores gaussianos

Comenzamos recordando la definición de un vector normal (o gaussiano) multidimensional.

**Definición 1** (a) *Una variable aleatoria  $X$  tiene distribución normal con media  $\mu$  y varianza  $\sigma^2$  cuando tiene densidad dada por*

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right).$$

(b) *Dados un vector  $\mu \in \mathbf{R}^n$  y una matriz  $n \times n$  semidefinida positiva  $\Sigma$ , decimos que el vector aleatorio  $X = (X_1, \dots, X_n)$  tiene distribución normal multidimensional con parámetros  $(\mu, \Sigma)$ , cuando para todo  $\alpha \in \mathbf{R}^n$  se tiene<sup>1</sup>*

$$\mathbf{E}e^{i\alpha^t X} = \exp\left(i\alpha^t \mu - \frac{1}{2}\alpha^t \Sigma \alpha\right) \quad (3.1)$$

(c) *Decimos que el vector aleatorio  $X = (X_1, \dots, X_n)$  tiene distribución normal multidimensional estándar cuando  $\mu = 0$  y  $\Sigma = I_n$  (matriz identidad  $n \times n$ ).*

**Proposición 3** *Son equivalentes:*

- (a) *el vector aleatorio  $X = (X_1, \dots, X_n)$  tiene distribución normal multidimensional con parámetros  $\mu = \mathbf{E}X$  y  $\Sigma = \text{cov}X = (\text{cov}(X_i, X_j))_{1 \leq i, j \leq n}$ ,*

---

<sup>1</sup>Identificamos los vectores  $x$  de  $\mathbf{R}^n$  con las matrices columna, correspondientemente  $x^t$  denota la matriz traspuesta de  $x$ .



(b) para todo  $\alpha \in \mathbf{R}^n$  el vector  $\alpha^t X$  tiene distribución normal en  $\mathbf{R}$ .

*Demostración.* Vemos que (a) $\Rightarrow$ (b) se obtiene de la fórmula (3.1) y el teorema de unicidad de las funciones características. Para ver (b) $\Rightarrow$ (a), elegimos  $\alpha \in \mathbf{R}^n$  fijo y arbitrario. Tenemos  $\mathbf{E}\alpha^t X = \alpha^t \mathbf{E}X$ , y  $\text{var}(\alpha^t X) = \sum_{i,j=1}^n \alpha_i \alpha_j \text{cov}(X_i, X_j)$ . Si definimos  $\mu = \mathbf{E}X$  y  $\Sigma = (\text{cov}(X_i X_j))$  tenemos que se verifica (3.1).  $\square$

Vista la identificación de las distribuciones con sus funciones características, de la definición surge que los parámetros  $\mu$  y  $\Sigma$  caracterizan la distribución gaussiana. Mas aún, de la proposición anterior obtenemos que

$$\mu = \mathbf{E}X, \quad \Sigma = \text{cov}X = \mathbf{E}(X - \mu)(X - \mu)^t,$$

es decir,  $\mu$  es el vector de las esperanzas de las coordenadas, y  $\Sigma$  la matriz que reúne a las covarianzas de las coordenadas.

## 3.2. Simulación de vectores gaussianos

### 3.2.1. Simulación de variables gaussianas: el método de Box-Muller

Comenzamos simulando un par de variables gaussianas estándar, basándonos en el siguiente resultado.

**Proposition 1** Sean  $U, V$  dos variables uniformes independientes. Entonces

$$(X, Y) = (\sqrt{-2 \log U} \sin(2\pi V), \sqrt{-2 \log U} \cos(2\pi V)),$$

tiene la distribución de un vector bi-variado normal estándar, es decir

$$(X, Y) \sim \mathcal{N}\left((0, 0), \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right).$$

*Demostración.* Para un conjunto  $A \subset \mathbf{R}^2$ , tenemos

$$\mathbf{P}((X, Y) \in A) = \iint_{(\sqrt{-2 \log u} \sin(2\pi v), \sqrt{-2 \log u} \cos(2\pi v)) \in A} dudv.$$

Hacemos un cambio de variable, de acuerdo a

$$x = \sqrt{-2 \log u} \sin(2\pi v), \quad y = \sqrt{-2 \log u} \cos(2\pi v).$$

Calculamos ahora  $u, v$  en función de  $x, y$ . Obtenemos

$$u = e^{\frac{1}{2}(x^2+y^2)}, \quad v = \frac{1}{2\pi} \arctan(y/x).$$

El Jacobiano del cambio de variable es

$$J(x, y) = \begin{bmatrix} e^{\frac{1}{2}(x^2+y^2)}(-x) & e^{\frac{1}{2}(x^2+y^2)}(-y) \\ \frac{1}{2\pi(1+(y/x)^2)}(-y/x^2) & \frac{1}{2\pi(1+(y/x)^2)}(1/x) \end{bmatrix} = \frac{1}{2\pi} e^{\frac{1}{2}(x^2+y^2)}.$$

Entonces

$$\begin{aligned} \mathbf{P}((X, Y) \in A) &= \iint_{(\sqrt{-\log u} \sin(2\pi v), \sqrt{-\log u} \cos(2\pi v)) \in A} dudv \\ &= \iint_{(x, y) \in A} \frac{1}{2\pi} e^{\frac{1}{2}(x^2+y^2)} dx dy, \end{aligned}$$

que prueba que el vector  $(X, Y)$  es normal estándar, dado que el integrando es la correspondiente densidad.  $\square$

### 3.2.2. Simulation of bi-dimensional Gaussian vectors

We want to simulate a bi-dimensional gaussian vector with mean  $(\mu_1, \mu_2)$  and covariance matrix

$$\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \quad (3.2)$$

It is simple to check that if  $(Z_1, Z_2)$  are independent standard normal, then

$$\begin{aligned} X_1 &= \mu_1 + \sigma_1 Z_1, \\ X_2 &= \mu_2 + \sigma_2(\rho Z_1 + \sqrt{1-\rho^2} Z_2) \end{aligned} \quad (3.3)$$

has the desired properties.

### 3.2.3. Simulación de vectores gaussianos: el método de Cholesky

Un vector aleatorio normal  $X$  está caracterizado por su esperanza  $\mu$  y su matriz de covarianza  $\Sigma$ , que es una matriz semi-definida positiva. Decimos que el vector normal es *estándar* cuando  $\Sigma$  es la matriz identidad.

Cuando  $X$  es no degenerado, tenemos  $\det(\Sigma) \neq 0$ , resultando una matriz definida positiva. El siguiente resultado fundamenta el método de simulación de Cholesky.

**Proposición 4** Dado el vector gaussiano  $X$  en  $\mathbf{R}^n$  con parametros  $(\mu, \Sigma)$ , si  $\det \Sigma \neq 0$  existe una matriz  $A$  de tamaño  $n \times n$  tal que

$$X = \mu + AZ, \quad (3.4)$$

donde la igualdad es en distribución, y  $Z$  es un vector normal estándar en  $\mathbf{R}^n$ .

*Demostración.* Como  $\Sigma$  es definida positiva es diagonalizable en una base ortonormal. Es decir, existe una matriz de pasaje ortogonal  $P$  y una matriz diagonal  $D$  con diagonal estrictamente positiva tal que  $\Sigma = P^t D P$ . Tomando  $\sqrt{D}$  la matriz diagonal con entradas diagonales correspondientes a las raíces cuadradas de las entradas de  $D$ . Podemos escribir entonces, con  $A = \sqrt{D} R$ , que  $\Sigma = A^t A$ . Si  $Z$  es un vector con coordenadas independientes normales es un vector normal estándar, y se verifica (3.4).  $\square$

Observamos ahora que, si  $Z$  es un vector is a standard normal vector in  $\mathbf{R}^d$ , (i.e.  $Z \sim \mathcal{N}(0, I_n)$  donde  $I_n$  es la matriz identidad de tamaño  $n$ ), tenemos que

$$X = \mu + A^t Z \sim \mathcal{N}(\mu, \Sigma).$$

De hecho,  $\mathbf{E}X = \mu + A^t \mathbf{E}Z = \mu$ , y

$$\mathbf{cov}X = \mathbf{E}(A^t Z)(A^t Z)^t = A^t \mathbf{E}Z Z^t A = A^t I_n A = \Sigma.$$

Entonces, como sabemos simular  $Z$  (por el método de Box-Muller, por ejemplo) podemos también simular  $X$ .

**Ejercicio 10.** *Multivariate normal vectors.* Use the Cholesky method to sample a vector  $(X, Y, Z)$  normally distributed as

$$(X, Y, Z) \sim \mathbf{N} \left( \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1 & -1/2 \\ 1/3 & -1/2 & 1 \end{pmatrix} \right)$$

Check your results computing the empirical variances and covariances.

**Ejercicio 11.** Check that, if  $(Z_1, Z_2)$  are independent standard normal, the random vector defined in (3.3) has expectation  $(\mu_1, \mu_2)$  and covariance matrix given in (3.2).

**Ejercicio 12.** *Bivariate normal vectors* (a) Produce a sample of bi-dimensional vectors  $(X, Y)$  normally distributed, according to

$$(X, Y) \sim \mathbf{N} \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0,8 \\ 0,8 & 1 \end{pmatrix} \right).$$

Compute the empirical correlation to check your results.

(b) Plot your sample also to check your results.

## Capítulo 4

# Procesos gaussianos

### 4.1. Procesos estocásticos y su simulación

Estudiaremos *procesos estocásticos con tiempo continuo*, es decir, familias de variables aleatorias  $X = (X_t)_{t \geq 0}$ , definidas en un espacio de probabilidad común  $(\Omega, \mathcal{F}, \mathbf{P})$ , cuyo índice  $t$  toma valores en la semirrecta real no negativa  $[0, \infty)$ . Una manera alternativa de ver un proceso aleatorio de tiempo continuo es considerar fijo cada suceso elemental  $\omega$  del espacio de estados  $\Omega$ , obteniéndose una función  $X(t, \omega)$ <sup>1</sup>. Una tal función se denomina *trayectoria* del proceso, y aquí supondremos que todas las trayectorias de los procesos considerados son continuas por la derecha con límites a la izquierda.

Diremos que un proceso estocástico  $X$  tiene *incrementos independientes* cuando para cualquier elección de índices  $0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \dots \leq s_n < t_n$ , las variables aleatorias

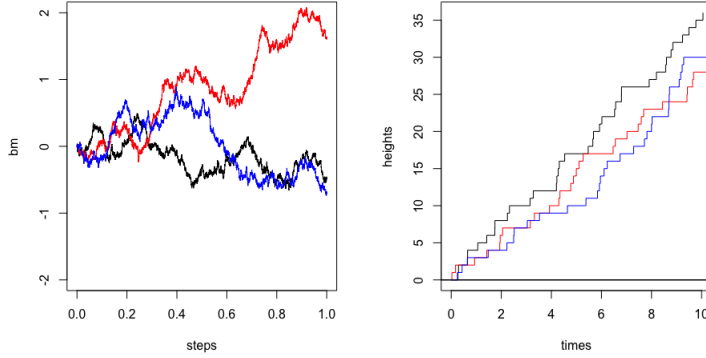
$$X_{t_1} - X_{s_1}, X_{t_2} - X_{s_2}, \dots, X_{t_n} - X_{s_n} \text{ son mutuamente independientes. } (4.1)$$

Un proceso estocástico tiene *incrementos estacionarios* (también decimos *incrementos homogéneos en el tiempo*) cuando cualesquiera sean  $t \geq 0$ , y  $h \geq 0$

$$\text{la distribución de } X_{t+h} - X_t \text{ es idéntica a la de } X_h. \quad (4.2)$$

---

<sup>1</sup>las variables aleatorias de un proceso estocástico  $X$  serán notadas indistintamente mediante  $X_t$ ,  $X_t(\omega)$  o  $X(t, \omega)$ .



Vemos dos ejemplos de trayectorias de procesos estocásticos. A la izquierda tenemos trayectorias de un proceso con trayectorias continuas, a la derecha un proceso con trayectorias discontinuas, o *saltos*.

When all the finite distributions are gaussian vectors, the process is called a *Gaussian process*. As a gaussian vector is determined by its mean and covariance matrix, a gaussian process is determined by the two functions

$$\begin{aligned}\mu(t) &= \mathbf{E}(X(t)), \quad \text{for all } t \in I \\ r(s, t) &= \mathbf{cov}(X(s), X(t)), \quad \text{for all } (s, t) \in I \times I.\end{aligned}$$

In particular, the variance of a coordinate is given by

$$\text{var}(X(t)) = r(t, t).$$

As we are able to simulate random gaussian vectors with the Cholesky method, to simulate a trajectory of a continuous gaussian process  $\{X(t): 0 \leq t \leq T\}$  with mean  $\mu(t)$  and covariance  $r(s, t)$ , we proceed in the following way:

- Select  $n$  and time grid  $I_n = \{Ti/n: i = 0, \dots, n\}$
- Simulate a gaussian vector  $(X_0, \dots, X_n)$  with
  - mean  $(\mu(Ti/n): i = 0, \dots, n)$
  - covariance matrix  $\Sigma_n = (r(Ti/n, Tj/n): i, j = 0, \dots, n)$

As the process is continuous, we can plot  $X(t)$  considering the polygonal chain defined at  $I_n$ , by

$$X(t_i) = X_i, \quad i = 0, \dots, n.$$

## 4.2. Proceso de Wiener

En esta sección estudiaremos el proceso de Wiener, o Movimiento Browniano. Se trata de un modelo probabilístico para la evolución temporal de un sistema sujeto a cambios instantáneos. Su relevancia radica en dos hechos: por una parte este modelo juega un rol central en toda la teoría de los procesos estocásticos (cálculo de Itô, difusiones, martingalas, procesos de incrementos independientes, procesos autosimilares) y por otro, encuentra aplicaciones en las mas diversas ramas del conocimiento (física, biología, economía, etc.).

La denominación de este proceso se debe a las investigaciones del botánico inglés Robert Brown, que en 1828 observó y describió el movimiento caótico de una partícula de polen suspendida en agua, destacando la naturaleza física (y no biológica) del movimiento observado. Norbert Wiener, en 1923, construyó el primer modelo matemático para la descripción de dicha dinámica, similar a la que presentamos aquí. Comenzamos con la definición.

**Definición 2** Diremos que un proceso estocástico  $W = (W(t))_{t \geq 0}$  es un proceso de Wiener o un movimiento Browniano si cumple las siguientes propiedades

- (a)  $\mathbf{P}(W(0) = 0) = 1$ .
- (b)  $W$  tiene trayectorias continuas.
- (c)  $W$  tiene incrementos independientes y estacionarios.
- (d) Para cada  $t > 0$  la variable aleatoria  $W(t)$  tiene distribución normal con parámetros  $(0, t)$ .

Comencemos observando que las propiedades (c) y (d) son equivalentes a la propiedad

- (e) Para todos  $0 = t_0 \leq t_1 \leq t_2 \leq \dots \leq t_n$  el vector aleatorio

$$V = (W(t_1), W(t_2), \dots, W(t_n))$$

tiene distribución multidimensional normal centrada, con matriz de covarianzas  $\Gamma_V = ((t_i \wedge t_j))_{(i,j=1,\dots,n)}$ .

que a su vez, es equivalente a la condición

(f) Para todos  $0 = t_0 \leq t_1 \leq t_2 \leq \dots \leq t_n$  el vector aleatorio

$$U = (W(t_1), W(t_2) - W(t_1), \dots, W(t_n) - W(t_{n-1}))$$

es gaussiano centrado, con coordenadas independientes y varianzas  $\text{var}(W(t_k) - W(t_{k-1})) = t_k - t_{k-1}$  ( $k = 1, \dots, n$ ), es decir, con matriz de covarianza  $\Gamma_U = \mathbf{diag}(t_1, t_2 - t_1, \dots, t_n - t_{n-1})$ .

La equivalencia de (e) y (f) resulta de la relación  $U = AV$ , donde

$$A = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ -1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & -1 & 1 \end{bmatrix}$$

Veamos por ejemplo (f) $\Rightarrow$ (e). Para  $\alpha \in \mathbf{R}^n$ , tenemos  $m_U = \mathbf{E}U = \mathbf{E}AV = \mathbf{A}\mathbf{E}V = 0$ . Verificamos multiplicando que  $\Gamma_U = \mathbf{A}\Gamma_V\mathbf{A}^t$ , entonces

$$\Gamma_U = \Gamma_{AV} = \mathbf{E}(AV(AV)^t) = \mathbf{A}\Gamma_V\mathbf{A}^t.$$

Para (e) $\Rightarrow$ (f) utilizamos que la matriz  $A$  es invertible.

Veamos ahora que (f) $\Rightarrow$ (c)+(d). Efectivamente (f) implica que los incrementos  $W(t_{i+1}) - W(t_i)$  son centrados e independientes, con varianza igual a  $t_{i+1} - t_i$ , es decir, homogéneos, concluyendo (c). Como además son gaussianos, obtenemos (d). Si (c) y (d) son ciertas, un vector con coordenadas independientes y gaussianas es un vector gaussiano con coordenadas independientes (basta verificar la definición de vector gaussiano, que resulta centrado y con la matriz de covarianza correspondiente).

#### 4.2.1. Simulation of Brownian motion

The independence of the increments gives a second way of simulating Brownian motion (the first is the Cholesky method). To simulate  $W$  over a time interval  $[0, T]$ , we use the same mesh  $I_n = \{Ti/n : i = 0, \dots, n\}$  but simulate the increments of the process and accumulate them.

- Set  $W(0) = 0$ , choose  $n$
- for  $i = 1$  in  $1:n$  simulate  $\Delta = \mathbf{N}(0, T/n)$  and set  $W(iT/n) = W((i-1)T/n) + \Delta$ .

In this way we obtain a discretization  $\{W(iT/n) : i = 0, \dots, n\}$  of a trajectory of a Brownian motion. In R, we can use the command `cumsum` :

```
n<-1e4
t<-1
bm<-c(0, cumsum(rnorm(n,0,sqrt(t/n))))
```



### 4.3. Ejercicios

**Ejercicio 13.** *Brownian motion.* Plot a large number of trajectories of a Brownian motion  $\{W(t)\}$  in an interval  $[0, 1]$ . Observe the behavior close to  $t = 0$ . On the same plot, add the curves  $y(t) = \sqrt{2t}$  with a different color. Find a good scale of the axis, conjecture the value of:

$$\ell = \lim_{t \rightarrow 0} \frac{W(t)}{\sqrt{2t}}, a.s.$$

(As a consequence of the *Law of the iterated logarithm* we get that  $\ell = 0$ .)

**Ejercicio 14.** *Transformations of Brownian motion.* Consider a Brownian motion  $\{W(t): t \geq 0\}$ .

(a) *Scaling property.* Prove that, given  $c > 0$ , the process

$$X(t) = \frac{1}{\sqrt{c}}W(ct)$$

is also a Brownian motion.

(b) *Translation.* Prove that, given  $a > 0$ , the process

$$X(t) = W(a + t) - W(a)$$

is also a Brownian motion.

(c) *Time inversion.* Assuming that the process

$$X(t) = \begin{cases} 0, & \text{when } t = 0 \\ tW(1/t), & \text{when } t > 0 \end{cases}$$

has continuous paths, prove that it is also a Brownian motion.

## Capítulo 5

# Otros procesos gaussianos

### 5.1. Sobre las trayectorias del movimiento Browniano

Una de las características más interesantes del proceso de Wiener es la naturaleza de sus trayectorias. Consideramos un intervalo  $[0, T]$ , y una sucesión de particiones

$$\lambda^n = \{0 = t_0^n < t_1^n < \dots < t_{k(n)}^n = T\} \quad (n = 1, 2, \dots), \quad (5.1)$$

cuya norma  $|\lambda^n| = \max\{t_k^n - t_{k-1}^n : k = 1, \dots, k(n)\}$  tiende a cero si  $n \rightarrow \infty$ , y tales que se verifica  $\lambda^n \subset \lambda^{n+1}$  ( $n = 1, 2, \dots$ ) es decir, cada partición se obtiene de la anterior agregando puntos. Para una función  $f: [0, T] \rightarrow \mathbf{R}$  con derivada  $f'(t)$  continua en  $[0, T]$ , cuando  $n \rightarrow \infty$ , tenemos

$$\sum_{k=1}^{k(n)} |f(t_k^n) - f(t_{k-1}^n)| = \sum_{k=1}^{k(n)} |f'(\theta_k^n)|(t_k^n - t_{k-1}^n) \rightarrow \int_0^T |f'(t)| dt.$$

(Aquí aplicamos el teorema del valor medio,  $\theta_k^n \in [t_{k-1}^n, t_k^n]$  para cada  $n$  y cada  $k$ .) El límite obtenido es la *variación* de la función  $f(t)$  en el intervalo  $[0, T]$ . En forma similar, si  $|f'(t)| \leq M$  ( $0 \leq t \leq T$ ), tenemos

$$\sum_{k=1}^{k(n)} (f(t_k^n) - f(t_{k-1}^n))^2 \leq \sum_{k=1}^{k(n)} M^2 (t_k^n - t_{k-1}^n)^2 \leq M^2 T |\lambda^n| \rightarrow 0,$$

si  $n \rightarrow \infty$ , y decimos, que la *variación cuadrática* de la función  $f(t)$  en el intervalo  $[0, T]$  es nula.

El siguiente teorema muestra que las trayectorias de un proceso de Wiener presentan un comportamiento diferente: su variación en un intervalo  $[0, T]$  no existe, es infinita; y su variación cuadrática en un intervalo  $[0, T]$  es igual a  $T$ .

**Teorema 1 (Propiedades de las trayectorias)**

Consideremos un proceso de Wiener  $\{W_t\}$  y una sucesión creciente de particiones  $\{\lambda^n\}$  como en (5.1), cuyas normas  $\{|\lambda^n|\}$  tienden a cero si  $n \rightarrow \infty$ . Se verifica

$$V_n = \sum_{k=1}^{k(n)} |W_{t_k^n} - W_{t_{k-1}^n}| \rightarrow \infty \quad (n \rightarrow \infty) \quad c.s. \quad (5.2)$$

$$Q_n = \sum_{k=1}^{k(n)} (W_{t_k^n} - W_{t_{k-1}^n})^2 \rightarrow T \quad (n \rightarrow \infty) \quad \text{en media cuadrática.} \quad (5.3)$$

Además, si  $\sum_{n=1}^{\infty} |\lambda^n| < \infty$ , la convergencia en (5.3) es casi segura.

*Demostración.* Comencemos con la demostración de (5.2). En primer lugar observemos que como las particiones son crecientes, aplicando la propiedad triangular, se obtiene que  $V_n \leq V_{n+1}$  ( $n = 1, 2, \dots$ ), es decir, la sucesión  $\{V_n\}$  es no decreciente, casi seguramente. Queremos demostrar que  $\mathbf{P}(V_n \rightarrow \infty) = 1$ . Esto es equivalente a demostrar que dado  $K > 0$  arbitrario, se verifica  $\mathbf{P}(\cup_{n=1}^{\infty} \cap_{m=n}^{\infty} \{V_m(\omega) > K\}) = 1$ . Esta igualdad, tomando complementos, es equivalente a  $\mathbf{P}(\cap_{n=1}^{\infty} \cup_{m=n}^{\infty} \{V_m(\omega) \leq K\}) = 0$ . Como la sucesión  $\{V_n\}$  es no decreciente, tenemos

$$\mathbf{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \{V_m(\omega) \leq K\}\right) \leq \mathbf{P}\left(\bigcup_{m=n}^{\infty} \{V_m(\omega) \leq K\}\right) = \mathbf{P}(V_n(\omega) \leq K).$$

En conclusión, para demostrar (5.2), verificamos que  $\mathbf{P}(V_n(\omega) \leq K) \rightarrow 0$  ( $n \rightarrow \infty$ ). Como  $\{W_t\}$  tiene incrementos independientes, aplicando la fórmula (??), obtenemos

$$\text{var}V_n = \sum_{k=1}^{k(n)} \text{var}|W_{t_k^n} - W_{t_{k-1}^n}| \leq \sum_{k=1}^{k(n)} \mathbf{E}|W_{t_k^n} - W_{t_{k-1}^n}|^2 = T.$$

Por otra parte, tenemos  $\mathbf{E}|W_{t_k^n} - W_{t_{k-1}^n}| = \sqrt{t_k^n - t_{k-1}^n} \mathbf{E}|Z|$ , donde  $Z$  es una variable aleatoria con distribución normal estándar, y  $\mathbf{E}|Z| = \sqrt{2/\pi}$ .

Entonces, como  $\sqrt{t_k^n - t_{k-1}^n} \geq (t_k^n - t_{k-1}^n)/\sqrt{|\lambda^n|}$ , tenemos

$$\mathbf{E}V_n = \sum_{k=1}^{k(n)} \mathbf{E}|W_{t_k^n} - W_{t_{k-1}^n}| = \mathbf{E}|Z| \sum_{k=1}^{k(n)} \sqrt{t_k^n - t_{k-1}^n} \geq \frac{T\mathbf{E}|Z|}{\sqrt{|\lambda^n|}} \rightarrow \infty,$$

si  $n \rightarrow \infty$ . Para  $n$  suficientemente grande se verifica  $\mathbf{E}V_n > K$ , y aplicando la desigualdad de Chebishev, obtenemos

$$\begin{aligned} \mathbf{P}(V_n \leq K) &\leq \mathbf{P}(|V_n - \mathbf{E}V_n| \geq \mathbf{E}V_n - K) \\ &\leq \frac{1}{(\mathbf{E}V_n - K)^2} \text{var}V_n \rightarrow 0 \quad (n \rightarrow \infty), \end{aligned}$$

concluyendo la demostración de (5.2).

Veamos ahora la demostración de (5.3). Las variables aleatorias  $Y_n = (W_{t_k^n} - W_{t_{k-1}^n})^2 - (t_k^n - t_{k-1}^n)$  ( $k = 1, \dots, k(n)$ ) son independientes, y verifican

$$\mathbf{E}Y_n = 0, \quad \text{var}Y_n = (t_k^n - t_{k-1}^n)^2 \mathbf{E}(Z^2 - 1)^2,$$

si  $Z$  designa una variable aleatoria con distribución normal estándar. Como en este caso los momentos de orden dos coinciden con las varianzas, tenemos

$$\begin{aligned} \mathbf{E}(Q_n - T)^2 &= \mathbf{E}\left(\sum_{k=1}^{k(n)} Y_n\right)^2 = \sum_{k=1}^{k(n)} \mathbf{E}(Y_n)^2 = \sum_{k=1}^{k(n)} (t_k^n - t_{k-1}^n)^2 \mathbf{E}(Z^2 - 1)^2 \\ &\leq T|\lambda^n| \mathbf{E}(Z^2 - 1)^2 \rightarrow 0 \quad (n \rightarrow \infty), \end{aligned} \quad (5.4)$$

obteniendo la convergencia en media cuadrática.

La convergencia casi segura, bajo el supuesto  $\sum_{n=1}^{\infty} |\lambda^n| < \infty$ , se obtiene de la siguiente forma. Sabemos que  $\mathbf{P}(Q_n \rightarrow T) = 1$  si (y solo si) para todo  $\varepsilon > 0$ , se cumple  $\mathbf{P}\left(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \{|Q_n - T| > \varepsilon\}\right) = 0$ . Tenemos

$$\begin{aligned} \mathbf{P}\left(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \{|Q_n - T| > \varepsilon\}\right) &\leq \mathbf{P}\left(\bigcup_{n=m}^{\infty} \{|Q_n - T| > \varepsilon\}\right) \\ &\leq \sum_{n=m}^{\infty} \mathbf{P}(|Q_n - T| > \varepsilon) \\ &\leq \frac{1}{\varepsilon^2} \sum_{n=m}^{\infty} \mathbf{E}|Q_n - T|^2 \leq t\mathbf{E}(Z^2 - 1) \sum_{n=m}^{\infty} |\lambda^n| \rightarrow 0, \end{aligned}$$

si  $m \rightarrow \infty$ , donde utilizamos la acotación obtenida en (5.4). De aquí se obtiene la convergencia casi segura<sup>1</sup>, concluyendo la demostración.  $\square$

<sup>1</sup>Hemos demostrado que  $\sum_{n=1}^{\infty} \mathbf{P}(\mathbf{A}_n) < \infty$  implica  $\mathbf{P}(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \mathbf{A}_n) = 0$ , que es la primer parte del llamado *lema de Borel–Cantelli*.

## 5.2. Geometric Brownian Motion

Given a Brownian motion  $W$ , the process

$$S(t) = S(0)e^{\sigma W(t) + \mu t}, \quad S(0) \text{ a positive constant.}$$

is named a *Geometric Brownian motion (GBM)*. It is used in Black-Scholes model of a financial market, that has to investment possibilities:

- A deterministic *savings account*, that evolves according to

$$B(t) = B(0)e^{rt}$$

- A stock, that has a random evolution, according to

$$S(t) = S(0)e^{\sigma W(t) + \mu t}.$$

## 5.3. Brownian bridge

A *Brownian bridge*  $\{R(t) : 0 \leq t \leq 1\}$  is a continuous gaussian stochastic process defined by

$$\mu(t) = 0, \quad r(t, s) = \min(s, t) - st.$$

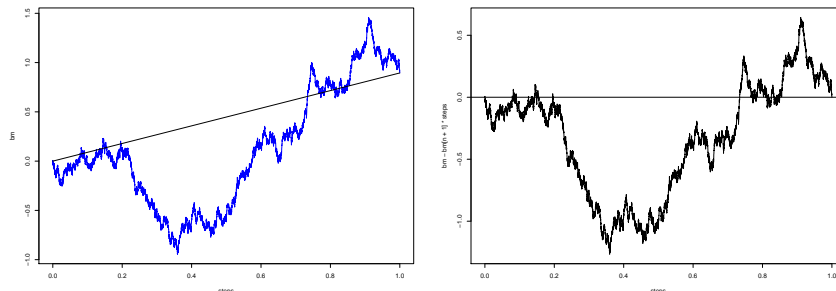
Observe that, as  $\text{var}(R(0)) = \text{var}(R(1)) = 0$ , we have

$$\mathbf{P}(R(0) = R(1) = 0).$$

A key property follows. If  $\{W(t) : 0 \leq t \leq 1\}$  is a Brownian motion<sup>2</sup> in  $[0, 1]$ , then

$$R(t) = W(t) - tW(1)$$

is a Brownian bridge. This property gives a way of simulating trajectories of the Brownian bridge:



<sup>2</sup>Brownian motion in intervals  $[a, b]$  is defined similarly.

On the left we plot a trajectory of a Brownian motion and also the line  $tW(1)$ . On the left we plot the difference  $W(t) - tW(1)$ , that is a trajectory of a Brownian bridge.

### 5.3.1. Kolmogorov test

A key role in statistics is played by the random variable

$$K = \max_{0 \leq t \leq 1} |R(t)|$$

To explain this result, let us define the empirical distribution function of a sample of iid rvs  $X_1, \dots, X_n$  as

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq x\}}$$

i.e. we count how many values  $X_i$  are smaller or equal than  $x$  (and divide by  $n$ ).

**Teorema 1** *Assume that the random sample of iid rv has continuous distribution  $F$ . Then*

$$K_n := \sqrt{n} \sup_{x \in \mathbf{R}} |F_n(x) - F(x)| \xrightarrow{d} K.$$

The distribution of  $K$  is explicit, through the formula

$$\mathbf{P}(K \geq x) = 2 \sum_{k=1}^{\infty} (-1)^{k-1} e^{-2k^2 x^2}$$

and notably, does not depend on  $F$ .

### 5.3.2. Application

If we suspect that our number generator does not generate, for instance, uniform random variables, we perform a *hypothesis test* in the following way.

- We produce a sample  $(X_1, \dots, X_n)$  and compute  $K_n$ .
- We construct a confidence interval for  $K$  with confidence  $1 - \alpha$ , of the form

$$\mathbf{P}(K \geq k_0) = \alpha.$$

- If our computed value  $K_n$  is larger than  $k_0$ , we reject the hypothesis that the random variables are uniform.

Observe that what is tested is the distribution, but the independence is assumed to hold.

## 5.4. Stationary gaussian processes

A gaussian process defined in  $I = \mathbf{R}$ , is named stationary, when,

$$\mu(t+h) = \mu(t), \quad r(s+h, t+h) = r(s, t), \quad \text{for all } h \in \mathbf{R}$$

(the same to say is that  $r(t, t+h)$  does not depend on  $t$ ). This means that  $\mu(t)$  is constant, and  $r(s, t)$  is a function that only depends on  $t-s$ . We will use the notation

$$r(s, t) = \rho(\tau),$$

where  $\tau = t-s$ . Observe that  $\rho(\tau) = \rho(-\tau)$ . We denote

$$\rho(h) = r(t, t+h)$$

So, a gaussian stationary process is determined once we give its (constant) mean  $\mu$  and covariance function  $\rho$ .

## 5.5. Proceso de Slepian

En primer lugar definimos un movimiento Browniano  $\{W(t) : t \in \mathbf{R}\}$  de la siguiente forma:

$$W(t) = \begin{cases} W^{(1)}(t), & \text{cuando } t \geq 0, \\ W^{(2)}(-t), & \text{cuando } t < 0. \end{cases}$$

donde  $W^{(i)}$ ,  $i = 1, 2$  son movimientos Brownianos independientes. El proceso de Slepian lo definimos entonces como

$$S(t) = W(t+1) - W(t), \quad t \in \mathbf{R}.$$

**Proposición 5** *El proceso  $\{S(t) : t \in \mathbf{R}\}$  es un proceso estacionario gaussiano con media  $\mu = 0$  y función de covarianza triangular, es decir*

$$\rho(\tau) = \begin{cases} 1 - |\tau|, & \text{cuando } |\tau| \leq 1, \\ 0, & \text{en otro caso.} \end{cases}$$

## 5.6. Ornstein-Uhlenbeck process

An Ornstein-Uhlenbeck (OU) process with parameters  $(a, b, \sigma)$  is a stationary gaussian process with mean  $\mu = b$  and covariance function

$$\rho(h) = \frac{\sigma^2}{2a} e^{-a|h|},$$

**Proposition 2** *Let  $W = \{W(t) : t \geq 0\}$  be a Brownian motion. Then*

$$X(t) = b + \frac{\sigma e^{-at}}{\sqrt{2a}} W(e^{2at})$$

*is an OU process with parameters  $(a, b, \sigma)$  defined for  $t \in \mathbf{R}$ .*

As  $X(t)$  is gaussian, the proof consists in the computation of

$$\mathbf{E}(X(t)) = b + \frac{\sigma e^{-at}}{\sqrt{2a}} \mathbf{E}(W(e^{2at})) = b,$$

and, for  $s < t$ ,

$$\begin{aligned} r(s, t) &= \mathbf{E}(X(s) - b)(X(t) - b) \\ &= \frac{\sigma^2 e^{-a(t+s)}}{2a} \mathbf{E}[(W(e^{2as})(W(e^{2at})))] \\ &= \frac{\sigma^2 e^{-a(t+s)}}{2a} e^{-2a \times \min(s, t)} \\ &= \frac{\sigma^2}{2a} e^{-a(t-s)}. \end{aligned}$$

Suppose we want to simulate an OU process  $X$  in an interval  $[0, T]$ . We choose  $N$  and simulate the vector

$$X(0), X(T/n), X(2T/n), \dots, X(T)$$

in the following way:

- Simulate

$$W(1), W(e^{2aT/N}), W(e^{4aT/N}), \dots, W(e^{2aT})$$

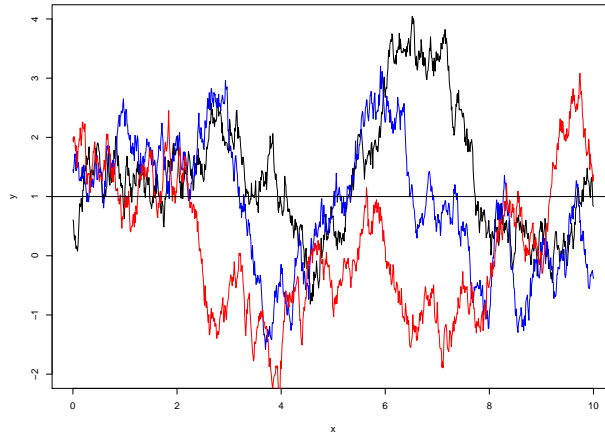
using the property of independent increments



- Set

$$X(kT/n) = b + \frac{e^{-akT/n}}{\sqrt{2a}} W(e^{2akT/N})$$

for  $k = 0, \dots, N$ .



We simulate three trajectories of the OU process in  $[0, T]$  with  $n = 1000$ ,  $a = b = 1$  and  $\sigma = 0,1$ .

## 5.7. Ejercicios

**Ejercicio 15.** Simulate several trajectories of an OU process by the Cholesky method and also using the method described in the notes. Plot 10 trajectories with each method.

**Ejercicio 16.** *Brownian motion: Properties of the trajectories.* Given a BM  $\{W(t): t \geq 0\}$  and  $n \in \mathbf{N}$  we define the discrete time random processes, depending on  $q = 1, 2, 3$ , by

$$V_n^q(k) = \sum_{i=1}^k |W(i/2^n) - W((i-1)/2^n)|^q, \quad \text{for } k = 0, \dots, 2^n.$$

(a) We want to establish the limit behavior for  $q = 1, 2, 3$ :

$$\lim_{n \rightarrow \infty} V_n^q(2^n) \quad \text{when } n \rightarrow \infty.$$

For this purpose, plot trajectories of the processes  $\{V_n^q(k): 0 \leq k \leq 2^n\}$  for  $q = 1, 2, 3$  and  $5 \leq n \leq 12$ . To see all the values, use first the command `plot` with `ylim=c(0,10)` and then use the command `lines`.

(b) Based on your simulations, conjecture the limits for  $q = 1, 2, 3$  and compute the expectations  $\mathbf{E}(V_n^q(2^n))$  to support your conjecture.

**Ejercicio 17.** *Brownian bridge.*

(a) Run a simulation of the Brownian bridge to determine the density of the random variable

$$K = \max_{0 \leq t \leq 1} |R(t)|.$$

and estimate  $\mathbf{E}(K)$  and  $\text{var}(K)$ .

(d) Determine the level  $k_0$  such that

$$\mathbf{P}(K \geq k_0) = 0,05.$$

(Use the command `quantile`. True value:  $k_0 = 1,36$ ).

**Ejercicio 18.** *Testing the random number generator.* Kolmogorov Theorem on the empirical distribution states that

$$K_n := \sqrt{n} \sup_{x \in \mathbf{R}} |F_n(x) - F(x)| \xrightarrow{d} K.$$

We assume that our random number generator `runif` gives independent and identically distributed random variables, we want to test whether this random variables are uniform. To perform this test:

(a) Given a sample  $X_1, \dots, X_n$ , prove that

$$k_n = \sqrt{n} \sup_{0 \leq x \leq 1} |F_n(x) - x| = \sqrt{n} \max\{|X_k - (k-1)/n|, |X_k - k/n|: k = 1, \dots, n\}.$$

(b) Given a sample generated by `runif(n)`, compute the statistic  $k_n$

(c) Plot in the same window, for  $0 \leq x \leq 1$ , the functions  $F_n(x)$  for the obtained sample, and the uniform distribution  $F(x) = x$ ,  $0 \leq x \leq 1$ .

(d) If this statistic is larger than  $k_0$  computed in the previous exercise, you have ground to suspect that your generator does not produce truly uniform variables, with a confidence of 95 %.

(e) You can check your results with `ks.test`.

**Ejercicio 19.** (a) Demostrar que el polinomio trigonométrico de la forma

$$X(t) = \sum_{k=1}^n a_n \cos(2\pi kt) + a_n \operatorname{sen}(2\pi kt)$$

donde  $(a_n)_{n \geq 1}, (b_n)_{n \geq 1}$  son sucesiones de variables aleatorias normales estándar es un proceso estacionario gaussiano. Calcular la función  $\rho$  de covarianza.

(e) Simular varias trayectorias de este proceso para  $n = 5$ .