



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

LICENCIATURA EN MATEMÁTICAS
FACULTAD DE CIENCIAS

Árboles continuos y la excursión browniana
normalizada

TESIS

que para obtener el grado académico de:
Matemático

P R E S E N T A:
Andrés Muñoz Medina

TUTOR:
Dr. Gerónimo Francisco Uribe Bravo



2010



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Índice general

Introducción	3
Capítulo 1. Preliminares	5
1. Propiedad de Markov	5
2. Propiedad de Markov fuerte del movimiento browniano	15
Capítulo 2. Excursiones en el movimiento browniano	21
1. Propiedades del movimiento browniano	21
2. El movimiento browniano y su mínimo	25
3. Excursión normalizada del movimiento browniano.	28
Capítulo 3. El árbol aleatorio continuo	35
1. Árboles discretos	35
2. Árboles reales	36
3. El árbol aleatorio continuo	48
Capítulo 4. Construcciones alternas del CRT	55
1. Creciendo el árbol binario uniforme en n hojas	55
2. Codificación de los árboles finitos	59
3. El árbol continuo de Aldous	63
Capítulo 5. Cuadrangulaciones	69
1. Mapas, cuadrangulaciones, árboles y móviles	69
2. Una biyección entre cuadrangulaciones y árboles bien etiquetados	73
3. Una biyección entre cuadrangulaciones y móviles	75
Conclusiones	79
Bibliografía	81

Introducción

En 1991 David Aldous introdujo en [Ald91] el Árbol Aleatorio Continuo a la teoría de gráficas aleatorias, y desde ese momento se han publicado numerosos artículos relacionando el CRT con diversos conceptos probabilísticos tales como la excursión browniana [Chu76], procesos de Rayleigh [EPW06] y super procesos [LG05] entre otros; sin embargo todos estos trabajos recurren a conceptos de probabilidad más avanzados que los que se estudian normalmente a nivel licenciatura; como teoría de procesos de Feller y teoría de excursiones por mencionar sólo algunos. El propósito de este trabajo es presentar resultados tomados de los artículos mencionados anteriormente, pero con pruebas un poco más intuitivas y más claras a mi parecer. En esta tesis se hacen estudios sobre 2 objetos probabilísticos, a saber la excursión browniana normalizada y el CRT así como las conexiones que hay entre ellos. Para poder establecer éstas, analizamos someramente la teoría sobre árboles reales y árboles combinatorios.

El primer capítulo presenta los preliminares para la tesis, donde se habla de procesos de Markov a tiempo continuo. El objetivo de este capítulo no es de ninguna manera el ahondar en conceptos de procesos de Markov, sino el dar una base teórica bien fundamentada en teoría de la medida sobre estos procesos y establecer una conexión clara con las cadenas de Markov para un entendimiento íntegro del siguiente capítulo donde se trabaja con 2 procesos de Markov: el movimiento browniano y la excursion browniana normalizada. De hecho, el logro principal del segundo capítulo es ser capaz de definir la excursión browniana normalizada sin necesidad de recurrir a teoría de excursiones y dar una construcción obtenida mediante transformaciones y condicionamientos a las trayectorias del browniano.

Los árboles discretos ordenados y los árboles reales son discutidos en el 3er capítulo y se hace un énfasis especial en la codificación de árboles reales mediante funciones continuas. Ésta codificación nos permite definir el CRT cómo

lo hace Le Gall en [LG05], a saber como un cociente del $[0, 1]$, utilizando la excursión browniana normalizada. Ésto justifica el hecho de haber dedicado un capítulo completo a la excursión. Obtenemos además las distribuciones finito dimensionales del CRT con un cálculo simple y bastante directo esperando así que queden claras estas distribuciones. Al terminar este capítulo hacemos un estudio breve de los artículos de Aldous ([Ald91], [Ald93]), construimos el CRT como originalmente se hizo y probamos que el CRT construido en el capítulo anterior y el CRT de Aldous son equivalentes en cierta manera, es decir, veremos que si e es una excursión browniana entonces el proceso $2e$ codifica al CRT de Aldous. También describimos un algoritmo para la construcción de los muestreos del CRT.

Por último nos desviamos un poco del camino de los árboles y definimos lo que son las cuadrangulaciones en el último capítulo, así como las relaciones que existen entre éstas y ciertos tipos de árboles combinatorios. Concluimos con una idea de lo que se está investigando actualmente en esta área de la probabilidad.

Preliminares

Este capítulo se enfocará en dar las bases necesarias de procesos de Markov para comprender los cálculos realizados en la demostración de las propiedades de dos procesos en particular, a saber el movimiento browniano y la excursión browniana normalizada.

1. Propiedad de Markov

Antes de iniciar con la teoría de procesos de Markov daremos unas definiciones básicas.

Definición 1.1. *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, a un conjunto de sub σ -álgebras $(\mathcal{F}_t)_{t \geq 0}$ que cumplen:*

$$\mathcal{F}_s \subset \mathcal{F}_t \text{ si } s \leq t$$

se le conoce como filtración de Ω .

A un espacio de probabilidad filtrado lo denotaremos por $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$

Definición 1.2. *Un proceso estocástico es un conjunto de variables aleatorias $(X_t)_{t \in T}$ definidas en el mismo espacio de probabilidad e indexadas por un conjunto T .*

En este trabajo T siempre será el conjunto de los naturales o el de los reales positivos. Por último diremos que un proceso definido en $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ es adaptado si la σ -álgebra generada por $(X_s)_{s \leq t}$ (la cual denotaremos por $\sigma(X_s)_{s \leq t}$) está contenida en \mathcal{F}_t . Claramente si $\mathcal{F}_t = \sigma(X_s)_{s \leq t}$ entonces el proceso será adaptado, en este caso diremos que \mathcal{F}_t es la filtración natural del proceso.

Trabajaremos con procesos estocásticos definidos en el espacio de trayectorias $\mathbb{R}^{\mathbb{R}^+}$ dotado de la sigma álgebra $\mathcal{F} = \mathcal{B}^{\mathbb{R}^+}$ y una medida de probabilidad \mathbb{P} , donde \mathcal{B} denota a los borelianos en los reales.

Definición 1.3. Una función $N : \mathbb{R} \times \mathcal{B} \rightarrow \mathbb{R}$ se llama kernel si para toda $x \in \mathbb{R}$ fija, la función $N(x, \cdot)$ es una medida en \mathcal{B} y para toda $A \in \mathcal{B}$ la función $N(\cdot, A)$ es medible. Un kernel π es llamado probabilidad de transición si para toda x , $\pi(x, \mathbb{R}) = 1$

Intuitivamente, una probabilidad de transición es una función que a cada punto x en el espacio de estados le asigna una medida, la cual representa la probabilidad de estar en un conjunto dado que en este momento me encuentro en x .

Definición 1.4. Sea $f : \mathbb{R} \rightarrow \mathbb{R}_+$ una función medible y N un kernel. Definimos la función Nf como sigue:

$$Nf(x) = \int_{\mathbb{R}} f(y)N(x, dy).$$

Proposición 1.1. La función Nf es medible.

DEMOSTRACIÓN. Supongamos primero que f es una función indicadora, digamos 1_A , con $A \in \mathcal{B}$; entonces $Nf(x) = \int_{\mathbb{R}} 1_A(y)N(x, dy) = N(x, A)$ la cual, por definición, es una función medible para todo A en \mathcal{B} . Usando el hecho de que la integral es lineal y que toda función medible positiva es el límite creciente de funciones simples obtenemos el resultado buscado. \square

Análogamente, si M es un kernel, probaremos que la función $NM(x, A) = \int_{\mathbb{R}} M(y, A)N(x, dy)$ es un kernel. Si x es fijo entonces claramente $NM(x, \cdot)$ es una función positiva y que al vacío le asigna el 0. Sea $\{A_n\}$ una sucesión de conjuntos disjuntos en \mathcal{B} entonces:

$$\begin{aligned} NM\left(x, \bigcup A_n\right) &= \int_{\mathbb{R}} M\left(y, \bigcup A_n\right) N(x, dy) = \int_{\mathbb{R}} \sum_{n=1}^{\infty} M(y, A_n)N(x, dy) \\ &= \sum_{n=1}^{\infty} \int_{\mathbb{R}} M(y, A_n)N(x, dy) = \sum_{n=1}^{\infty} NM(x, A_n) \end{aligned}$$

El intercambio entre suma e integral en las igualdades anteriores es posible ya que todos los sumandos son positivos y el teorema de convergencia monótona nos dice que en este caso si es posible realizarlo. De la cuenta anterior se sigue que NM es un kernel.

Sea $\mathbf{X} = (X_t)_{t \geq 0}$, un proceso estocástico y \mathcal{F}_t la filtración natural del proceso, si para cada $t > s \geq 0$ definimos una probabilidad de transición $P_{s,t}$, diremos que \mathbf{X} es un proceso de Markov con respecto a la filtración \mathcal{F}_t y la medida de

probabilidad \mathbb{P} si para toda función f medible positiva y para toda $t > s \geq 0$ se tiene que:

$$\mathbb{E}[f(X_t)|\mathcal{F}_s] = P_{s,t}f(X_s) \quad \text{c.s con respecto a } \mathbb{P}$$

Decimos que \mathbf{X} es un proceso homogéneo si $P_{s,t} = P_{0,t-s}$. En este caso definimos la función de distribución para $x \in \mathbb{R}$ fijo: $F(x, y, t) := P_t(x, (-\infty, y])$; así mismo definimos la densidad de transición al tiempo t como

$$p(x, y, t) = \frac{d}{dy}F(x, y, t)$$

en caso de que ésta exista.

La propiedad de Markov nos dice entonces que la esperanza de una función de X_t dada toda la información hasta el tiempo s no es sino la 'esperanza' de la función dada sólo la información del presente. De esta forma queda clara la conexión entre procesos de Markov y cadenas de Markov a tiempo discreto.

Daremos ahora como ejemplo de un proceso de Markov al movimiento browniano. El movimiento browniano como sabemos se puede definir como sigue:

Definición 1.5. *El movimiento browniano es un proceso $(W_t)_{t \geq 0}$ con las siguientes propiedades:*

1. *El proceso tiene incrementos independientes e idénticamente distribuidos.*
2. *$W_t \sim Normal(0, t)$*
3. *$W_0 = 0$ c.s.*
4. *El proceso tiene trayectorias continuas c.s.*

Un ejemplo del movimiento browniano puede verse en la figura 1.

Cabe aclarar que un proceso con valores en \mathbb{R} se dice que tiene trayectorias continuas casi seguramente si para $\omega \in \Omega$ la función $t \rightarrow X_t(\omega)$ es continua excepto en un conjunto de medida 0.

Para probar que el movimiento browniano es en efecto un proceso de Markov será necesario el siguiente:

Lema 1.1. *Un proceso estocástico X_t es de Markov si y sólo si para cualesquiera $0 \leq t_1 < t_2 \dots < t_n$ y f_i medible positiva se tiene que:*

$$\mathbb{E} \left[\prod_{i=0}^n f_i(X_i) \right] = \int_{\mathbb{R}} v(dx_0) f_0(x_0) \int_{\mathbb{R}} P_{0t_1}(x_0, dx_1) f_1(x_1) \dots \int_{\mathbb{R}} P_{t_{n-1}t_n}(x_{n-1}, dx_n) f_n(x_n)$$

Donde $v(A) = \mathbb{P}(X_0 \in A)$

DEMOSTRACIÓN. Supongamos primero que X_t es de Markov. Si $n = 0$ el resultado es evidente. Para el caso general notemos que por propiedades básicas de la esperanza condicional se cumple que:

$$\begin{aligned} \mathbb{E} \left[\prod_{i=0}^n f_i(X_{t_i}) \right] &= \mathbb{E} \left[\mathbb{E} \left[\prod_{i=0}^n [f_i(X_{t_i}) | \mathcal{F}_{t_{n-1}}] \right] \right] \\ &= \mathbb{E} \left[\prod_{i=0}^{n-1} f_i(X_{t_i}) \mathbb{E} [f_n(X_{t_n}) | \mathcal{F}_{t_{n-1}}] \right] \\ &= \mathbb{E} \left[\prod_{i=0}^{n-1} f_i(X_{t_i}) \cdot P_{t_{n-1}t_n} f(X_{t_{n-1}}) \right]. \end{aligned}$$

Donde la segunda igualdad se da pues $\prod_{i=0}^{n-1} f_i(X_{t_i})$ es una función $\mathcal{F}_{t_{n-1}}$ medible y en la tercera igualdad aplicamos la propiedad de Markov al tiempo t_{n-1} . El producto de la última expresión tiene sólo $n - 1$ factores (el último factor es $f_{n-1}(X_{t_{n-1}}) \cdot P_{t_{n-1}t_n} f(X_{t_{n-1}})$, que es una función de $X_{t_{n-1}}$). Por lo tanto podemos aplicar la hipótesis de inducción para ver que la expresión anterior es igual a:

$$\begin{aligned} &\int_{\mathbb{R}} v(dx_0) f_0(x_0) \int_{\mathbb{R}} P_{0t_1}(x_0, dx_1) f_1(x_1) \\ &\quad \dots \int_{\mathbb{R}} P_{t_{n-2}t_{n-1}}(x_{n-2}, dx_{n-1}) f_{n-1}(x_{n-1}) P_{t_{n-1}t_n} f_n(x_{n-1}) \\ &= \int_{\mathbb{R}} v(dx_0) f_0(x_0) \int_{\mathbb{R}} P_{0t_1}(x_0, dx_1) f_1(x_1) \dots \\ &\quad \int_{\mathbb{R}} P_{t_{n-2}t_{n-1}}(x_{n-2}, dx_{n-1}) f_{n-1}(x_{n-1}) \int_{\mathbb{R}} P_{t_{n-1}t_n}(x_{n-1}, dx_n) f_n(x_n) \end{aligned}$$

Inversamente queremos probar que para todo conjunto $A \in \mathcal{F}_s$ y para toda función f medible se cumple que:

$$\mathbb{E}[1_A f(X_t)] = \mathbb{E}[1_A P_{st} f(X_s)]$$

Por el lema de las clases monótonas sabemos que es suficiente considerar conjuntos de la forma:

$$A = (X_{t_1} \in A_1, \dots, X_{t_n} \in A_n)$$

Con $0 \leq t_0, \dots < t_n \leq s$ entonces la ecuación anterior se convierte en:

$$\mathbb{E} \left[\prod_{i=0}^n 1_{X_i \in A_i} f(X_t) \right] = \mathbb{E} \left[\prod_{i=0}^n 1_{X_i \in A_i} P_{st} f(X_s) \right]$$

Pero esto se ve aplicando la expresión del enunciado a ambos lados de la igualdad. \square

Gracias a los teoremas de aproximación para funciones medibles positivas, utilizando el lema anterior podemos asegurar que un proceso es de Markov si y sólo si para cualesquiera $0 < t_1, \dots < t_n$ y $A_1, \dots, A_n \subset \mathbb{R}$ se tiene que:

$$\mathbb{P}(X_{t_1} \in A_1, \dots, X_{t_n} \in A_n) = \int_{\mathbb{R}} v(dx) \int_{A_1} P_{0t_1}(x, dx_1) \dots \int_{A_n} P_{t_{n-1}t_n}(x_{n-1}, dx_n)$$

Si X_t es un proceso homogéneo y existen las densidades de transición, de lo anterior se sigue que una condición suficiente para que un proceso sea de Markov es que se cumpla la siguiente igualdad:

$$\mathbb{P}(X_{t_1} \in dx_1, \dots, X_{t_n} \in dx_n) = \int_{\mathbb{R}} v(dx) p(x, x_1, t_1) p(x_1, x_2, t_2 - t_1) \dots p(x_{n-1}, x_n, t_n - t_{n-1})$$

Proposición 1.2. *El movimiento browniano es un proceso de Markov homogéneo con probabilidades de transición $P_t(x, A) = \int_A p(x, y, t) dy$ donde $p(x, y, t)$ es la densidad de una variable normal con media x y varianza t .*

Es inmediato ver que la probabilidad de transición va de acuerdo con el punto 2 de la definición de W_t .

DEMOSTRACIÓN. Sea $U_i = W_{t_i} - W_{t_{i-1}}$, por hipótesis se tiene que las variables U_i son independientes con densidad $f_i(u_i) = p(0, u_i, t_i - t_{i-1})$ por lo tanto su densidad conjunta está dada por

$$\mathbb{P}(U_1 \in du_1 \dots U_n \in du_n) = \prod_{i=1}^n f_i(u_i)$$

Si ahora utilizamos el teorema del cambio de variable vemos que:

$$\begin{aligned} \mathbb{P}(W_1 \in dx_1, \dots, W_n \in dx_n) &= \prod_{i=1}^n f_i(x_i - x_{i-1}) \\ &= \prod_{i=1}^n p(x_{i-1}, x_i, t_i - t_{i-1}) \end{aligned}$$

Pero el lema 1.1 nos asegura que la última igualdad implica que el movimiento browniano es un proceso de Markov. \square

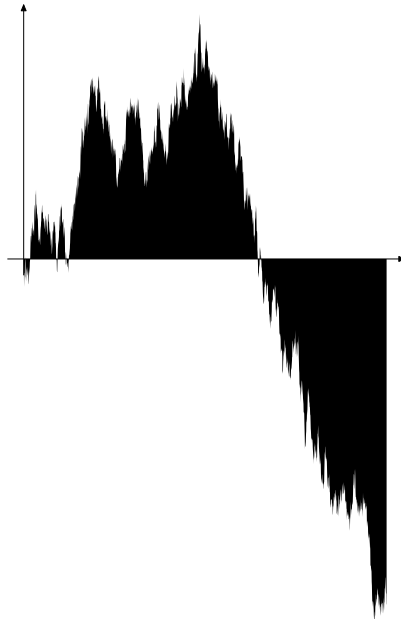


FIGURA 1. Ejemplo de un movimiento browniano.

Proposición 1.3. *Si W_t un movimiento browniano y $s > 0$, entonces el proceso $B_t = W_{s+t} - W_s$ es un movimiento browniano independiente de \mathcal{F}_s*

DEMOSTRACIÓN. Para probar que B_t cumple las condiciones del movimiento browniano y que es un proceso independiente de \mathcal{F}_s sólo será necesario calcular las densidades conjuntas de: $W_{s_1}, \dots, W_{s_n}, W_s, B_{t_1}, \dots, B_{t_m}$ para cualesquiera $0 \leq s_1 < \dots < s_n \leq s$ y t_1, \dots, t_n .

Sabemos ya que:

$$\begin{aligned} \mathbb{P}(W_{s_1} \in dx_1 \dots, W_{s_n} \in dx_n, W_s \in dx, W_{s+t_1} \in dy_1, \dots, W_{s+t_m} \in dy_m) \\ = p(0, x_1, s_1) \dots p(x_n, x, s - t_n) p(x, y_1, t_1) \dots p(y_{m-1}, y_m, t_m - t_{m-1}) \end{aligned}$$

Haciendo el cambio de variable:

$$B_{t_i} = W_{s+t_i} - W_s$$

Obtenemos la densidad conjunta:

$$\begin{aligned} \mathbb{P}(W_{s_1} \in dx_1 \dots, W_{s_n} \in dx_n, W_s \in dx, B_{t_1} \in dy_1, \dots, B_{t_m} \in dy_m) \\ = p(0, x_1, s_1), \dots, p(x_n, x, s - s_n) p(x, x + y_1, t_1), \\ \dots, p(x + y_{m-1}, x + y_m, t_m - t_{m-1}) \end{aligned}$$

Si integramos con respecto a x vemos que la densidad buscada inicialmente es igual a:

$$p(0, x_1, s_1), \dots, p(x_{n-1}, x_n, t_n - t_{n-1}) p(0, y_1, t_1), \dots, p(y_{m-1}, y_m, t_m - t_{m-1})$$

Lo cual prueba que B_t tiene las mismas distribuciones finito dimensionales que W_t , además claramente empieza en 0 y tiene trayectorias continuas. Por lo tanto B_t es un movimiento browniano. Además por el cálculo anterior sabemos que eventos de la forma $(W_{s_1} \in A_1, \dots, W_{s_n} \in A_n)$ son independientes con los eventos $(B_{t_1} \in C_1, \dots, B_{t_m} \in C_m)$ y por el teorema de las clases monótonas podemos concluir que B_t es independiente de \mathcal{F}_s como queríamos. \square

Continuemos ahora con el estudio de procesos de Markov más generales pues resulta más fácil introducir ciertos conceptos y realizar demostraciones que requerirían de más cuentas si las probáramos sólo para el browniano. En lo que resta del capítulo trataremos solamente con procesos de Markov homogéneos.

Así, por el lema 1.1 y por el teorema de extensión de Kolmogorov (véase [RY91] o [Tud02]), para toda medida de probabilidad ν en \mathbb{R} existe una medida de probabilidad en el espacio de trayectorias dada por:

$$(1) \quad P_\nu(X_{t_0} \in A_0, \dots, X_{t_n} \in A_n) = \int_{\mathbb{R}} \nu(dx) \int_{A_0} P_{t_0}(x, dx_0) \dots \int_{A_n} P_{t_n - t_{n-1}}(x_{n-1}, dx_n)$$

Si $\nu = \epsilon_x$, la medida puntual en x , tiene sentido hablar de la medida P_x , i.e la medida de probabilidad en el espacio de trayectorias que comienzan en x

c.s. Si Z es una variable aleatoria, denotaremos a las esperanzas con respecto a \mathbb{P}_v y \mathbb{P}_x respectivamente por $\mathbb{E}_v[Z]$, $\mathbb{E}_x[Z]$. En particular si $Z = 1_{X_t \in A}$ se tiene que $\mathbb{E}_x[Z] = \mathbb{P}_x(X_t \in A) = P_t(x, A)$. Esto en particular nos dice que para cualquier $A \in \mathcal{B}$, el mapeo $x \rightarrow \mathbb{E}_x[1_{X_t \in A}]$ es medible.

Proposición 1.4. *Para toda x y para toda Z variable aleatoria el mapeo $x \rightarrow \mathbb{E}_x[Z]$ es medible y además para cualquier medida v :*

$$\mathbb{E}_v[Z] = \int_{\mathbb{R}} v(dx) \mathbb{E}_x[Z]$$

DEMOSTRACIÓN. Sea $\mathcal{M} = \{A \in \mathcal{F} \mid \mathbb{E}_v[1_A] = \int_{\mathbb{R}} v(dx) \mathbb{E}_x[1_A]\}$. Probaremos que es una clase monótona.

Claramente $\Omega \in \mathcal{M}$ pues ambos lados de la ecuación son 1. Supongamos entonces que $A \subset B \in \mathcal{M}$ entonces:

$$\begin{aligned} \mathbb{E}_v[1_{B \setminus A}] &= \mathbb{E}_v[1_B - 1_A] \\ &= \mathbb{E}_v[1_B] - \mathbb{E}_v[1_A] \\ &= \int_{\mathbb{R}} v(dx) \mathbb{E}_x[1_B - 1_A] \\ &= \int_{\mathbb{R}} v(dx) \mathbb{E}_x[1_{B \setminus A}] \end{aligned}$$

Es decir que $B \setminus A \in \mathcal{M}$. Por último sea $(A_n) \in \mathcal{M}$ una sucesión creciente de conjuntos entonces:

$$\begin{aligned} \mathbb{E}_v[1_{\bigcup A_n}] &= \mathbb{P}_v\left(\bigcup A_n\right) \\ &= \lim_{n \rightarrow \infty} \mathbb{P}_v(A_n) \quad (\text{Continuidad de la medida}) \\ &= \lim_{n \rightarrow \infty} \int_{\mathbb{R}} v(dx) \mathbb{P}_x(A_n) \\ &= \int_{\mathbb{R}} v(dx) \lim_{n \rightarrow \infty} \mathbb{P}_x(A_n) \\ &= \int_{\mathbb{R}} v(dx) \mathbb{E}_x[1_{\bigcup A_n}] \end{aligned}$$

Cabe aclarar que el intercambio del límite con la integral está permitido pues $\mathbb{P}_x(A_n)$ es una sucesión creciente y positiva. Por los cálculos anteriores sabemos que $\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$ y por lo tanto \mathcal{M} es una clase monótona.

Tomemos entonces $A = (X_0 \in A_0, \dots, X_n \in A_n)$; se tiene que $\mathbb{P}_x(A)$ está dado por la ecuación 1 con $v = \epsilon_x$, en este caso utilizando la proposición 1.1 se tiene que el mapeo $x \rightarrow \mathbb{P}_x[A]$ es medible y que la igualdad se da trivialmente.

Utilizando el lema de las clases monótonas se tiene que la propiedad buscada se cumple para todo A medible y por la linealidad de la integral y la convergencia monótona se sigue que la propiedad se cumple para cualquier variable aleatoria. \square

Definición 1.6. Sea X_t el proceso coordinado en $(\mathbb{R}^{\mathbb{R}_+}, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, dado por $X_t(\omega) = \omega(t)$. Definimos el operador de traslación de trayectorias:

$$\theta_t : \mathbb{R}^{\mathbb{R}_+} \longrightarrow \mathbb{R}^{\mathbb{R}_+}$$

de la siguiente manera $X_s \circ \theta_t(\omega) = X_{s+t}(\omega)$

La función θ_t hace que veamos el proceso $(X_s)_{s \geq 0}$ a partir del tiempo t .

Proposición 1.5. Si Z es una función medible con respecto a la sigma álgebra generada por el proceso X , entonces:

$$\mathbb{E}_v[Z \circ \theta_s | \mathcal{F}_s] = \mathbb{E}_{X_s}[Z]$$

DEMOSTRACIÓN. Queremos probar que:

$$\mathbb{E}_v[Z \circ \theta_s \cdot Y] = E_v[E_{X_s}[Z] \cdot Y]$$

Para cualquier función positiva Y y medible con respecto a \mathcal{F}_s . Pero por el lema de las clases monótonas sólo es necesario demostrarlo para el caso en que $Y = \prod_{j=1}^n f_j(X_{t_j})$ y $Z = \prod_{i=1}^m g_i(X_{t_i})$ donde f_j, g_i son funciones medibles con respecto a la σ -álgebra generada por X y $0 \leq t_1 < \dots < t_n \leq s$; sin embargo este resultado se obtiene precisamente del lema 1.1. \square

Ahora que hemos introducido la familia de probabilidades \mathbb{P}_x para un proceso de Markov y el operador de traslación, podemos ver que significa esto en el movimiento browniano.

Hasta ahora los procesos estocásticos fueron definidos en el espacio canónico de trayectorias $\mathbb{R}^{\mathbb{R}_+}$ y el movimiento browniano no fue la excepción; sin embargo el movimiento browniano pedía la condición de que $W_0 = 0$ casi seguramente; ésto corresponde a la medida de probabilidad \mathbb{P}_0 . Por lo que podríamos hablar de un movimiento browniano iniciado en x al considerar las mismas probabilidades de transición pero con medida inicial ϵ_x (la medida puntual en x), dotando así al espacio canónico con la medida \mathbb{P}_x .

Calculemos entonces

$$\mathbb{P}_x(W_t \in A)$$

Por definición sabemos que

$$\mathbb{P}_x(W_t \in A) = P_t(x, A) = \int_A \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}}$$

Es fácil ver entonces que

$$\mathbb{P}_x(W_{t_1} \in A_1, \dots, W_{t_n} \in A_n) = \mathbb{P}_0(W_{t_1} + x \in A_1, \dots, W_{t_n} + x \in A_n)$$

Por lo que es lo mismo 'ver' al proceso $W_t + x$ bajo la ley \mathbb{P}_0 que al proceso W_t bajo la ley \mathbb{P}_x , esto nos será de mucha utilidad en la sección siguiente. Por conveniencia, cuando hablemos del movimiento browniano, si se utiliza la medida \mathbb{P} estaremos refiriéndonos a la medida \mathbb{P}_0 .

Como último ejemplo volveremos a calcular la distribución conjunta de W_s, W_{t+s} utilizando el operador de traslación de trayectorias para dejar en claro como se ve ésto en el movimiento browniano.

$$\begin{aligned} \mathbb{P}(W_s \in A, W_{s+t} \in B) &= \mathbb{E}[1_{W_s \in A} 1_{W_t \circ \theta_s \in B}] \\ &= \mathbb{E}[1_{W_s \in A} \mathbb{E}[1_{W_t \in B} \circ \theta_s | \mathcal{F}_s]] \end{aligned}$$

Que por la propiedad de Markov sabemos que es igual a

$$\begin{aligned} \mathbb{E}[1_{W_s \in A} \mathbb{E}_{W_s}[1_{W_t \in B}]] &= \int_A \mathbb{P}_0(W_s \in dx) \mathbb{P}_0(W_t + x \in B) dx \\ &= \int_A \frac{1}{\sqrt{2\pi s}} e^{-\frac{x^2}{2s}} \int_B \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}} dy dx \end{aligned}$$

Tal como lo habíamos calculado antes.

Concluiremos esta sección con un lema que será ampliamente utilizado en el próximo capítulo.

Lema 1.2. Sean $A \in \mathcal{F}_t$, $C \in \mathcal{B}$ y $B = \cap_{n=1}^{\infty} (X_{t+t_n} \in A_n)$ con $A_n \in \mathcal{B}$; si $\mathbb{P}(A, X_t \in dy)$ existe entonces:

$$(2) \quad \mathbb{P}_x(A, X_t \in C, B) = \int_C \mathbb{P}_x(A, X_t \in dy) P_y(B') dy$$

donde $B' = \cap_{n=1}^{\infty} (X_{t_n} \in A_n)$.

DEMOSTRACIÓN. El resultado puede obtenerse con una aplicación simple de la propiedad de Markov al tiempo t . Primero expresemos la probabilidad buscada como la esperanza de una función indicadora y condicionemos con respecto a \mathcal{F}_t . Entonces:

$$\begin{aligned}
 \mathbb{P}_x(A, X_t \in C, B) &= \mathbb{E}_x[1_A 1_{X_t \in C} 1_{\cap_{n=1}^\infty (X_{t+t_n} \in A_n)}] \\
 &= \mathbb{E}_x[1_A 1_{X_t \in C} \mathbb{E}[1_{\cap_{n=1}^\infty (X_{t_n} \circ \theta_t \in A_n)} | \mathcal{F}_t]] \\
 &= \mathbb{E}_x[1_A 1_{X_t \in C} \mathbb{E}[1_{\cap_{n=1}^\infty X_{t_n} \in A_n} \circ \theta_t | \mathcal{F}_t]]
 \end{aligned}$$

(3)

Después utilizando el lema anterior vemos que la expresión anterior es igual a:

$$\begin{aligned}
 &\mathbb{E}_x [1_A 1_{X_t \in C} E_{X_t} [1_{B'}]] \\
 &= \int_C \mathbb{P}_x(A, X_t \in dy) P_y(B') dy
 \end{aligned}$$

□

2. Propiedad de Markov fuerte del movimiento browniano

En esta sección vamos a extender la propiedad de Markov a tiempos de paro. Para empezar veamos que es un tiempo de paro.

Definición 1.7. *Una variable aleatoria $T : \Omega \rightarrow \overline{\mathbb{R}}_+$ se dice que es un tiempo de paro si para toda t se cumple que $\{T \leq t\} \in \mathcal{F}_t$*

Es decir un tiempo de paro es un tiempo aleatorio del cual se puede saber toda su información sólo con la información conocida hasta el momento. Así podemos decir que para un proceso estocástico $\mathbf{X} = (X_t)_{t \geq 0}$, $X_T(\omega) = X_t(\omega)$ si $T(\omega) = t$ y $X_T(\omega) = \Delta$ si $T(\omega) = \infty$ donde Δ es un punto que no se encuentra en el espacio de estados E . También podemos definir la sigma álgebra \mathcal{F}_T como sigue:

$$\mathcal{F}_T = \{A | A \cap \{T \leq s\} \in \mathcal{F}_s \forall s \in \mathbb{R}\}$$

En la sección anterior probamos que $W_{t+s} - W_s$ era un movimiento browniano independiente de \mathcal{F}_s . En esta sección veremos que esto también es cierto para un tiempo de paro T .

Proposición 1.6. *Sea W_t un movimiento browniano y T un tiempo de paro finito casi seguramente, entonces el proceso $B_t = W_{t+T} - W_T$ es un movimiento browniano independiente de \mathcal{F}_T .*

DEMOSTRACIÓN. Supongamos primero que T es un tiempo de paro discreto, es decir sus valores los toma en un conjunto discreto D . Sea $A \in \mathcal{F}_T$,

$A = \bigcup_{d \in D} A \cap \{T = d\}$. Entonces para cualesquiera $\Gamma_1, \dots, \Gamma_n \in \mathcal{B}$ y $0 < t_1 < \dots < t_n$ se tiene:

$$\begin{aligned} \mathbb{P}((B_{t_i} \in \Gamma_i), A) &= \sum_{d \in D} \mathbb{P}((W_{t_i+d} - W_d \in \Gamma_i), A, T = d) \\ &= \sum_{d \in D} \mathbb{P}(W_{t_i+d} - W_d \in \Gamma) \mathbb{P}(A, T = d) \end{aligned}$$

Esto ocurre pues por definición $A \cap \{T = d\} \in \mathcal{F}_d$ y por la proposición 1.3 es un evento independiente de $(W_{t_i+d} - W_d \in \Gamma_i)$. La expresión anterior es además igual a:

$$\begin{aligned} &\int_{\Gamma_1} p(0, x_1, t_1) \dots \int_{\Gamma_n} p(x_{n-1}, x_n, t_n - t_{n-1}) dx_n \dots dx_1 \sum_{d \in D} \mathbb{P}(A, T = d) \\ &= \left(\int_{\Gamma_1} p(0, x_1, t_1) \dots \int_{\Gamma_n} p(x_{n-1}, x_n, t_n - t_{n-1}) dx_n \dots dx_1 \right) \mathbb{P}(A) \end{aligned}$$

Mediante el lema de las clases monótonas es fácil ver que el proceso B_t es independiente de \mathcal{F}_T ; además si $A = \Omega$ encontramos que las distribuciones finito dimensionales de B_t son las mismas que las de W_t , y como las otras propiedades del movimiento browniano se cumplen trivialmente concluimos que B_t es un movimiento browniano.

Supongamos ahora que T es un tiempo finito casi seguramente. Definimos los tiempos de paro discretos T_n como:

$$T_n = \frac{[2^n T] + 1}{2^n}$$

Donde $[x]$ denota la parte entera de x . Es bastante claro que $T_n \rightarrow T$ de forma decreciente y por lo tanto $\mathcal{F}_T \subset \mathcal{F}_{T_n}$. Entonces si $B_t^{(n)} = W_{t+T_n} - W_{T_n}$, por lo dicho anteriormente sabemos que $B_t^{(n)}$ es independiente de \mathcal{F}_{T_n} y a su vez esto implica que para toda n el proceso $B_t^{(n)}$ es independiente de \mathcal{F}_T , también por la continuidad de las trayectorias del movimiento browniano se tiene que $B_t^{(n)}$ converge a B_t puntualmente y por lo tanto B_t es independiente de \mathcal{F}_T y como $B_t^{(n)}$ se distribuye Normal(0, t) para toda n entonces también lo hace B_t . Sólo nos resta probar entonces que B_t tiene incrementos independientes. Para esto notemos lo siguiente, si T es un tiempo de paro finito, entonces claramente también lo es $T + t$ para toda t , por lo tanto si definimos $B_{t_1}^t = B_{t_1+t+T} - B_{t+T} = B_{t+t_1} - B_t$, este proceso es independiente de \mathcal{F}_{t+T} , y en particular es independiente de B_t . Utilizando este argumento recursivamente

podemos facilmente ver que el proceso B_t tiene incrementos independientes y por lo tanto es un movimiento browniano. \square

Lema 1.3. *Sea W_t un movimiento browniano y $T_a = \inf\{t|W_t = a\}$ con $a > 0$, si $x > a$ entonces $\mathbb{P}(W_t > x) = \mathbb{P}(W_t < 2a - x, T_a < t)$, a esto se le conoce como el principio de reflexión.*

DEMOSTRACIÓN. Como las trayectorias de W_t son continuas entonces notemos que $T_a = \inf\{t|W_t \geq a\}$ por lo tanto:

$$(T_a > t) = \bigcup_{q \in \mathbb{Q}, q \leq t} (W_q < a) \in \mathcal{F}_t$$

Por lo tanto T_a es un tiempo de paro, además si $x > a$ entonces $(W_t > x, T_a < t) = (W_t > x)$ pues la continuidad de las trayectorias fuerza a haber cruzado a si en este momento estoy por arriba de él, entonces utilizando el resultado anterior tenemos que:

$$\begin{aligned} \mathbb{P}(W_t > x, T_a < t) &= \mathbb{P}(W_{(t-T_a)+T_a} - W_{T_a} > x - a, T_a < t) \\ &= \mathbb{P}(B_{t-T_a} > x - a)\mathbb{P}(T_a < t) \\ &= \mathbb{P}(B_{t-T_a} < a - x)\mathbb{P}(T_a < t) \\ &= \mathbb{P}(W_t < 2a - x, T_a < t) \end{aligned}$$

Probando así el lema. El resultado es análogo si $a < 0$ y $x < a$. Sólo necesitamos cambiar el sentido de las desigualdades. \square

Como aplicación del principio de reflexión calcularemos la distribución del máximo acumulativo del movimiento browniano.

Corolario 1.1. *Sea W_t un movimiento browniano, entonces*

$$\mathbb{P}\left(\sup_{s < t} W_t > a\right) = \mathbb{P}(|W_t| > a) = 2\mathbb{P}(W_t > a)$$

DEMOSTRACIÓN.

$$\begin{aligned} \mathbb{P}\left(\sup_{s < t} W_t > a\right) &= \mathbb{P}(T_a < t) \\ &= \mathbb{P}(T_a < t, W_t > a) + \mathbb{P}(T_a < t, W_t < a) \\ &= \mathbb{P}(W_t > a) + \mathbb{P}(T_a < t, W_t < 2a - a) \\ &= \mathbb{P}(W_t > a) + \mathbb{P}(W_t > a) = 2\mathbb{P}(W_t > a) \end{aligned}$$

\square

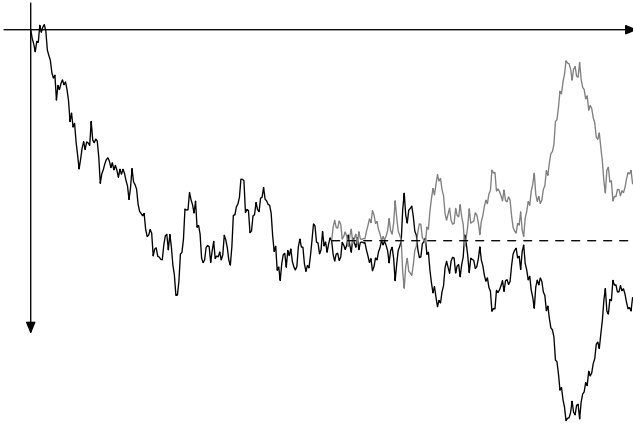


FIGURA 2. El principio de reflexión en el movimiento browniano

Proposición 1.7. Sea W_t un movimiento browniano y $a > 0$, si T_a es como en el lema anterior entonces la densidad de T_a está dada por:

$$f(t) = \frac{a}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}}$$

DEMOSTRACIÓN.

$$\begin{aligned} \mathbb{P}(T_a < t) &= \mathbb{P}\left(\sup_{s < t} W_t > a\right) \\ &= \mathbb{P}(|W_t| > a) = \mathbb{P}\left(\frac{1}{\sqrt{t}}|W_t| > \frac{a}{\sqrt{t}}\right) \\ &= \mathbb{P}\left(|W_1| > \frac{a}{\sqrt{t}}\right) \\ &= \int_{\frac{a}{\sqrt{t}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \end{aligned}$$

Si ahora derivamos con respecto de t la expresión anterior obtenemos que la densidad está dada por:

$$f(t) = \frac{a}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}}$$

□

Notemos que por la propiedad de simetría del movimiento browniano hubiésemos podido calcular la densidad anterior para $a < 0$ reflejando el proceso con respecto al eje de las abscisas. Por lo que la densidad del primer tiempo de arribo a a está dada en general por

$$f(t) = \frac{|a|}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}}$$

Por último si recordamos que en la sección anterior vimos que el proceso W_t+x bajo \mathbb{P}_0 era lo mismo que el proceso W_t bajo la ley \mathbb{P}_x entonces podemos concluir de forma sencilla que:

$$\mathbb{P}_x(T_0 \in A) = \mathbb{P}_0(T_{-x} \in A)$$

Además es fácil establecer un resultado análogo al principio de reflexión pero con la medida \mathbb{P}_x .

Excursiones en el movimiento browniano

En este capítulo definiremos uno de los principales objetos de estudio de este trabajo, a saber: la excursión browniana normalizada. Primeramente enunciaremos algunas propiedades básicas del movimiento browniano, así como la densidad conjunta del browniano con su mínimo acumulativo. Después nos enfocaremos en definir la excursión del browniano que contiene un tiempo determinista, y por último daremos las distribuciones finito dimensionales de la excursión browniana normalizada.

1. Propiedades del movimiento browniano

Sea W_t un movimiento browniano y $Y_t = |W_t|$. Definimos:

$$\begin{aligned} d_t &= \inf\{s \geq t \mid W_s = 0\} \\ g_t &= \sup\{s \leq t \mid W_s = 0\} \end{aligned}$$

Es fácil ver que el tiempo aleatorio d_t representa el primer regreso a cero del proceso después del tiempo t , mientras que g_t representa el último paso por el cero antes del tiempo t . Calculemos ahora la densidad conjunta de (g_t, W_t) y de (d_t, W_t) .

Recordemos que en el capítulo anterior definimos $T_0 := \inf\{t > 0 \mid W_t = 0\}$. Tomemos $x > 0$, entonces

$$\begin{aligned} \mathbb{P}(W_t \geq x, d_t \geq s) &= \mathbb{P}(W_t \geq x, T_0 \circ \theta_t \geq s - t) \\ &= \mathbb{E}[1_{W_t \geq x} 1_{T_0 \circ \theta_t \geq s - t}] \\ &= \mathbb{E}[\mathbb{E}[1_{W_t \geq x} 1_{T_0 \circ \theta_t \geq s - t} \mid \mathcal{F}_t]] \\ &= \mathbb{E}[1_{W_t \geq x} \mathbb{E}[1_{T_0 \geq s - t} \circ \theta_t \mid \mathcal{F}_t]] \end{aligned}$$

Pero una simple aplicación de la propiedad de Markov al tiempo t nos permite ver que lo anterior es en efecto:

$$\begin{aligned}
\mathbb{E} [1_{W_t \geq x} \mathbb{P}_{W_t}(T_0 \geq s - t)] &= \\
&= \int_x^\infty \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}} \mathbb{P}_y(T_0 \geq s - t) dy \\
&= \int_x^\infty \int_{s-t}^\infty \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}} \frac{y}{\sqrt{2\pi r^3}} e^{-\frac{y^2}{2r}} dr dy
\end{aligned}$$

Debido a la simetría del movimiento browniano, es claro que si $x < 0$ entonces la probabilidad anterior es la misma a excepción de que x se cambia por $-x$. Derivando con respecto a x y con respecto a s obtenemos la densidad conjunta:

$$f_{W,d}(x, s) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \frac{|x|}{\sqrt{2\pi(s-t)^3}} e^{-\frac{x^2}{2(s-t)}}$$

Para calcular la densidad conjunta de (g_t, W_t) utilizaremos la propiedad de inversión temporal del movimiento browniano. Sea $x > 0$, entonces:

$$\begin{aligned}
&\mathbb{P}(g_t \leq s, W_t \geq x) \\
&= \mathbb{P}\left(\inf_{s \leq r \leq t} W_r > 0, W_t \geq x\right) \\
&= \mathbb{P}\left(\inf_{s \leq r \leq t} rW_{1/r} > 0, tW_{1/t} \geq x\right) \\
&= \mathbb{P}\left(\inf_{1/t \leq r \leq 1/s} W_r > 0, W_{1/t} \geq \frac{x}{t}\right) \\
&= \mathbb{P}\left(d_{1/t} \geq \frac{1}{s}, W_{1/t} \geq \frac{x}{t}\right) \\
&= \int_{\frac{1}{s}}^\infty \int_{\frac{x}{t}}^\infty \frac{1}{\sqrt{2\pi 1/t}} e^{-\frac{y^2}{2/t}} \frac{|y|}{\sqrt{2\pi(r-1/t)^3}} e^{-\frac{y^2}{2(r-1/t)}} dy dr
\end{aligned}$$

Nuevamente derivando dos veces obtenemos la densidad buscada:

$$f_{W,g}(x, s) = \frac{1}{\sqrt{2\pi s}} \frac{|x|}{\sqrt{2\pi(t-s)^3}} e^{-\frac{x^2}{2(t-s)}}$$

Si utilizamos el lenguaje de procesos de Markov entonces estas dos densidades resultan muy intuitivas pues si

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

es la densidad de transición para el movimiento browniano y

$$(4) \quad g(0, y, t) = \frac{|y|}{\sqrt{2\pi t^3}} e^{-\frac{x^2}{2t}}$$

es la densidad del primer regreso a cero dado que el proceso empieza en y , entonces se tiene:

$$(5) \quad f_{W,d}(x, s) = p(0, x, t)g(0, x, s - t)$$

y

$$(6) \quad f_{W,g}(x, s) = p(0, 0, s)g(0, x, t - s)$$

La ecuación 5 nos dice que la densidad $f_{W,d}$ es la probabilidad de ir de 0 a x en el tiempo t y después regresar por primera vez al cero en $s - t$ unidades de tiempo, lo cual está de acuerdo con nuestra intuición ya que W_t es un proceso de Markov. La segunda fórmula puede ser un poco menos intuitiva; sin embargo la heurística detrás de la ecuación 6 es que la función de densidad del último paso por el 0 en el movimiento browniano antes del tiempo t es la misma que la función de densidad del primer regreso a 0. Ésto ocurre debido a la propiedad de inversión temporal. Entonces el lado derecho de la ecuación 6 denota la probabilidad de ir del 0 al 0 en el tiempo s y después ir al estado x en el tiempo $t - s$ sin volver a pasar por el cero.

Corolario 2.1. *La densidad conjunta de Y_t, d_t y la de Y_t, g_t están dadas respectivamente por*

$$f_{Y,d}(x, s) = 2p(0, x, t)g(0, x, s - t)1_{x>0}$$

y

$$f_{Y,g}(x, s) = 2p(0, 0, s)g(0, x, t - s)1_{x>0}$$

Proposición 2.1. *La densidad conjunta de g_t, d_t está dada por*

$$\mathbb{P}(g_t \in ds, d_t \in dr) = \frac{1}{2\pi\sqrt{s(r-s)^3}}1_{s<t<r}$$

DEMOSTRACIÓN. Veamos primero que

$$(7) \quad \mathbb{P}(g_t < s, d_t > r) = \mathbb{P}(W_s > 0, d_s > r) + \mathbb{P}(W_s < 0, d_s > r)$$

La igualdad se da debido a que $g_t < s$ y $d_t > r$ si y sólo si $W_s \neq 0$ y el movimiento browniano no se anula en el intervalo $[s, r]$; de lo contrario $g_t \geq s$ o $d_t \leq r$; por lo tanto $\{W_s \neq 0, d_s > r\}$ es igual al evento del lado izquierdo de la ecuación 7. Por último podemos separar la condición $W_s \neq 0$ en $W_s < 0$ ó $W_s > 0$.

Sustituyendo los valores que ya conocemos del lado derecho de la ecuación 7 obtenemos:

$$\mathbb{P}(g_t < s, d_t > r) = 2 \int_0^\infty \int_r^\infty \frac{1}{\sqrt{2\pi s}} e^{-\frac{x^2}{2s}} \frac{|x|}{\sqrt{2\pi(u-s)^3}} e^{-\frac{x^2}{2(u-s)}} dx du$$

Diferenciando dos veces obtenemos la formula buscada. \square

Definición 2.1. Definimos la longitud de la excursión del movimiento browniano como $L_t = d_t - g_t$.

Corolario 2.2. La densidad conjunta de g_t, L_t está dada por

$$\mathbb{P}(g_t \in ds, L_t \in dl) = \frac{1}{2\pi\sqrt{sl^3}} 1_{s < t} 1_{l > t-s}$$

DEMOSTRACIÓN. Simplemente aplíquese el teorema de cambio de variable a la función dada por :

$$\begin{aligned} g_t &= g_t \\ L_t &= d_t - g_t \end{aligned}$$

\square

Para concluir esta sección vamos a calcular la densidad de transición del movimiento browniano condicionado a no anularse.

Lema 2.1. Sean $x > 0, y > 0$ y T_0 el primer regreso a cero entonces:

$$\mathbb{P}_x(W_t > y, T_0 > t) = \mathbb{P}_x(W_t > y) - \mathbb{P}_x(W_t < -y)$$

DEMOSTRACIÓN. Ya que $-y < 0$ podemos aplicar el principio de reflexión, a saber:

$$\mathbb{P}_x(W_t < -y) = \mathbb{P}_x(W_t > 2(0) - (-y), T_0 < t) = \mathbb{P}_x(W_t > y, T_0 < t).$$

Por lo tanto:

$$\begin{aligned} \mathbb{P}_x(W_t > y) &= \mathbb{P}_x(W_t > y, T_0 > t) + \mathbb{P}_x(W_t > y, T_0 < t) \\ &= \mathbb{P}_x(W_t > y, T_0 > t) + \mathbb{P}_x(W_t < -y) \end{aligned}$$

Despejando obtenemos el resultado. \square

A la derivada con respecto a y de la probabilidad anterior la denotaremos por

$$q(x, y, t) = \frac{1}{\sqrt{2\pi t}} \left(e^{-\frac{(x-y)^2}{2t}} - e^{-\frac{(x+y)^2}{2t}} \right)$$

y representa la densidad de transición del movimiento browniano condicionado a no anularse.

Corolario 2.3. *La densidad conjunta de Y_{t_1}, Y_{t_2} y que el proceso no se anule en $[t_1, t_2]$ está dada por:*

$$(8) \quad \mathbb{P}(Y_{t_1} \in dy_1, T_0 \circ \theta_{t_1} > t_2, Y_{t_2} \in dy_2) = 2(p, 0, y_1)q(y_1, y_2, t_2 - t_1)$$

DEMOSTRACIÓN. Por la propiedad de Markov y el lema anterior se tiene que:

$$\mathbb{P}(Y_{t_1} > y_1, T_0 \circ \theta_{t_1} > t_2, Y_{t_2} > y_2) = \int_{y_1}^{\infty} 2p(0, x_1, t_1) \int_{y_2}^{\infty} q(x_1, x_2, t_2 - t_1) dx_1 dx_2$$

Si derivamos la expresión de la derecha con respecto a x_1 y x_2 obtenemos el resultado buscado. \square

2. El movimiento browniano y su mínimo

Esta sección estará dedicada a establecer resultados sobre el comportamiento del movimiento browniano y su mínimo acumulativo así como resultados sobre el mínimo del movimiento browniano en cierto intervalo de tiempo, los cuales serán muy útiles en la siguiente sección.

Proposición 2.2. *Sea $m_W(t) = \inf_{s \leq t} W_s$ el mínimo acumulativo del movimiento browniano hasta el tiempo t , sean $x \in \mathbb{R}$ y $y < 0$ con $x > y$ entonces:*

$$\mathbb{P}(W_t > x, m_W(t) < y) = \mathbb{P}(W_t < 2y - x)$$

DEMOSTRACIÓN. Notemos que $m_W(t) < y$ si y sólo si $T_y < t$ donde T_y es el primer tiempo de llegada a y . Por lo tanto:

$$\mathbb{P}(W_t > x, m_W(t) < y) = \mathbb{P}(W_t > x, T_y < t)$$

Por otro lado como $2y - x < y$ podemos aplicar el principio de reflexión, a saber:

$$\begin{aligned} \mathbb{P}(W_t < 2y - x) &= \mathbb{P}(W_t > 2y - (2y - x), T_y < t) \\ &= \mathbb{P}(W_t > x, T_y < t) \\ &= \mathbb{P}(W_t > x, m_W(t) < y). \end{aligned}$$

\square

Utilizando esta igualdad en las probabilidades podemos calcular la densidad conjunta de $W_t, m_W(t)$ simplemente derivando con respecto a x, y a la función

$$\mathbb{P}(W_t < 2y - x) = \int_{-\infty}^{2y-x} \frac{1}{\sqrt{2\pi t}} e^{-\frac{z^2}{2t}} dz$$

Corolario 2.4. *La densidad conjunta de $W_t, m_W(t)$ está dada por*

$$\varphi(x, y, t) = \frac{2(x-2y)}{\sqrt{2\pi t^3}} e^{-\frac{(x-2y)^2}{2t}}$$

Probaremos ahora una propiedad importante de la función φ

Proposición 2.3. *Para cualesquiera $x, y, t, l > 0$ se tiene que*

$$l\varphi(\sqrt{l}x, \sqrt{l}y, tl) = \varphi(x, y, t)$$

La demostración es una simple cuenta:

$$\begin{aligned} l\varphi(\sqrt{l}x, \sqrt{l}y, tl) &= \frac{2\sqrt{l^3}(x-2y)}{\sqrt{2\pi t^3 l^3}} e^{-\frac{(x-2y)^2 l}{2tl}} \\ &= \frac{2(x-2y)}{\sqrt{2\pi t^3}} e^{-\frac{(x-2y)^2}{2t}} \\ &= \varphi(x, y, t) \end{aligned}$$

Nos interesa ahora calcular la densidad conjunta de W_{t_1}, W_{t_2} y el mínimo del movimiento browniano en este intervalo de tiempo.

Definición 2.2. *Sea $g : \mathbb{R} \rightarrow \mathbb{R}$ una función y $t_1, t_2 \in \mathbb{R}$, denotamos por $m_g(t_1, t_2) = \inf_{t_1 \leq s \leq t_2} g(s)$ al mínimo de la función en el intervalo $[t_1, t_2]$.*

Proposición 2.4. *Sea $m_W(t_1, t_2)$ el mínimo del movimiento browniano en el intervalo $[t_1, t_2]$, entonces la densidad conjunta de $W_{t_1}, m_W(t_1, t_2), W_{t_2}$ está dada por*

$$\mathbb{P}(W_{t_1} \in dx_1, m_W(t_1, t_2) \in dy, W_{t_2} \in dx_2) = p(0, x_1, t_1)\varphi(x_2 - x_1, y - x_1, t_2 - t_1)$$

DEMOSTRACIÓN. Aplicando directamente el lema 1.2, el cual es una simple consecuencia de la propiedad de Markov, podemos observar que:

$$\begin{aligned} &\mathbb{P}(W_{t_1} > x_1, m_W(t_1, t_2) < y, W_{t_2} > x_2) \\ &= \int_{x_1}^{\infty} p(0, y_1, t_1)\mathbb{P}_{y_1}(m_W(t_2 - t_1) < y, W_{t_2 - t_1} > x_2)dy_1 \\ &= \int_{x_1}^{\infty} p(0, y_1, t_1)\mathbb{P}(m_W(t_2 - t_1) < y - y_1, W_{t_2 - t_1} > x_2 - y_1)dy_1 \\ &= \int_{x_1}^{\infty} \int_{-\infty}^{y-y_1} \int_{x_2-y_1}^{\infty} p(0, y_1, t_1)\varphi(y_2, z, t_2 - t_1)dy_2 dz dy_1 \end{aligned}$$

Derivando esta expresión con respecto a x_1, y, x_2 obtenemos la densidad buscada. \square

Corolario 2.5. *Sea $m_Y(t_1, t_2)$ el mínimo del proceso Y_t en el intervalo $[t_1, t_2]$, entonces para cualesquiera $x_1, x_2 > y > 0$ se tiene que:*

$$\mathbb{P}(Y_{t_1} \in dx_1, m_Y(t_1, t_2) \in dy, Y_{t_2} \in dx_2) = 2p(0, x_1, t_1)\varphi(x_2 - x_1, y - x_1, t_2 - t_1)$$

y

$$\begin{aligned} \mathbb{P}(Y_{t_1} \in dx_1, m_Y(t_1, t_2) = 0, Y_{t_2} \in dx_2) = \\ 2p(0, x_1, t_1)(p(x_1, x_2, t_2 - t_1) - q(x_1, x_2, t_2 - t_1)) \end{aligned}$$

Esto se sigue por la simetría del movimiento browniano:

$$\begin{aligned} \mathbb{P}(Y_{t_1} > x_1, m_Y(t_1, t_2) < y, Y_{t_2} > x_2) \\ = \mathbb{P}(W_{t_1} > x_1, 0 < m_W(t_1, t_2) < y, W_{t_2} > x_2) \\ + \mathbb{P}(W_{t_1} < -x_1, 0 > \sup_{t_1 < r < t_2} W_r > -y, W_{t_2} < -x_2) \\ = \mathbb{P}(W_{t_1} > x_1, 0 < m_W(t_1, t_2) < y, W_{t_2} > x_2) \\ + \mathbb{P}(-W_{t_1} > x_1, 0 < \inf_{t_1 < r < t_2} -W_r < y, -W_{t_2} > x_2) \\ = 2\mathbb{P}(W_{t_1} > x_1, 0 < m_W(t_1, t_2) < y, W_{t_2} > x_2) \end{aligned}$$

Derivando esta probabilidad obtenemos el resultado buscado. Para la segunda ecuación basta notar la igualdad entre los siguientes eventos:

$$\{Y_{t_1} \in dx_1, m_Y(t_1, t_2) = 0, Y_{t_2} \in dx_2\} = \{Y_{t_1} \in dx_1, T_0 \circ \theta_{t_1} < t_2, Y_{t_2} \in dx_2\}$$

Como además sabemos que:

$$\begin{aligned} \mathbb{P}(Y_{t_1} \in dx_1, Y_{t_2} \in dx_2) = \\ \mathbb{P}(Y_{t_1} \in dx_1, T_0 \circ \theta_{t_1} < t_2, Y_{t_2} \in dx_2) + \mathbb{P}(Y_{t_1} \in dx_1, T_0 \circ \theta_{t_1} > t_2, Y_{t_2} \in dx_2) \end{aligned}$$

entonces podemos usar la ecuación 8 para ver que la probabilidad buscada es

$$2p(0, x_1, t_1)(p(x_1, x_2, t_2 - t_1) - q(x_1, x_2, t_2 - t_1))$$

Corolario 2.6. *Para $x_1, x_2 > y > 0$ se tiene que:*

$$P_{x_1}(m_Y(t) \in dy, Y_t \in dx_2) = \varphi(x_2 - x_1, y - x_1, t)$$

y

$$P_{x_1}(m_Y(t) = 0, Y_t \in dx_2) = p(x_1, x_2, t) - q(x_1, x_2, t)$$

Notemos que en las demostraciones de esta sección sólo utilizamos la propiedad de Markov del movimiento browniano; por lo tanto se tiene que en general si X_t es un proceso de Markov y $m_X(t_1, t_2)$ denota el ínfimo del proceso en el intervalo $[t_1, t_2]$ entonces:

Lema 2.2.

$$\begin{aligned} \mathbb{P}(X_{t_1} \in dx_1, m_X(t_1, t_2) \in dy, X_{t_2} \in dx_2) = \\ \mathbb{P}(X_{t_1} \in dx_1)P_{x_1}(m_X(0, t_2 - t_1) \in dy, X_{t_2 - t_1} \in dx_2) \end{aligned}$$

3. Excursión normalizada del movimiento browniano.

Ahora tenemos los elementos suficientes para definir el proceso de excursión del movimiento browniano.

En el resto de la sección mantendremos al número $t_0 > 0$ fijo.

Definición 2.3. *Definimos el proceso de excursión del movimiento browniano que contiene a t_0 para $t \geq 0$ como $Z_r = Y_{(g_{t_0+r}) \wedge d_{t_0}}$*

Para poder calcular las distribuciones finito dimensionales del proceso de excursión es necesario eliminar la aleatoriedad del intervalo $[0, L_{t_0}]$, por esta razón calculemos lo siguiente

Proposición 2.5. *Sean $s < t_0, l > t_0 - s$ y $s < t_1 < \dots < t_n < s + l$, entonces*

$$(9) \quad \begin{aligned} \mathbb{P}(g_{t_0} \in ds, Y_{t_1} \in dx_1, m_Y(t_1, t_2) \in dy_1, \dots, Y_{t_n} \in dx_n, L_{t_0} \in dl) = \\ 2p(0, 0, s)g(0, x_1, t_1 - s)\varphi(x_2 - x_1, y_1 - x_1, t_2 - t_1) \dots \\ \varphi(x_n - x_{n-1}, y_{n-1} - x_{n-1}, t_n - t_{n-1})g(0, x_n, l - t_n). \end{aligned}$$

DEMOSTRACIÓN. Es fácil ver que el evento

$$(g_{t_0} < s, Y_{t_1} > x_1, m_Y(t_1, t_2) < y_1, \dots, Y_{t_n} > x_n, L_{t_0} > l)$$

es igual al evento

$$(g_{t_1} < s, Y_{t_1} > x_1, 0 < m_Y(t_1, t_2) < y_1, Y_{t_2} > x_2, \dots, \\ Y_{t_{n-1}} > x_{n-1}, 0 < m_Y(t_{n-1}, t_n) < y_{n-1}, Y_{t_n} > x_n, T_0 \circ \theta_{t_n} > s + l)$$

ya que el último es sólo es una reescritura del primero, es decir, ambos eventos describen la misma trayectoria del proceso Y_t evitando que éste se anule en todo el intervalo $[s, s + l]$. Ahora bien, dado que $\{g_{t_1} < s, Y_{t_1} < x_1\} = \{\inf_{s < r < t_1} Y_r > 0, Y_{r_1} < x_1\} \in \mathcal{F}_{t_1}$, podemos utilizar entonces la propiedad de Markov, vía el lema 1.2, en conjunto con el principio de inducción y las ecuaciones 6, 4 para obtener que la probabilidad del segundo evento esta dada por:

$$\int_0^s \int_{x_1}^{\infty} 2p(0, 0, \omega)g(0, w_1, t_1 - \omega) \int_{x_2}^{\infty} \int_0^{y_1} \varphi(w_2 - w_1, z_1 - w_1, t_2 - t_1) \\ \dots \int_{x_n}^{\infty} \int_0^{y_{n-1}} \varphi(w_n - w_{n-1}, z_{n-1} - w_{n-1}, t_n - t_{n-1}) \\ \int_{l-t_n}^{\infty} g(x_n, 0, \lambda)d\lambda dz_{n-1}dw_n \dots dz_1dw_2dw_1d\omega.$$

Derivando con respecto a las variables que nos interesan obtenemos la ecuación 9. \square

Si integramos la ecuación 9 con respecto a las variables y_1, \dots, y_{n-1} obtenemos:

$$\mathbb{P}(g_{t_0} \in ds, Y_{t_1} \in x_1, \dots, Y_{t_n} \in x_n, L_{t_0} \in dl) \\ = 2p(0, 0, s)g(0, x_1, r_1)q(x_1, x_2, t_2 - t_1) \dots q(x_{n-1}, x_n, t_n - t_{n-1})g(0, x_n, l - r_n)$$

Corolario 2.7. *La densidad finito dimensional del proceso Y_t en el intervalo $[s, s+l]$, bajo la condición $g_{t_0} = s, L_{t_0} = l$ está dada por:*

$$\mathbb{P}(Y_{t_1} \in dx_1, \dots, Y_{t_n} \in dx_n | g_{t_0} = s, L_{t_0} = l) \\ = 4\pi\sqrt{sl^3}p(0, 0, s)g(0, x_1, r_1)q(x_1, x_2, r_2 - r_1) \dots \\ (10) \quad q(x_{n-1}, x_n, r_n - r_{n-1})g(x_n, 0, l - r_n)$$

Calcularemos ahora la distribución finito dimensional para el proceso de excursión y sus mínimos.

Proposición 2.6. *Con las hipótesis de la definición del proceso de excursión y $0 < t_1 < t_2 < \dots < t_n < l$ se tiene que:*

$$\mathbb{P}(Z_{t_1} \in dz_1, m_Z(t_1, t_2) \in dy_1, \dots \\ m_Z(t_{n-1}, t_n) \in dy_{n-1}, Z_{t_n} \in dz_n | g_{t_0} = s, L_{t_0} = dl) \\ = 4\pi\sqrt{sl^3}p(0, 0, s)g(0, z_1, t_1)\varphi(z_2 - z_1, y_1 - z_1, t_2 - t_1) \dots \\ \varphi(z_n - z_{n-1}, y_{n-1} - z_{n-1}, t_n - t_{n-1})g(0, z_n, l - t_n)$$

Antes de hacer la demostración de la proposición trataremos de explicar que significa la expresión de arriba. El lado izquierdo denota la probabilidad de que el proceso Y_t sea cero en el tiempo s , que no se anule en todo el intervalo $[s, s+l]$; y que a los tiempos $s+t_i$ el proceso se encuentre en los valores z_i , mientras que los mínimos en los intervalos $[s+t_i, s+t_{i+1}]$ valen y_i . Finalmente se pide también que el proceso muera en cero al tiempo $l+s$. El lado derecho de la ecuación nos da la probabilidad de que el último retorno a cero antes del

tiempo t_1 ocurra al tiempo s (ecuación 5) que después el proceso vaya de z_i a z_{i+1} en el tiempo $t_{i+1} - t_i$ manteniendo el mínimo en estos intervalos igual a $y_i > 0$, y que finalmente el proceso vaya de z_n a cero en el tiempo $l - t_n$. Pensándolo en el contexto de procesos de Markov esta igualdad tiene mucho sentido.

DEMOSTRACIÓN. Notemos que el resultado será una consecuencia directa del corolario anterior ya que

$$\begin{aligned} & \mathbb{P}(Z_{t_1} \in dz_1, m_Z(t_1, t_2) \in dy_1, \dots, Z_{t_n} \in dz_n | g_{t_0} = s, L_{t_0} = dl) \\ &= \mathbb{P}(Y_{s+t_1} \in dz_1, m_Y(s+t_1, s+t_2) \in dy_1, \dots, Y_{s+t_n} \in dz_n | g_{t_0} = s, L_{t_0} = l) \end{aligned} \quad (11)$$

Y por el corolario anterior la probabilidad anterior está dada por:

$$\begin{aligned} & 4\pi\sqrt{sl^3}p(0, 0, s)g(0, z_1, t_1)\varphi(z_2 - z_1, y_1 - z_1, t_2 - t_1) \dots \\ & \varphi(z_{n-1}, z_n, t_n - t_{n-1})g(z_n, 0, l - t_n) \end{aligned}$$

□

Calcularemos ahora las distribuciones finito dimensionales de la excursión normalizada del movimiento browniano.

Definición 2.4. *Definimos la excursión browniana normalizada como el proceso estocástico $(e_t)_{t \in [0,1]}$ dado por:*

$$e_t = \frac{Z_{tL_{t_0}}}{\sqrt{L_{t_0}}}$$

La excursión normalizada es un proceso con trayectorias continuas definido en el intervalo $[0, 1]$ tal que $e_0 = 0 = e_1$ y está condicionado a permanecer positivo.

Sabemos que el proceso Z_t depende fuertemente de los valores de g_{t_0} y de L_{t_0} , vamos a probar que para la excursión browniana normalizada no es el caso.

Proposición 2.7. *Dados $0 < t_1 < \dots < t_n < 1$ entonces se tiene que:*

$$\begin{aligned} & \mathbb{P}(e_{t_1} \in dx_1, m_e(t_1, t_2) \in dy_1, \dots, m_e(t_{n-1}, t_n) \in dy_{n-1}, e_{t_n} \in dx_n) \\ &= 2\sqrt{2\pi}g(0, x_1, t_1)\varphi(x_2 - x_1, y_1 - x_1, t_2 - t_1) \dots \\ & \varphi(x_n - x_{n-1}, y_{n-1} - x_{n-1}, t_n - t_{n-1})g(0, x_n, 1 - t_n) \end{aligned}$$

DEMOSTRACIÓN. Sean $s < t_0, l > t_0 - s$ entonces:

$$\begin{aligned}
& \mathbb{P}(e_{t_1} > x_1, m_e(t_1, t_2) < y_1, \dots, m_e(t_{n-1}, t_n) < y_{n-1}, e_{t_n} > x_n | g_{t_0} = s, L_{t_0} = l) \\
&= \mathbb{P}(Z_{lt_1} > \sqrt{l}x_1, m_Z(lt_1, lt_2) < \sqrt{l}y_1 \dots, \\
&\quad m_Z(lt_{n-1}, lt_n) < y_{n-1}, Z_{lt_n} > \sqrt{l}x_n | g_{t_0} = s, L_{t_0} = l) \\
&= \int_{\sqrt{l}x_1}^{\infty} 4\pi\sqrt{sl^3}p(0, 0, s)g(0, z_1, lt_1) \int_0^{\sqrt{l}y_1} \varphi(z_2 - z_1, w_1 - z_1, l(t_2 - t_1)) \dots \\
&\quad \int_0^{\sqrt{l}y_{n-1}} \int_{\sqrt{l}x_n}^{\infty} \varphi(z_n - z_{n-1}, w_{n-1} - z_{n-1}, l(t_n - t_{n-1})) \\
&\quad g(z_n, 0, (1 - t_n)l) dz_n dw_{n-1} \dots dw_1 dz_1.
\end{aligned}$$

Si diferenciamos con respecto a x_1, \dots, x_n y con respecto a y_1, \dots, y_{n-1} obtenemos:

$$\begin{aligned}
& \mathbb{P}(e_{t_1} \in dx_1, m_e(t_1, t_2) \in dy_1, \dots, \\
&\quad m_e(t_{n-1}, t_n) \in dy_{n-1}, e_{t_n} \in dx_n | g_{t_0} = s, L_{t_0} = l) \\
&= 4\pi\sqrt{sl^3}p(0, 0, s)\sqrt{l}g(0, \sqrt{l}x_1, lt_1)l\varphi((x_2 - x_1)\sqrt{l}, (y_1 - x_1)\sqrt{l}, (t_2 - t_1)l) \dots \\
&\quad l\varphi(\sqrt{l}(x_n - x_{n-1}), \sqrt{l}(y_{n-1} - x_{n-1}), (t_n - t_{n-1})l)g(\sqrt{l}x_n, 0, (1 - t_n)l).
\end{aligned}$$

Utilizando la proposición 2.3 y el hecho de que

$$p(0, 0, s) = \frac{1}{\sqrt{2\pi s}}$$

puede verse que la ecuación anterior es igual a:

$$\begin{aligned}
& 2\sqrt{2\pi l^3}\sqrt{l}g(0, \sqrt{l}x_1, lt_1)g(0, \sqrt{l}x_n, (1 - t_n)l)\varphi(x_2 - x_1, y_1 - x_1, t_2 - t_1) \dots \\
&\quad \varphi(x_n - x_{n-1}, y_{n-1} - x_{n-1}, t_n - t_{n-1}) \\
&= 2\sqrt{2\pi l^3}\sqrt{l} \frac{\sqrt{l}x_1}{\sqrt{2\pi l^3}(t_1)^3} e^{-\frac{x_1^2 l}{t_1}} \frac{\sqrt{l}x_n}{\sqrt{2\pi l^3}(1 - t_n)^3} e^{-\frac{x_n^2 l}{l(1-t_n)}} \\
&\quad \varphi(x_2 - x_1, y_1 - x_1, t_2 - t_1) \dots \varphi(x_n - x_{n-1}, y_{n-1} - x_{n-1}, t_n - t_{n-1}) \\
&= 2\sqrt{2\pi}g(0, x_1, t_1)\varphi(x_2 - x_1, y_1 - x_1, t_2 - t_1) \dots \\
&\quad \varphi(x_n - x_{n-1}, y_{n-1} - x_{n-1}, t_n - t_{n-1})g(0, x_n, 1 - t_n)
\end{aligned}$$

Como esta densidad condicional no depende de los parametros s, l entonces la densidad finitodimensional está dada por la ecuación propuesta.

□

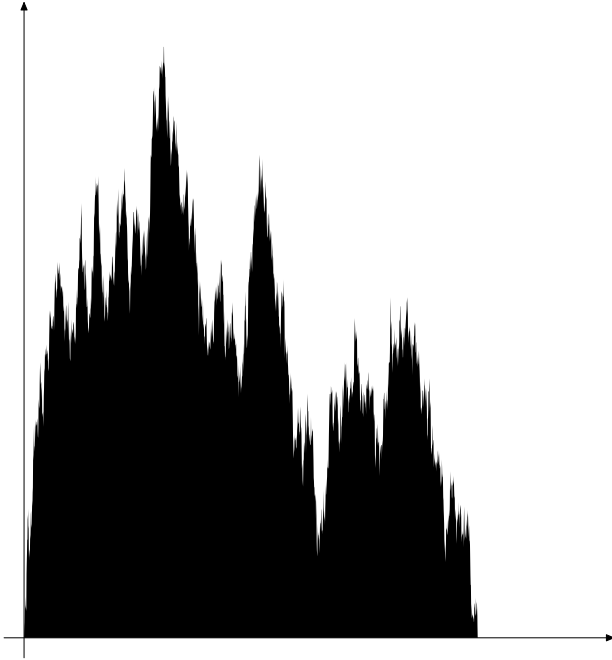


FIGURA 1. Ejemplo de una excursión browniana normalizada

Para concluir este capítulo calcularemos las distribuciones de muestreos uniformes en la excursión browniana normalizada, la razón de ésto es que nos será de gran utilidad al tratar de estudiar distribuciones finito dimensionales en árboles continuos.

Iniciaremos recalcando ciertas relaciones que se dan entre las funciones g y φ introducidas en este capítulo. Sabemos, por lo dicho en el capítulo anterior que:

$$\mathbb{P}_x(T_0 \in dt) = \frac{|x|}{\sqrt{2\pi t^3}} \exp\left(-\frac{x^2}{2t}\right) = g(x, 0, t)$$

Por otro lado sabemos que la densidad conjunta del mínimo acumulativo y el browniano iniciado en x al tiempo t , esta dada por:

$$\begin{aligned}\varphi(y-x, z-x, t) &= \frac{2(x+y-2z)}{\sqrt{2\pi t^3}} \exp\left(-\frac{(x+y-2z)^2}{2t}\right) \\ &= 2g(0, x+y-2z, t) \\ &= 2\mathbb{P}_0(T_{x+y-2z} \in dt) = 2\mathbb{P}_{x+y-2z}(T_0 \in dt)\end{aligned}$$

Lema 2.3. *Sea e una excursión browniana normalizada y U_1, \dots, U_n variables independientes e independientes de e uniformemente distribuidas en $(0, 1)$. Sea $m_i = m_e(U_{(i)}, U_{(i+1)})$ el mínimo de la excursión en el intervalo $[U_{(i)}, U_{(i+1)}]$, donde $U_{(k)}$ denota el k -ésimo estadístico de orden. Se cumple entonces que:*

$$\begin{aligned}&\mathbb{P}(e_{U_{(1)}} \in dx_1, m_1 \in dy_1, \dots, m_{n-1} \in dy_{n-1}, e_{U_{(n)}} \in dx_n) \\ &= n!2^{n+1}(x_1 + \dots + x_n - y_1 - \dots - y_{n-1}) \exp(-2(x_1 + \dots + x_n - y_1 - \dots - y_{n-1})^2)\end{aligned}$$

DEMOSTRACIÓN. Haremos el caso para $n = 2$ con lo cual el lector se convencerá de que el caso general también es cierto. Condicionemos entonces con respecto a los estadísticos de orden. Como los estadísticos son independientes de la excursión esto es posible:

$$\begin{aligned}&\mathbb{P}(e_{U_{(1)}} \in dx_1, m_1 \in dy_1, e_{U_{(2)}} \in dx_2 | U_{(1)} = u, U_{(2)} = v) \\ &= \mathbb{P}(e_u \in dx_1, m_e(u, v) \in dy_1, e_v \in dx_2) \\ &= 2\sqrt{2\pi}g(0, x_1, u)\varphi(x_2 - x_1, y_1 - x_1, v - u)g(x_2, 0, 1 - v) \\ &= 2\sqrt{2\pi}g(0, x_1, u)2g(0, x_1 + x_2 - 2y_1, v - u)g(x_2, 0, 1 - v)\end{aligned}$$

Integrando con respecto a u y v y multiplicando por $2!$ (la densidad de los estadísticos de orden) tenemos que la densidad buscada está dada por:

$$2!4\sqrt{2\pi} \int_0^1 \int_u^1 g(0, x_1, u)g(0, x_1 + x_2 - 2y_1, v - u)g(x_2, 0, 1 - v)dudv$$

pero es bastante claro que $\int_u^1 g(0, x_1 + x_2 - 2y_1, v - u)g(x_2, 0, 1 - v)dv$ es una convolución y por lo tanto es la densidad de la suma de las variables aleatorias $T_{x_1+x_2-2y_1}$ y T_{x_2} , así tenemos que debe ser igual a:

$$\mathbb{P}(T_{x_1+2x_2-2y_1} = T_{x_1+x_2-2y_1} + T_{x_2} \in d(1-u)) = g(0, x_1 + 2x_2 - 2y_1, 1 - u)$$

Por lo tanto la densidad buscada se convierte en

$$\begin{aligned} & 2!4\sqrt{2\pi} \int_0^1 g(0, x_1, u)g(0, x_1 + 2x_2 - 2y_1, 1 - u)du \\ &= 2!4\sqrt{2\pi}g(0, 2x_1 + 2x_2 - 2y_1, 1) \\ &= 2!8(x_1 + x_2 - y_1) \exp(-2(x_1 + x_2 - y_1)^2). \end{aligned}$$

El caso general se demuestra haciendo uso del mismo razonamiento, es decir, utilizando convoluciones y sumas de variables aleatorias. \square

El árbol aleatorio continuo

En este capítulo formalizaremos la idea de lo que es un árbol en ambos sentidos, el discreto y el continuo, así mismo introduciremos el concepto del árbol aleatorio continuo, el cuál es en cierta forma el límite de árboles binarios uniformes. Aplicaremos los conceptos obtenidos en el capítulo anterior para obtener las distribuciones finito dimensionales del árbol aleatorio continuo (CRT por sus siglas en inglés). Tratamos de obtener las distribuciones finito dimensionales del árbol de una forma técnicamente menos demandante de lo que se hace por ejemplo en [LG05], donde se utiliza teoría de excursiones y conceptos más avanzados de teoría de procesos estocásticos, haciendo así inaccesibles los resultados a personas sin éstos conocimientos, los cuales pueden tomar años en ser dominados.

1. Árboles discretos

En esta sección se dará una formalización del concepto de árbol discreto. Aunque no utilizaremos este concepto de manera exhaustiva, será útil para dar cierta intuición sobre lo que es un árbol real.

Primero introduciremos el conjunto de etiquetas :

$$\mathcal{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n$$

Donde por convención $\mathbb{N}^0 = \emptyset$, de esta manera un elemento de \mathcal{U} es de la forma $u = (u_1, u_2, \dots, u_n)$ y hacemos $|u| = n$, donde $|u|$ denota la "generación" de u . Si $u = (u_1, \dots, u_n)$ y $v = (v_1, \dots, v_n)$ entonces su concatenación está dada por $uv = (u_1, \dots, u_n, v_1, \dots, v_n)$; en particular $\emptyset u = u \emptyset = u$. También definimos el mapeo $\pi : \mathcal{U} - \{\emptyset\} \rightarrow \mathcal{U}$ dado por $u = (u_1, \dots, u_n) \rightarrow (u_1, \dots, u_{n-1})$ y diremos que $\pi(u)$ es el padre de u .

Definición 3.1. *Un árbol ordenado, enraizado (finito) es un subconjunto $\mathfrak{t} \subset \mathcal{U}$ que cumple las siguientes propiedades:*

1. $\emptyset \in \mathfrak{t}$ al elemento \emptyset se le conocerá como la raíz del árbol.

2. Si $u \in \mathbf{t}$ no es la raíz, entonces $\pi(u) \in \mathbf{t}$
3. Para cada $u \in \mathbf{t}$ existe un número entero $k_u(\mathbf{t}) \geq 0$ tal que para todo $j \in \mathbb{N}$, $u_j \in \mathbf{t}$ si y sólo si $1 \leq j \leq k_u(\mathbf{t})$.

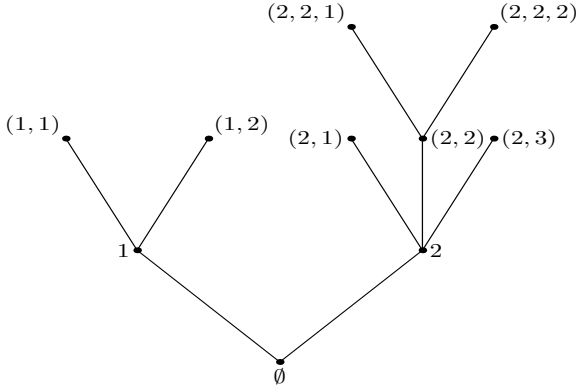


FIGURA 1. Ejemplo de un árbol discreto

Un ejemplo claro se puede ver en la figura 1, el número $k_u(t)$ denota el “número de hijos” de un individuo. Un árbol ordenado discreto puede verse como el árbol genealógico de una familia.

2. Árboles reales

En esta sección introducimos el importante concepto de árbol real o árbol continuo. La idea intuitiva es la siguiente: tomemos un árbol discreto \mathbf{t} y unamos cada par de vértices $v, \pi(v)$ mediante una arista, el árbol real será entonces el conjunto de vértices y todos los puntos de las aristas. Esta idea intuitiva se formaliza más adelante y se generaliza permitiendo que el árbol tenga una infinidad de hojas y que haya vértices con una infinidad de hijos. Una vez definidos los árboles reales nos daremos a la tarea de tratar de codificarlos.

Definición 3.2. *Un espacio métrico compacto (\mathcal{T}, d) se dice que es un árbol continuo si para todo $a, b \in \mathcal{T}$ se cumple que:*

1. Existe una única isometría $f_{ab} : [0, d(a, b)] \rightarrow \mathcal{T}$ tal que $f_{ab}(0) = a$ y $f_{ab}(d(a, b)) = b$
2. Si $g : [0, 1] \rightarrow \mathcal{T}$ es una función continua inyectiva tal que $g(0) = a, g(1) = b$ entonces $g([0, 1]) = f_{ab}([0, d(a, b)])$.

Para hacer que un árbol sea enraizado sólo distinguimos un vértice especial $\rho = \rho(\mathcal{T})$. En adelante todos los árboles serán enraizados. Aunque esta parece una definición bastante abstracta, después de introducir algunos conceptos más veremos que es bastante natural. Si $a, b \in \mathcal{T}$ entonces denotaremos por $[[a, b]]$ a $f_{ab}([0, d(a, b)])$ este será el segmento con la distancia más corta que une a los puntos a y b . En particular $[[\rho, a]]$ es el segmento de línea que une a la raíz con a y lo pensaremos como la línea ancestral de a . Por la inyectividad de f_{ab} se sigue que existe un único punto c tal que $[[\rho, c]] = [[\rho, a]] \cap [[\rho, b]]$ y será su ancestro común más reciente y lo denotaremos por $a \wedge b$. También podemos definir una relación de orden parcial en \mathcal{T} dada por $a \preceq b$ si y sólo si $a \in [[\rho, b]]$. Ahora damos una explicación de la definición de un árbol real. La primera condición nos dice por un lado que la distancia entre un punto a y la raíz está dada por la longitud del segmento $[[\rho, a]]$; para fijar ideas pensemos que \mathcal{T} está en un espacio euclideo, entonces puede pensarse en $[[\rho, a]]$ como el segmento que une a ρ con a y la distancia entre estos 2 puntos como la longitud euclidiana del segmento. Dicho esto entonces la distancia entre a, b está dada por la suma de los segmentos $[[a, c]]$ y $[[c, b]]$. La segunda propiedad nos dice que los únicos caminos sin auto-intersecciones en (\mathcal{T}, d) son de la forma $[[a, b]]$, como puede verse en la figura 2.

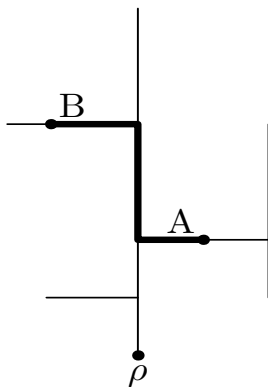


FIGURA 2. Ejemplo de un árbol real donde se puede apreciar que él único camino sin auto-intersecciones entre A y B es de la forma $[[A, B]]$

Por definición la multiplicidad de un vértice a será la cantidad de componentes conexas de $\mathcal{T} \setminus a$, un punto distinto de ρ de multiplicidad 1 se llama hoja.

Debido a que uno de los objetivos principales de este capítulo es establecer la convergencia de una clase de árboles reales, entonces volveremos al espacio de los árboles reales un espacio métrico, y la manera más natural de hacerlo será con la métrica de Gromov-Hausdorff entre espacios métricos compactos. Si (E, δ) es un espacio métrico denotaremos por $\delta_H(K, K')$ a la distancia de Hausdorff entre subespacios compactos de E :

$$\delta_H(K, K') = \inf\{\epsilon > 0 \mid K \subset U_\epsilon(K'), K' \subset U_\epsilon(K)\}$$

donde $U_\epsilon(K) = \{x \mid \delta(x, K) < \epsilon\}$.

Sean entonces $\mathcal{T}, \mathcal{T}'$ dos árboles reales con respectivas raíces ρ, ρ' . Definimos la distancia entre ambos por:

$$d_{GH}(\mathcal{T}, \mathcal{T}') = \inf\{\delta_H(\phi(\mathcal{T}), \phi'(\mathcal{T}')) \vee \delta(\phi(\rho), \phi'(\rho'))\}$$

Donde $\phi : \mathcal{T} \rightarrow E$ y $\phi' : \mathcal{T}' \rightarrow E$ corren sobre todos los encajes en todos los espacios métricos posibles E . Diremos que dos árboles $\mathcal{T}_1, \mathcal{T}_2$ son equivalentes si existe una isometría que preserve la raíz entre ellos, es claro que en este caso $d_{GH}(\mathcal{T}_1, \mathcal{T}_2) = 0$. Puede verse en [Gro99] que d_{GH} define una métrica en el conjunto de clases de equivalencia de los árboles reales.

2.1. Codificando árboles reales mediante funciones continuas.

Sea $g : [0, \infty) \rightarrow [0, \infty)$ una función continua con soporte compacto y tal que $g(0) = 0$. Sea $m_g(s, t)$ como en la sección anterior el mínimo de la función g en el intervalo $[s \wedge t, s \vee t]$; definimos una pseudo-métrica en $[0, \infty)$ como sigue

$$\delta_g(s, t) = g(s) + g(t) - 2m_g(s, t)$$

Es claro que esta función es mayor o igual que 0 y es simétrica. Veamos ahora que también cumple la desigualdad del triángulo: sean $s < t < u$, entonces

$$\begin{aligned} m_g(s, u) &= m_g(s, t) \\ &\quad \text{ó} \\ m_g(s, u) &= m_g(t, u) \end{aligned}$$

sin pérdida de generalidad supongamos lo segundo, entonces:

$$\begin{aligned} \delta_g(s, u) &= g(s) + g(u) - 2m_g(s, u) \\ &\leq g(s) + g(u) - 2m_g(s, u) + 2g(t) - 2m_g(s, t) \\ &= g(s) + g(t) - 2m_g(s, t) + g(t) + g(u) - 2m_g(t, u) \\ &= \delta_g(s, t) + \delta_g(t, u) \end{aligned}$$

Si definimos una relación de equivalencia en $[0, \infty)$ (\sim_g) como sigue : $s \sim_g t$ si y sólo si $g(s) = g(t) = m_g(s, t)$ entonces el espacio $\mathcal{T}_g = [0, \infty) / \sim_g$ se convierte un espacio métrico con la métrica d_g inducida por la pseudo-métrica δ_g . Este espacio métrico es un árbol enraizado, donde la raíz es la clase de equivalencia del 0 la cual denotaremos por ρ , como veremos a continuación.

Sea $p_g : [0, \infty) \rightarrow \mathcal{T}_g$ la proyección canónica sobre el espacio cociente. Sea ζ el supremo del soporte de g , entonces es claro que $p_g(s) = \rho$ para toda $s \geq \zeta$, en particular $p([0, \zeta]) = \mathcal{T}_g$ es compacto.

Para demostrar que \mathcal{T}_g es en efecto un árbol vamos a necesitar de un lema, el cual, aunque interesante por si mismo, tiene una demostración bastante técnica; sin embargo nos da mucha información sobre la estructura de \mathcal{T}_g . Aunque en la demostración se dirá que el objeto \mathcal{T}_g es de hecho un árbol, el lector deberá notar que es sólo por abuso de lenguaje y nunca se utilizan las propiedades dadas en la definición de árbol real.

Lema 3.1. (*Reenraizamiento de un árbol*)

Sea s_0 un elemento en $[0, \zeta]$, para todo $r \geq 0$ sea \bar{r} el único elemento en $[0, \zeta]$ que hace que $r - \bar{r}$ sea un múltiplo entero de ζ . Definimos el mapeo

$$g^{s_0}(s) = g(s_0) + g(\overline{s_0 + s}) - 2m_g(s_0, \overline{s_0 + s})$$

Para todo $s \in [0, \zeta]$ y $g^{s_0}(s) = 0$ para toda $s \geq \zeta$. Entonces g^{s_0} es una función continua con soporte compacto tal que $g^{s_0}(0) = 0$ por lo tanto podemos definir $\mathcal{T}_{g^{s_0}}$ más aún tenemos que

$$(12) \quad d_{g^{s_0}}(s, t) = d_g(\overline{s_0 + s}, \overline{s_0 + t})$$

Además existe una única isometría R entre $\mathcal{T}_g, \mathcal{T}_{g^{s_0}}$, tal que

$$(13) \quad R(p_{g^{s_0}}(s)) = p_g(\overline{s_0 + s})$$

en particular $R(p_{g^{s_0}}(0)) = p_g(\overline{s_0}) = p_g(s_0)$, es decir que $T_{g^{s_0}}$ es el árbol T_g reenraizado en $p_g(s_0)$ como se ve en la figura 3

DEMOSTRACIÓN. Es inmediato que la función g^{s_0} cumple las propiedades necesarias para definir correctamente al árbol $\mathcal{T}_{g^{s_0}}$. Ahora bien probaremos la ecuación 12 para el caso en el que $s, t \in [0, \zeta - s_0]$. Con estas condiciones pueden ocurrir dos cosas.

Si $m_g(s_0 + s, s_0 + t) \geq m_g(s_0, s_0 + s)$ entonces $m_g(s_0, s_0 + r) = m_g(s_0, s_0 + s)$ para toda $r \in [s, t]$ y por lo tanto

$$m_{g^{s_0}}(s, t) = g(s_0) + m_g(s_0 + s, s_0 + t) - 2m_g(s_0, s_0 + s).$$

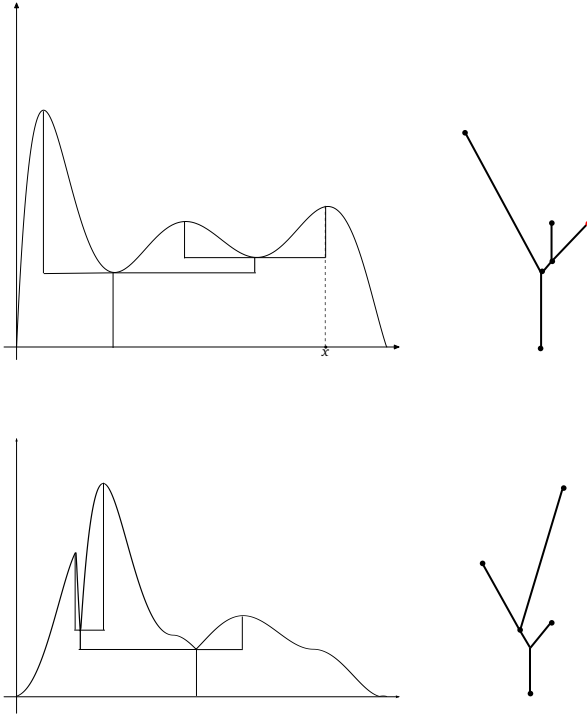


FIGURA 3. Una función g reenraizada en el punto x , así mismo se puede ver el árbol codificado por la función g y el reenraizamiento del árbol en el vértice rojo para obtener el árbol codificado por la función g^{s_0} .

De aquí se sigue que :

$$\begin{aligned}
d_{g^{s_0}}(s, t) &= \\
g^{s_0}(s) + g^{s_0}(t) - 2m_{g^{s_0}}(s, t) &= \\
2g(s_0) + g(s_0 + s) - 2m_g(s_0, s_0 + s) + g(s_0 + t) - 2m_g(s_0, s_0 + t) \\
&\quad - 2g(s_0) - 2m_g(s_0 + s, s_0 + t) + 4m_g(s_0, s_0 + s) \\
&= g(s_0 + s) + g(s_0 + t) - 2m_g(s_0 + s, s_0 + t) \\
&= d_g(s_0 + s, s_0 + t)
\end{aligned}$$

Si $m_g(s_0 + s, s_0 + t) < m_g(s_0, s_0 + s)$, entonces g^{s_0} se minimiza donde la diferencia

$$g(s_0 + r) - 2m_g(s_0 + r)$$

es más pequeña. Veamos donde ocurre esto. Sabemos que $g(s_0) \geq m_g(s_0, s_0 + s)$ y como además $m_g(s_0 + s, s_0 + t) < m_g(s_0, s_0 + s)$ entonces existe un primer punto $r_0 \in [s, t]$ tal que

$$g(s_0 + r_0) = m_g(s_0, s_0 + s) = m_g(s_0, s_0 + r_0)$$

Vamos a probar que $g^{s_0}(r_0) = m_{g^{s_0}}(s, t)$.

Tomemos $r \in [s, t]$ y supongamos que $g(s_0 + r) \geq m_g(s_0, s_0 + s)$, entonces:

$$\begin{aligned}
g(s_0 + r) - 2m_g(s_0, s_0 + r) &\geq m_g(s_0, s_0 + s) - 2m_g(s_0, s_0 + s) \\
&= g(s_0 + r_0) - 2m_g(s_0, s_0 + r_0)
\end{aligned}$$

Por otro lado si $g(s_0 + r) < m_g(s_0, s_0 + s)$ entonces:

$$\begin{aligned}
g(s_0 + r) - 2m_g(s_0, s_0 + r) &\geq -g(s_0 + r) \\
&> -m_g(s_0, s_0 + s) \\
&= g(s_0 + r_0) - 2m_g(s_0, s_0 + r_0)
\end{aligned}$$

Por lo tanto:

$$m_{g^{s_0}}(s, t) = g(s_0) - m_g(s_0, s_0 + s)$$

y finalmente llegamos a que:

$$\begin{aligned}
d_{g^{s_0}}(s, t) &= \\
&= g(s_0 + s) - 2m_g(s_0, s_0 + s) + g(s_0 + t) - 2m_g(s_0, s_0 + t) + 2m_g(s_0, s_0 + s) \\
&= d_g(s_0 + s, s_0 + t)
\end{aligned}$$

Las otras posibilidades para s, t se pueden tratar de manera similar. Del hecho de que $d_{g^{s_0}}(s, t) = 0$ sii $d_g(\overline{s_0 + s}, \overline{s_0 + t}) = 0$ se sigue que la relación en 13 define una única función biyectiva R . Además es claro que se trata de una isometría por la ecuación 12. □

Ahora pasamos a demostrar el hecho de que \mathcal{T}_g es un árbol en el sentido de la definición dada al inicio de esta sección. Iniciaremos con la definición de ciertos conceptos que nos harán más fácil demostrar este hecho.

Para cualesquiera $\sigma, \sigma' \in \mathcal{T}_g$ definimos una relación de orden parcial $\sigma' \preceq \sigma$ si y sólo si $d(\rho, \sigma) = d(\rho, \sigma') + d(\sigma', \sigma)$. Si $\sigma' = p_g(s), \sigma = p_g(t)$, entonces es claro que $\sigma' \preceq \sigma$ si y sólo si $m_g(s, t) = g(s)$. Ahora bien para todo σ_0, σ definimos $[[\sigma_0, \sigma]] := \{\sigma' \mid d(\sigma_0, \sigma) = d(\sigma_0, \sigma') + d(\sigma', \sigma)\}$ de esta manera si r es tal que $m_g(s, t) = g(r)$ entonces también es claro que $[[\rho, \sigma]] \cap [[\rho, \sigma']] = [[\rho, \gamma]]$ donde $\gamma = p_g(r)$ en este caso denotamos a γ por $\sigma \wedge \sigma'$.

Definimos $\mathcal{T}_g[\sigma] = \{\sigma' \mid \sigma' \preceq \sigma\}$ entonces los conjuntos $\mathcal{T}_g \setminus \mathcal{T}_g[\sigma], \mathcal{T}_g[\sigma] \setminus \{\sigma\}$ son abiertos disjuntos. Para ver que el primero de ellos es abierto notemos que si $p_g(s) = \sigma$ entonces el conjunto $E = \{u \in [0, \zeta] \mid m_g(s, u) = g(s)\}$ es compacto y además $p_g(E) = \mathcal{T}_g[\sigma]$, para ver esto sea $(u_n) \in E$ una sucesión convergente a u , entonces $m_g(s, u_n) = g(s)$ pero por la continuidad de la función $m_g(s, \cdot)$ entonces $m_g(s, u) = g(s)$ por lo tanto $u \in E$. Esto nos dice que E es cerrado y por lo tanto compacto; además por lo dicho en el párrafo anterior sabemos que $\sigma \preceq p_g(u)$ sii $m_g(s, u) = g(s)$, y por lo tanto $p_g(E) = \mathcal{T}_g[\sigma]$. Así $\mathcal{T}_g \setminus \mathcal{T}_g[\sigma]$ es abierto. Ahora bien es fácil por la definición de \mathcal{T}_g demostrar que $\mathcal{T}_g[\sigma] \setminus \sigma$ es abierto. Consideremos $\mathcal{U} := p_g^{-1}(T_g[\sigma] \setminus \sigma)$; por lo dicho anteriormente es claro que

$$\mathcal{U} = \{u \in [0, \zeta] \mid m_g(s, u) = g(s) < g(u)\}$$

Tomemos $u \in \mathcal{U}$, entonces $g(u) > g(s)$ y $g(t) \geq g(s)$ para toda $t \in [u, s]$. Por la continuidad de g existe $\epsilon > 0$ de tal que $g(t) \geq g(s)$ para toda $t \in [u - \epsilon, s]$ por lo tanto $B_\epsilon(u) \subset \mathcal{U}$ y con esto se prueba que la imagen inversa de $\mathcal{T}_g[\sigma] \setminus \sigma$ es abierta en \mathbb{R} y por definición de topología cociente, esto implica que $\mathcal{T}_g[\sigma] \setminus \sigma$ es abierto en \mathcal{T}_g .

Comenzaremos la tan anunciada demostración de que \mathcal{T}_g es un árbol. Sean $\sigma_1, \sigma \in \mathcal{T}_g$; por el lema 3.1 podemos suponer que $\sigma_1 = \rho$. Por lo tanto tenemos que probar la existencia de una única isometría $f = f_{\rho\sigma}$ de $[0, d_g(\rho, \sigma)]$ en \mathcal{T}_g tal que $f(0) = \rho$ y $f(d_g(\rho, \sigma)) = \sigma$. Para lograr esto, sea $s \in p_g^{-1}(\sigma)$, entonces $g(s) = d_g(\rho, \sigma)$; para toda $a \in [0, d_g(\rho, \sigma)]$ definimos:

$$v(a) = \inf\{r \in [0, s] \mid m_g(r, s) = a\}$$

Note que $g(v(a)) = a$. Definimos entonces la función f cómo $f(a) = p_g(v(a))$. Se tiene de esta manera que $f(0) = \rho$ y que $f(d_g(\rho, \sigma)) = \sigma$, la última igualdad se obtiene pues $m_g(v(g(s)), s) = g(s)$ implica que $p_g(v(g(s))) = p_g(s) = \sigma$. Además si $a, b \in [0, d_g(\rho, \sigma)]$ con $a \leq b$ es inmediato que $m_g(v(a), v(b)) = a$, entonces $d_g(f(a), f(b)) = g(v(a)) + g(v(b)) - 2a = b - a$, y esto prueba que f es una isometría.

Para demostrar la unicidad supongamos que existe otra isometría \tilde{f} que cumple las mismas propiedades que f . Si $a \in [0, d_g(\rho, \sigma)]$ entonces

$$d_g(\tilde{f}(a), \sigma) = d_g(\rho, \sigma) - a = d_g(\rho, \sigma) - d_g(\rho, \tilde{f}(a))$$

Por lo tanto $\tilde{f}(a) \preceq \sigma$. Sea t tal que $p_g(t) = \tilde{f}(a)$. Note que $g(t) = d_g(\rho, p_g(t)) = a$. Como $\tilde{f}(a) \preceq \sigma$ se sigue que $g(t) = m_g(t, s)$. Por otro lado también sabemos que $g(t) = a = g(v(a)) = m_g(v(a), s)$. De esto se sigue que $\tilde{f}(a) = p_g(t) = p_g(v(a)) = f(a)$ y esto completa la prueba.

Del párrafo anterior podemos obtener algo más, en particular que $f([0, d_g(\rho, \sigma)]) = [[\rho, \sigma]]$: En efecto, ya sabemos que $f([0, d_g(\rho, \sigma)]) \subset [[\rho, \sigma]]$. Inversamente si $\eta \preceq \sigma$, el final del párrafo anterior nos asegura que $\eta = f(d_g(\rho, \eta))$ concluyendo la prueba del comentario.

Para probar la segunda propiedad de la definición de árbol real elegimos una función inyectiva y continua $q : [0, 1] \rightarrow \mathcal{T}_g$, (sin pérdida de generalidad nuevamente podemos suponer que $q(0) = \rho$ y sea $\sigma = q(1)$), veamos que cumple que $q[0, 1] = [[\rho, \sigma]]$. Supongamos primero que existe un punto $\eta \in [[\rho, \sigma]] \setminus q[0, 1]$, entonces $q[0, 1]$ se queda contenido en la unión de los abiertos $\mathcal{T}_g \setminus \mathcal{T}_g[\eta]$, $\mathcal{T}_g[\eta] \setminus \eta$ además $q(0) = \rho \in \mathcal{T}_g \setminus \mathcal{T}_g[\eta]$ y $q(1) = \sigma \in \mathcal{T}_g[\eta] \setminus \eta$. Esto contradice el hecho de que $q[0, 1]$ es conexo.

Inversamente supongamos que $\exists a \in (0, 1)$ tal que $q(a)$ no pertenece a $[[\rho, \sigma]]$. Sea $\eta = q(a)$ y $\gamma = \eta \wedge \sigma$. De la definición de γ es inmediato que ver que $\gamma \in [[\rho, \eta]]$ y que $d_g(\eta, \sigma) = d_g(\eta, \gamma) + d_g(\gamma, \sigma)$; por lo tanto $\gamma \in [[\eta, \sigma]]$. Lo primero implica que γ pertenece a $q[0, a]$ por lo argumentado en el párrafo anterior. Lo segundo implica que via un renaizamiento, $\gamma \in q[a, 1]$ lo cual, por lo la inyectividad de q , no puede pasar a menos que $\gamma = q(a) = \eta$, contradiciendo el hecho de que η no pertenece a $[[\rho, \sigma]]$

Con esto hemos encontrado una manera práctica de trabajar con los árboles reales, de hecho todos los árboles reales enraizados pueden ser codificados mediante una función *caglad*; pero no se hará la demostración pues el objetivo de este trabajo no es sumergirse demasiado en las propiedades de estos espacios métricos, sino saber que es lo que pasa cuando se les da cierta aleatoriedad. Para una demostración de este hecho se refiere al lector al artículo [Duq08], donde puede encontrar una explicación muy detallada sobre árboles reales y su caracterización por medio de funciones *caglad*.

Sabiendo ahora que los árboles pueden ser codificados por funciones continuas una pregunta natural sale a relucir: ¿qué relación hay entre la distancia de Gromov-Hausdorff y la distancia entre las funciones codificantes? Para responder ésto utilizaremos una definición alternativa para d_{GH} . Recordemos que si $(\mathcal{T}_1, d_1), (\mathcal{T}_2, d_2)$ son dos espacios métricos compactos, una correspondencia entre T_1, T_2 es un subconjunto \mathcal{R} de $\mathcal{T}_1 \times \mathcal{T}_2$ para el cual se cumple que para todo $x_1 \in \mathcal{T}_1$ existe al menos un $x_2 \in \mathcal{T}_2$ tal que $(x_1, x_2) \in \mathcal{R}$ y análogamente para todo $y_2 \in \mathcal{T}_2$ existe un $y_1 \in \mathcal{T}_1$ tal que $(y_1, y_2) \in \mathcal{R}$. La distorsión de una correspondencia se define como

$$dist(\mathcal{R}) := \sup\{|d_1(x_1, x_2) - d_2(y_1, y_2)| \mid (x_1, y_1), (x_2, y_2) \in \mathcal{R}\}$$

Así se tiene entonces que para $(\mathcal{T}, \rho), (\mathcal{T}', \rho')$ dos árboles enraizados:

$$(14) \quad d_{GH}(\mathcal{T}, \mathcal{T}') = \frac{1}{2} \inf_{\mathcal{R} \in \mathcal{C}(\mathcal{T}, \mathcal{T}'), (\rho, \rho') \in \mathcal{R}} dist(\mathcal{R})$$

Donde $\mathcal{C}(\mathcal{T}, \mathcal{T}')$ es el conjunto de todas las correspondencias entre \mathcal{T} y \mathcal{T}' . Se refiere al lector al lema 2.3 de [EPW06] para ver la demostración de que ambas definiciones son equivalentes.

Lema 3.2. *Sean g y g' dos funciones con soporte compacto de $[0, \infty)$ a $[0, \infty)$ tales que $g(0) = g'(0) = 0$. Sean además $\mathcal{T} = \mathcal{T}_g$ y $\mathcal{T}' = \mathcal{T}_{g'}$. Entonces*

$$d_{GH}(\mathcal{T}, \mathcal{T}') \leq 2\|g - g'\|$$

DEMOSTRACIÓN. Definimos la siguiente correspondencia entre \mathcal{T} y \mathcal{T}' :

$$\mathcal{R} := \{(\sigma, \sigma') \in \mathcal{T} \times \mathcal{T}' \mid \sigma = p_g(t), \sigma' = p_{g'}(t), \text{ para algun } t \in [0, \infty)\}$$

Para acotar la distorsión de \mathcal{R} notemos que si $(\sigma, \sigma'), (\eta, \eta') \in \mathcal{R}$ entonces existen s, t tales que $\sigma = p_g(s), \sigma' = p_{g'}(s)$ y $\eta = p_g(t), \eta' = p_{g'}(t)$ pero además sabemos que

$$\begin{aligned}d_g(\sigma, \eta) &= g(s) + g(t) - 2m_g(s, t) \\d_{g'}(\sigma', \eta') &= g'(s) + g'(t) - 2m_{g'}(s, t)\end{aligned}$$

Por lo tanto

$$\begin{aligned}2d_{GH}(\mathcal{T}, \mathcal{T}') &\leq |d_g(\sigma, \eta) - d_{g'}(\sigma', \eta')| \\ &\leq |g(s) - g'(s)| + |g(t) - g'(t)| + 2|m_g(s, t) - m_{g'}(s, t)| \\ &\leq 4\|g - g'\|\end{aligned}$$

Y esto demuestra el lema. □

Como último análisis de la codificación de árboles estudiaremos brevemente las componentes conexas de $\mathcal{T}_g \setminus \sigma$. Por simplicidad de notación quitaremos el subíndice g de todos nuestros objetos.

Notemos primero que $\mathcal{T} \setminus \mathcal{T}[\sigma]$ es siempre conexo ya que: si $\phi, \phi' \notin \mathcal{T}[\sigma]$ entonces es claro que $\phi \wedge \phi' \notin \mathcal{T}[\sigma]$. Por lo tanto el camino obtenido mediante la concatenación de $[[\phi, \phi \wedge \phi']]$ y $[[\phi \wedge \phi', \phi']]$ conecta a los 2 puntos, i.e $\mathcal{T} \setminus \mathcal{T}[\sigma]$ es conexo.

Ahora bien sea $\rho \neq \sigma \in \mathcal{T}$; definamos:

$$\begin{aligned}I &= \inf \{t \in [0, \zeta] | p(t) = \sigma\} \\ S &= \sup \{t \in [0, \zeta] | p(t) = \sigma\}\end{aligned}$$

Demostraremos que el conjunto E es igual al intervalo $[I, S]$, recordemos que $E = \{x | m_g(x, s) = g(s)\}$ donde $p(s) = \sigma$. Sea $x \in [I, S]$ y supongamos que $m_g(x, s) \neq g(s)$. Sin pérdida de generalidad supongamos que $s < x$ entonces $\exists t \in (s, x]$ tal que $g(s) > g(t) = m_g(x, s)$ por otro lado por la propiedad de S existe $s' \leq S$ tal que $p(s') = \sigma$; sin embargo $g(t) = m_g(x, s) \geq m_g(s, s')$ y como $s \sim s'$ se tiene que $g(s) = g(s') = m_g(s, s')$, por lo tanto $g(s) > g(s')$ y esto es claramente una contradicción. Inversamente tomemos ahora $x \in E$ con $x \geq s$. Si $g(x) = g(s)$ entonces $p(x) = \sigma$ y por lo tanto $x \in [I, S]$. Si $g(x) < g(s)$ entonces $m_g(x, s) < g(s)$ lo cual es una contradicción. Por último si $g(x) > g(s)$ entonces $I \leq x$, además por la continuidad y el hecho de que $g(\zeta) = 0$ se tiene que el conjunto $A = \{t | x < t \text{ y } g(t) = g(s)\}$ es no vacío. Sea $I' = \inf A$ por continuidad $g(I') = g(s)$ además $\forall y \in [x, I'] g(y) > g(s)$ de lo contrario I' no sería ínfimo, por lo tanto $m_g(s, I') = g(s)$ y esto implica que $I' \in [I, S]$ y a su vez esto implica que $x \in [I, S]$. Así $[I, S] = E$, pero ya sabíamos que $p[I, S] = p(E) = \mathcal{T}[\sigma]$.

Con esto en mente demostraremos que la imagen bajo p de las componentes conexas de $[I, S] \setminus p^{-1}(\sigma)$ son las componentes conexas de $\mathcal{T}[\sigma] \setminus \sigma$. Claramente las imágenes de las componentes son conexas y hacen el total de $\mathcal{T}[\sigma] \setminus \sigma$ sólo nos falta probar que son disjuntas.

Sean $(a_1, b_1), (a_2, b_2)$ dos componentes conexas tales que $b_1 < a_2$ y sean $x \in (a_1, b_1), y \in (a_2, b_2)$ tales que $p(x) = p(y)$. Esto implica que:

$$(15) \quad g(x) = g(y) = m_g(x, y)$$

Dado que $b_1 \notin [I, S] \setminus p^{-1}(\sigma)$ se sigue que $p(b_1) = \sigma$ y por lo tanto $m_g(x, y) \leq m_g(x, b_1) = g(b_1)$ además si $g(x) < g(b_1)$ entonces existe $t \in (x, b_1)$ tal que $m_g(x, b_1) = g(t) < g(s)$ lo cual es una contradicción pues $t \in [I, S]$ se sigue entonces que $g(x) = g(b_1)$ y combinando este resultado con la ecuación 15 obtenemos que $p(x) = p(y) = \sigma$ lo cual es una clara contradicción. Por lo tanto $p(a_1, b_1) \cap p(a_2, b_2) = \emptyset$. Tenemos entonces clasificadas todas las componentes conexas de $\mathcal{T} \setminus \sigma$ a saber $\mathcal{T} \setminus \mathcal{T}[\sigma]$ y las enunciadas anteriormente.

También por todo lo dicho anteriormente es fácil establecer el siguiente:

Lema 3.3. *Un punto σ en un árbol es un nodo (un vértice de grado al menos 3) sii $[I, S]$ tiene al menos dos componentes conexas, lo cual ocurre sii existe $s \in p^{-1}(\sigma)$ que es un mínimo local de g . Además si n es el grado del nodo y m es la cantidad de puntos en $p^{-1}(\sigma)$ donde se alcanza un mínimo, entonces $m+2=n$.*

Para terminar esta sección definiremos lo que es un subárbol y daremos una nueva manera de codificarlos, la cual nos será muy útil al momento de trabajar con árboles aleatorios.

Sea \mathcal{T} un árbol real, y $x_1, \dots, x_n \in \mathcal{T}$; entonces el subárbol generado por estos vértices se define como

$$\mathcal{T}(x_1, \dots, x_n) = \bigcup_{m=1}^n [[\rho, x_m]]$$

Es fácil ver que este objeto dotado con la métrica d es un árbol y además puede ser representado por un árbol ordenado finito y por el conjunto de distancias indexado por los vértices del árbol discreto. Para formalizar lo dicho, se dará una construcción recursiva de el árbol discreto y de la familia de distancias (alturas).

Sea g una función continua de $[0, \infty)$ en $[0, \infty)$, si $t_0 < t_1 \dots < t_p$ definimos el **árbol marcado**

$$\theta(g; t_1, \dots, t_n) = (\tau(g; t_1, \dots, t_p), (h_v)_{v \in \tau(g; t_1, \dots, t_p)})$$

Donde $\tau(g; t_1, \dots, t_p)$ es un árbol ordenado finito como los definidos en la primera sección de este capítulo, y $h_v > 0$ denota la distancia entre v y $\pi(v)$ si $v \neq \emptyset$ y en este caso h_\emptyset denota la longitud de la raíz. En la figura 4 se puede ver un ejemplo de $\theta(g; t_1, t_2, t_3)$ un árbol con 3 hojas.

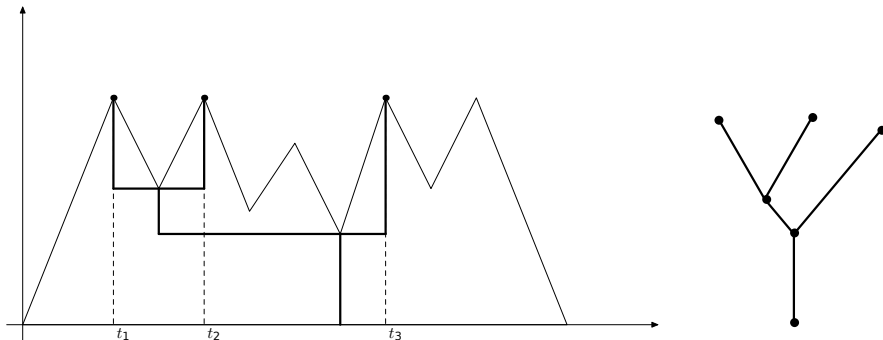


FIGURA 4. En la figura puede verse el árbol $\theta(g; t_1, t_2, t_3)$, nótese que sólo nos interesa la estructura de los mínimos entre estos 3 puntos deshaciéndonos de todo lo demás.

Con esto en mente para $p = 1$ definimos $\tau(g; t_1) = \emptyset$ y $h_\emptyset = g(t_1)$. Si $p \geq 2$, supongamos que el árbol $\theta(g; t_1, \dots, t_j)$ ya está definido para toda $j < p$ entonces existe un número $k \in \{1, \dots, p - 1\}$ y k números $1 \leq i_1 < \dots < i_k \leq p - 1$ tales que $m_g(t_i, t_{i+1}) = m(t_1, t_p)$ sii $i \in \{i_1, \dots, i_k\}$, (el número k denota el número de hijos del \emptyset); por convención se definen $i_0 = 0$ e $i_{k+1} = p$. Para toda $l \in \{1, \dots, k + 1\}$ definimos g^l con la siguiente fórmula:

$$g^l(t) = g((t \vee t_{i_{l-1}+1}) \wedge t_{i_l}) - m_g(t_1, t_p)$$

Como $i_l < p$ entonces podemos definir los árboles $\tau(g^l; t_{i_{l-1}+1}, \dots, t_{i_l})$. Finalmente podemos definir el árbol $\tau(g; t_1, \dots, t_p)$ como la concatenación de todos los árboles $\tau(g^l, t_{i_{l-1}+1}, \dots, t_{i_l})$ donde $1 \leq l \leq k$; esto quiere decir que

$$\tau(g; t_1, \dots, t_p) = \{\emptyset\} \bigcup_{l=1}^k \{lu \mid u \in \tau(g^l; t_{i_{l-1}+1}, \dots, t_{i_l})\}$$

En la figura 4 se puede ver esta representación para $p = 3$.

Más aún si para cada $1 \leq l \leq k$

$$\theta(g^l, t_{i_{l-1}+1}, \dots, t_{i_l}) = \left(\tau(g^l, t_{i_{l-1}+1}, \dots, t_{i_l}), (h_v^l)_{v \in \tau(g^l, t_{i_{l-1}+1}, \dots, t_{i_l})} \right)$$

definimos las marcas $h_v = h_u^l$ si $v = lu$ con $u \in \tau(g^l, t_{i_{l-1}+1}, \dots, t_{i_l})$ y $h_\emptyset = m_g(t_1, t_p)$. Ahora bien es claro que si g codifica al árbol \mathcal{T}_g entonces el árbol marcado $\theta(\tau(g; t_1, \dots, t_p), (h_v)_{v \in \tau(g; t_1, \dots, t_p)})$ codifica al árbol $\mathcal{T}_g(p_g(t_1), \dots, p_g(t_p))$. Para fijar ideas digamos que si en \mathbb{R}^2 a cada vértice v le pegamos un segmento de longitud h_v que lo una con su padre de forma que ninguno de estos segmentos se intersecte, entonces la unión de todos estos segmentos será un representante de la clase de equivalencia de $\mathcal{T}_g(p_g(t_1), \dots, p_g(t_p))$.

3. El árbol aleatorio continuo

Denotemos por \mathbf{A}_p al conjunto de árboles ordenados enraizados de p hojas (un vértice $u \in \mathbf{t}$ se llama hoja si $k_u(\mathbf{t}) = 0$, es decir, si no tiene hijos), análogamente el subconjunto de árboles marcados Θ_p estará formado por todos los árboles de la forma $\theta = (\tau, (h_v)_{v \in \tau})$ donde $\tau \in \mathbf{A}$ y $h_v > 0$. Asimismo definimos los subconjuntos \mathbf{A}_p^{bin} , Θ_p^{bin} de árboles binarios con p hojas, recordemos que un árbol binario es un árbol ordenado \mathbf{t} para el cual se cumple que $k_u(\mathbf{t}) = 0, 2$ para todo $u \in \mathbf{t}$. Es un hecho conocido que :

$$|\mathbf{A}_p^{bin}| = \frac{(2p-2)!}{(p-1)!p!} =: c_p$$

Estamos ahora en posición de definir el árbol aleatorio continuo (de ahora en adelante CRT) y de calcular sus distribuciones finito dimensionales, esto quiere decir muestrear al árbol y obtener la distribución de el árbol generado por este muestreo.

Sea e una excursión browniana, definimos entonces el CRT como

$$\mathcal{T}_e := p_e([0, 1])$$

Regresando al problema de calcular las leyes finito dimensionales, las cuales por definición son las leyes de los subárboles $\mathcal{T}_e(p_e(U_1), \dots, p_e(U_n))$. Pero además por lo dicho en la sección anterior esto es equivalente a calcular la distribución de:

$$(\tau(e_{U_1}, \dots, e_{U_n}), (h_v)_{v \in \tau})$$

Notemos algunas propiedades básicas de estos muestreos. Primero veamos que con probabilidad 0 alguno de los n puntos elegidos es un nodo. Supongamos que si entonces debe existir un punto $U_{(i)}$ donde se alcanza el mínimo de la excursión entre $U_{(i-1)}$ y $U_{(i+1)}$ de aquí se sigue que debería ocurrir que

$$m_e(U_{(i-1)}, U_{(i)}) = e_{U_{(i)}}$$

Pero esto es un evento de probabilidad cero.

Veamos ahora que si η es un nodo del árbol muestreado, entonces el grado de η es 3. Supongamos que el grado del nodo es mayor a 3 entonces deben existir i tal que

$$m_e(U_{(i-1)}, U_{(i)}) = m_e(U_{(i)}, U_{(i+1)})$$

Nuevamente este evento ocurre con probabilidad 0; por lo tanto todos los nodos son de orden 3. Esto quiere decir que con probabilidad 1 los árboles muestreados son árboles binarios con n hojas.

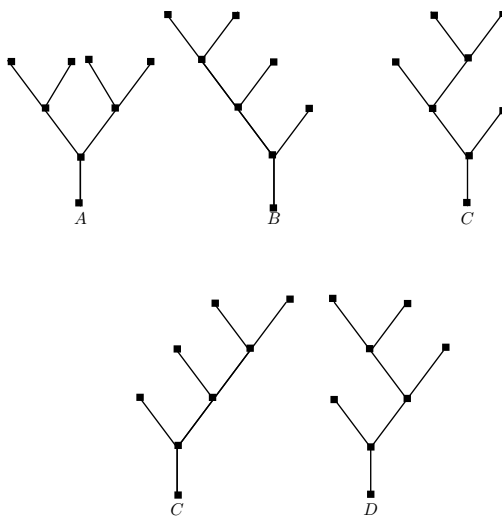


FIGURA 5. Las distintas estructuras combinatorias para un muestreo del CRT con 4 puntos

Hagamos un ejemplo para $n = 4$; por lo dicho arriba se tiene que la estructura combinatoria de $\theta(e; U_1, U_2, U_3, U_4)$ estará dada por alguno de los árboles mostrados en la figura 5. Sin pérdida de generalidad digamos que es el árbol A, dotemos a este árbol de alturas con las etiquetas h_1, \dots, h_7 de abajo hacia arriba y de izquierda a derecha como en la figura 6.

Sean h'_1, \dots, h'_7 números mayores que 0 y H_1, \dots, H_7 las variables aleatorias que representan a las alturas del árbol $\theta(e; U_1, \dots, U_4)$.

Notemos que $\tau = \tau(e, U_1, \dots, U_4) = A$ y $(H_v)_{v \in \tau} \in \prod_{v \in \tau} [h_v + h'_v]$ sii las siguientes condiciones se dan en la excursión browniana:

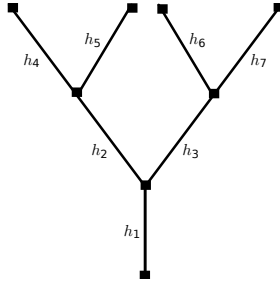


FIGURA 6. La manera en que se dota a un árbol con alturas

$$\begin{aligned}
 e_{U_{(1)}} &\in [h_1 + h_2 + h_4, h_1 + h_2 + h_4 + h'_4) \\
 m_1 &\in [h_1 + h_2, h_1 + h_2 + h'_2) \\
 e_{U_{(2)}} &\in [h_1 + h_2 + h_5, h_1 + h_2 + h_5 + h'_5) \\
 m_2 &\in [h_1, h_1 + h'_1) \\
 e_{U_{(3)}} &\in [h_1 + h_3 + h_6, h_1 + h_3 + h_6 + h'_6) \\
 m_3 &\in [h_1 + h_3, h_1 + h_3 + h'_3) \\
 e_{U_{(4)}} &\in [h_1 + h_3 + h_7, h_1 + h_3 + h_7 + h'_7)
 \end{aligned}$$

Estas condiciones se pueden ver explícitamente en la figura 7.

De lo expuesto anteriormente se puede llegar a la conclusión de que si $\Theta = (A, (h_v)_{v \in A})$ entonces podemos expresar la ley finito dimensional del CRT para el caso $n = 4$, a saber:

$$\begin{aligned}
 &\mathbb{P}(\theta(e_{U_1}, \dots, U_4) \in d\Theta) \\
 &= \lim_{h'_1, \dots, h'_7 \rightarrow 0} \frac{1}{h'_1 \dots h'_7} \mathbb{P}(\tau = A, (H_v)_{v \in A} \in \prod_{v \in \tau} [h_v, h_v + h'_v)) \\
 (16) \quad &= 2^5 4! (h_1 + \dots + h_7) \exp(-2(h_1 + \dots + h_7)^2).
 \end{aligned}$$

La última igualdad puede verse simplemente derivando las integrales que se obtendrían y después haciendo las sustituciones pertinentes en la densidad obtenida en el lema 2.3.

Notemos lo siguiente: la estructura combinatoria sólo influye en el orden en el que las alturas son colocadas en la excursión browniana; pero dado que la

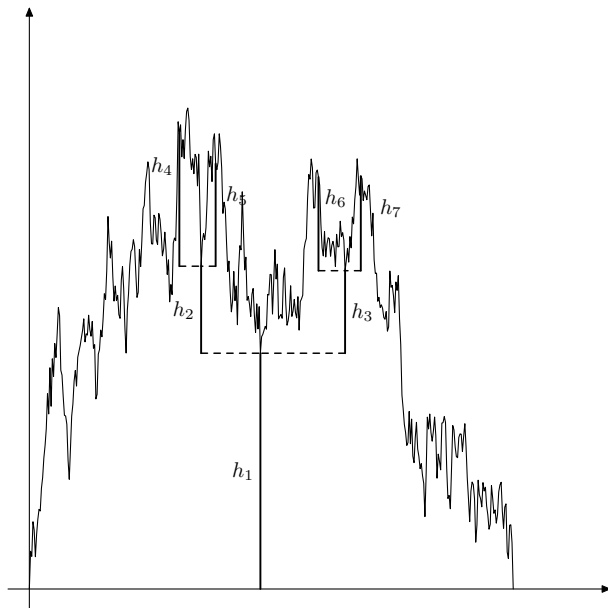


FIGURA 7. Ejemplo de la traducción de la distribución de alturas en la excursión browniana normalizada

ecuación 16 es una función simétrica de éstas, entonces la estructura combinatoria se vuelve irrelevante, lo que nos dice que esta ley induce una medida uniforme en el espacio \mathbf{A}_4^{bin} .

El resultado más general puede establecerse y demostrarse de la misma manera que el caso $n = 4$

Lema 3.4. *Las leyes finito dimensionales de \mathcal{T}_e están dadas por:*

$$\mathbb{P}(\theta(e; U_1, \dots, U_n) \in d\Theta) = 2^{n+1} n! L(\Theta) \exp(-2L(\Theta)^2)$$

Donde $\Theta = (\tau, (h_v)_{v \in \tau})$ es un árbol márcado y $L(\Theta)$ es la longitud del árbol

$$L(\Theta) := \sum_{v \in \tau} h_v$$

Lo último que realizaremos en este capítulo será demostrar que los árboles muestreados descritos anteriormente convergen casi seguramente al CRT.

Proposición 3.1. *La sucesión de árboles $\theta(e, U_{(1)}, \dots, U_{(n)}) \rightarrow \mathcal{T}_e$ conforme $n \rightarrow \infty$.*

DEMOSTRACIÓN. Para probar este hecho veremos que las funciones codificantes para los árboles $\theta(e, U_{(1)}, \dots, U_{(n)})$ convergen casi seguramente en la métrica uniforme, esto nos basta en vista del lema 3.2.

Sabemos que para una función continua f en $[0, 1]$ y para toda $\epsilon > 0$ existe $\delta > 0$, tal que para toda partición $\Delta = \{x_1, \dots, x_n\}$ del $[0, 1]$ con $\text{mesh}(\Delta) < \delta$ entonces la función g formada por trozos lineales que unen a los puntos $f(x_i), f(x_{i+1})$ cumple que $\|f - g\|_\infty < \epsilon$.

Con esto en mente definimos las variables X_k^n como sigue: dadas n variables $(U_i)_{i=1}^n$ uniformes $X_k^n = U_{(k+1)} - U_{(k)}$ con $U_0 = 0, U_{(n+1)} = 1$. Debido a lo dicho anteriormente nos basta probar que

$$Y_n = \max_k X_k^n \longrightarrow 0 \text{ c.s.}$$

conforme $n \longrightarrow \infty$. Ya que esto nos asegura que las funciones g_n formadas por trozos lineales que unen a los puntos $e_{U_{(i)}}, e_{U_{i+1}}$ convergen c.s. a e . Y por lo tanto las funciones codificantes de los árboles $\theta(e, U_{(1)}, \dots, U_{(n)})$ convergen c.s a e .

Para probar entonces lo que queremos acotamos

$$\begin{aligned} \mathbb{P}(Y_n > \rho) &= \mathbb{P}\left(\bigcup_{k=0}^n \{X_k^n > \rho\}\right) \\ &\leq \sum_{k=0}^n \mathbb{P}(X_k^n > \rho) \end{aligned}$$

Ahora bien, por el hecho de que estamos trabajando con variables uniformes e independientes se sigue que todas las variables X_i^n tienen la misma distribución marginal igual a la distribución del mínimo de las U_i' s por lo tanto

$$\sum_{k=0}^n \mathbb{P}(X_k^n > \rho) = (n+1)(1-\rho)^n$$

Con esta cota podemos ver que

$$\sum_{n=0}^{\infty} \mathbb{P}(Y_n > \rho) \leq \sum_{n=0}^{\infty} (n+1)(1-\rho)^n < \infty$$

para toda $0 < \rho \leq 1$. Por lo que el lema de Borel-Cantelli nos asegura

$$\mathbb{P}(\overline{\lim} Y_n > \rho) = 0$$

para toda $0 < \rho \leq 1$. Esto implica que $\lim Y_n = 0$ c.s. terminando así la demostración.

□

Construcciones alternas del CRT

Veremos en este capítulo la manera original en que se construyó el CRT, a saber, la publicada por David Aldous en [Ald91]. Esta construcción difiere de la construcción dada en el capítulo anterior en el hecho de que para ésta el CRT es un proceso límite, mientras que en el capítulo anterior construimos el CRT como un objeto, después lo discretizamos y demostramos que el CRT era el límite de estas discretizaciones.

Así mismo en este capítulo se dará una manera de codificar árboles discretos y un algoritmo para simular el árbol uniforme en n hojas; el cual resulta bastante fácil de implementar.

1. Creciendo el árbol binario uniforme en n hojas

Definición 4.1. *Sea τ un árbol ordenado finito y $\sigma, \sigma' \in \tau$; diremos que σ es hermano de σ' si $\pi(\sigma) = \pi(\sigma')$, i.e si tienen el mismo padre.*

Haremos un pequeño cambio en la definición de árbol finito ordenado agregándole un nuevo vértice al árbol que será el padre de la raíz. La raíz será denotada por 0 y el nuevo vértice por \emptyset . De esta forma $\pi(1) = \pi(2) = 0$ y $\pi(0) = \emptyset$, como se ve en la figura 1.

Si $\sigma = \pi(\sigma')$ definimos la arista entre estos dos vértices como el segmento $[[\sigma, \sigma']]$. En el dibujo ésta es el segmento que une a un padre con su hijo. Al conjunto de todas las aristas de un árbol τ lo denotaremos por E_τ . Denotaremos también al conjunto de todas las hojas de τ por L_τ .

Por último para un árbol τ y $\sigma \in \tau$ definimos a la decendencia de σ como sigue:

$$\tau[\sigma] = \{\sigma' : \pi^k(\sigma') = \sigma \text{ para alguna } k\}$$

Ahora que tenemos todos los elementos necesarios, daremos una transformación de A_n^{bin} en A_{n+1}^{bin} de tal forma que el árbol uniforme de n hojas crece en el árbol uniforme de $n + 1$ hojas.

Para n fijo construimos los conjuntos R, S de la siguiente manera:

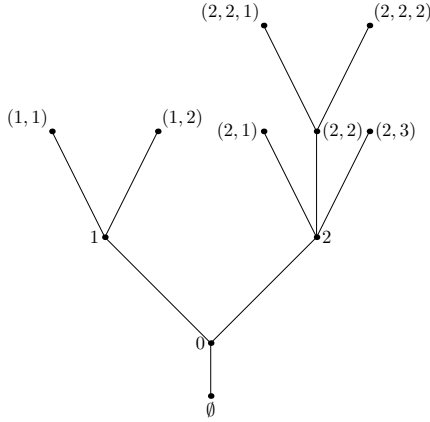


FIGURA 1

$$R \subset A_{n+1}^{bin} \times \bigcup_{\tau \in A_{n+1}^{bin}} L_{\tau}$$

y los elementos de R cumplen $(\tau, \sigma) \in R$ sii $\sigma \in L_{\tau}$. De una manera análoga definimos el conjunto

$$S \subset A_n^{bin} \times \bigcup_{\tau \in A_n^{bin}} E_{\tau} \times \{1, 2\}$$

diciendo que $(\tau, [[\sigma, \sigma']], x) \in S$ sii $[[\sigma, \sigma']] \in E_{\tau}$.

Ahora bien sea:

$$\phi : R \longrightarrow S$$

dada por la siguiente regla de correspondencia:

Sean $\tau \in A_{n+1}^{bin}$ y $\sigma \in \tau$ una hoja del árbol. Si $\sigma = (x_1, \dots, x_m, x_{m+1})$ entonces $\phi(\tau, \sigma) = (\theta_{\tau}, [[\pi^2(\sigma), \pi(\sigma)]], x_{m+1})$ donde θ_{τ} es el árbol de n hojas que se forma al quitar de τ la hoja σ y a su padre $\pi(\sigma)$ y a cualquier $\eta \in \tau[\pi(\sigma)]$ se le aplica la siguiente transformación

$$\eta = (x_1, \dots, x_m, y_1, \dots, y_k) \longrightarrow (x_1, \dots, x_m, y_2, \dots, y_k)$$

Es claro que la función está bien definida, es decir, que $\phi(R) \subset S$.

Veamos ahora que la función es biyectiva. Tomemos $(\tau, \sigma) \neq (\tau', \sigma')$, tenemos 2 casos:

Caso 1: $\sigma = (x_1, \dots, x_m) \neq (y_1, \dots, y_k) = \sigma'$

En este caso $(x_1, \dots, x_{m-1}) \neq (y_1, \dots, y_{k-1})$ o $x_m \neq y_k$, si pasa lo primero entonces

$$[[\pi^2(\sigma), \pi(\sigma)]] \neq [[\pi^2(\sigma'), \pi(\sigma')]]$$

Si pasa lo segundo entonces la tercera coordenada de $\phi(\tau, \sigma)$ es distinta de la de $\phi(\tau', \sigma')$. En ambos ocurre que $\phi(\tau, \sigma) \neq \phi(\tau', \sigma')$.

Caso 2: $\tau \neq \tau'$.

Podemos suponer que $\sigma = \sigma'$, de lo contrario regresamos al caso anterior. Sea $\sigma = (x_1, \dots, x_{m+1})$, nuevamente pueden ocurrir dos cosas:

$\tau \setminus \tau[\pi(\sigma)] \neq \tau' \setminus \tau'[\pi(\sigma')]$ o $\tau[\pi(\sigma)] \neq \tau'[\pi(\sigma')]$. Si pasa lo primero, ya que esa parte del árbol no se modifica para construir θ_τ y $\theta_{\tau'}$, ambos árboles también son distintos. Si pasa lo segundo entonces sabemos que existe $\eta \in \tau[\pi(\sigma)] \setminus \tau'[\pi(\sigma')]$ donde $\eta = (x_1, \dots, x_m, y_1, \dots, y_k)$. Supongamos que existe $\eta' = (x_1, \dots, x_m, y^*, y_2, \dots, y_k) \in \tau'$ con $y \neq y^*$.

Entonces ya que $\sigma \in \tau \cap \tau'$ y ambos son árboles binarios se sigue que $y_1 \neq x_{m+1}$ y $y^* \neq x_{m+1}$ pues σ es una hoja; por lo tanto $y_1 = y^*$, lo cual contradice lo que habíamos supuesto. Por lo tanto no puede existir η' con las características mencionadas. De este argumento se sigue que

$$\eta := (x_1, \dots, x_m, y_2, \dots, y_k) \in \theta_\tau \setminus \theta'_{\tau'}$$

Debido a los casos 1 y 2 podemos concluir que la función ϕ es inyectiva.

Probemos ahora que su imagen es S. Sea $\theta \in A_n^{bin}$, $[[\eta, \eta']] \in E_\theta$ y $x = 1, 2$. Construimos un árbol τ como sigue : η' y su hermano pertenecen a τ , además si $\eta' = (x_1, \dots, x_m)$ y $\eta = \pi(\eta')$ entonces para toda $\psi \notin \theta[\eta]$ se tiene que $\psi \in \tau$, por último agregamos a τ también el vértice (x_1, \dots, x_m, x) y si el vértice $(x_1, \dots, x_{m-1}, x_m, y_1, \dots, y_k) \in \theta$ entonces $(x_1, \dots, x_m, y, y_1, \dots, y_k)$ pertenecerá a τ , donde $y \in \{1, 2\}$ pero $y \neq x$. En el dibujo la construcción es más ilustrativa además puede verse que $\phi(\tau, (x_1, \dots, x_m, x)) = (\theta, [[\eta, \eta']], x)$ por lo tanto tenemos una biyección. Además en la demostración se dió una manera de crecer un árbol de n hojas en un árbol de $n+1$ hojas, a saber tomar una arista elegir un lado y agregar una hoja en esa arista. Demostraremos

ahora que aleatorizando esta forma podemos hacer crecer el árbol uniforme de n hojas.

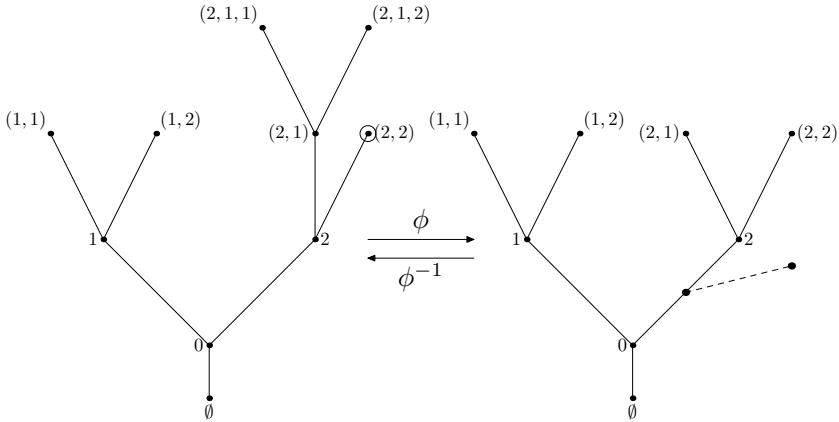


FIGURA 2

Sea T_n el árbol uniforme de n hojas, X una variable aleatoria uniforme independiente de T_n en $\{1, 2\}$ y Λ una variable aleatoria en $\bigcup_{\tau \in A_{n+1}^{bin}} E_\tau$ que cumple:

$$\mathbb{P}(T_n = \theta, \Lambda = [[\sigma, \sigma']], X = x) = \mathbb{P}(T_n = \theta) \frac{1}{2n - 1} \mathbb{P}(X = x)$$

Es decir una vez elegido un árbol, Λ denota elegir una arista del árbol uniformemente.

Proposición 4.1. *Sea $(T', \Sigma) = \phi^{-1}(T_n, \Lambda, X)$, entonces T' se distribuye uniformemente en A_{n+1}^{bin} .*

DEMOSTRACIÓN. Por ser ϕ biyectiva, es fácil ver que:

$$\mathbb{P}(T' = \tau, \Sigma = \sigma) = \frac{1}{2c_n(2n - 1)}$$

Tomando en cuenta que $(T', \Sigma) \in R$, se tiene que para cada τ fijo Σ sólo puede tomar $n + 1$ valores, a saber el número de hojas de τ ; por lo tanto integrando la distribución conjunta con respecto a σ obtenemos:

$$\mathbb{P}(T' = \tau) = \frac{n + 1}{2c_n(2n - 1)} = \frac{1}{c_{n+1}}$$

Por lo tanto $T' = T_{n+1}$ en distribución. □

Este resultado nos asegura que para construir el árbol uniforme en $n + 1$ hojas a partir del árbol uniforme de n hojas sólo necesitamos elegir una arista y un lado de ésta uniformemente y agregar ahí una hoja.

2. Codificación de los árboles finitos

En esta sección daremos una codificación de los árboles finitos ordenados, ésta será utilizada para poder simular el árbol uniforme de n hojas, ya que será más fácil implementar la simulación en una computadora siguiendo esta codificación.

Sea $\mathbf{t} = \{u_0 = \emptyset, u_1, u_2, \dots, u_n\}$ un árbol ordenado finito con sus vértices ordenados en orden lexicográfico. Definimos la sucesión (x_0, x_1, \dots, x_n) asociada a \mathbf{t} como sigue:

$$x_r = \mathbf{k}_{u_r}(\mathbf{t}) - 1$$

Notemos que en caso de un árbol binario se tiene una secuencia formada por 1's y -1 's y para el caso general la sucesión cumple que:

1. $x_k \geq -1$ para toda k
2. $x_n = -1$
3. $x_0 + \dots + x_m \geq 0 \ \forall m < n$
4. $x_0 + x_2 + \dots + x_n = -1$

La primera condición es obvia pues $\mathbf{k}_{u_r} \geq 0$ para toda r y la segunda también es fácil de ver ya que como los vértices están en orden lexicográfico es necesario que el último sea una hoja. Para la tercera y cuarta condición sea $V_m = \{u_0, \dots, u_m\}$; estimemos

$$\sum_{u \in V_n} \mathbf{k}_u(\mathbf{t})$$

Notemos que cada vértice u a excepción de la raíz es contado una sólo vez en la suma (pues lo cuenta $\mathbf{k}_{\pi(u)}(\mathbf{t})$); por lo tanto

$$\sum_{u \in V_n} \mathbf{k}_u(\mathbf{t}) = n$$

Y por lo tanto $\sum x_k = \sum_{u \in V_n} \mathbf{k}_u(\mathbf{t}) - n - 1 = -1$. Esto demuestra la condición 4. Por último para lo condición 3, como V_m tiene el orden lexicográfico se tiene que éste es un árbol por si sólo y además $\mathbf{k}_u(V_m) < \mathbf{k}_u(\mathbf{t})$ para algún $u \in V_m$ pues u_{m+1} tiene que ser hijo de algún elemento u de V_m y no está siendo contado en $\mathbf{k}_u(V_m)$ pero si en $\mathbf{k}_u(\mathbf{t})$, entonces se sigue que

$$\sum_{u \in V_m} \mathbf{k}_u(\mathbf{t}) > \sum_{u \in V_m} \mathbf{k}_u(V_m) = m$$

Y restando $m - 1$ en ambos lados de la desigualdad se obtiene la condición 3 de la lista.

Veamos ahora que a cada secuencia con las condiciones 1, 2, 3 y 4 le corresponde un árbol. Sea (x_0, \dots, x_n) una sucesión con las propiedades anteriores le asociamos el árbol $\mathbf{t} = \{u_0, \dots, u_n\}$ construido recursivamente como sigue: Si la secuencia consta de un sólo número $x = -1$ entonces el árbol asociado es \emptyset y es claro que la secuencia asociada al árbol es -1 . Ahora bien sea (x_0, \dots, x_n) una sucesión con las propiedades anteriores. Por la propiedad 3 se tiene que $x_0 \geq 0$ de donde $x_1 + \dots + x_n \leq -1$, si $x_1 = -1$ entonces x_1 cumple las 4 propiedades y podemos definir $\mathbf{t}_1 = \emptyset$ como el árbol que le corresponde a x_1 . Por otro lado si $x_1 > -1$ entonces existe r_1 el primer momento en que $x_1 + \dots + x_{r_1} = -1$, esto pasa ya que por la propiedad 1 los decrementos de la sucesión $s_k = x_1 + \dots + x_k$ son de tamaño 1 y como $s_1 > -1$ y $s_n \leq -1$, tiene que haber un momento r_1 en que $s_{r_1} = -1$. Así, es fácil ver que (x_1, \dots, x_{r_1}) es una sucesión con las 4 propiedades, por lo tanto podemos construir \mathbf{t}_1 recursivamente. Si $r_1 \neq n$ entonces nuevamente por la propiedad 3 tenemos que $x_0 + \dots + x_{r_1} \geq 0$ y por lo tanto $x_{r_1+1} + \dots + x_n \leq -1$. De la misma manera que antes podemos encontrar r_2 y construir \mathbf{t}_2 . En el momento en que $r_k = n$ habremos construido \mathbf{t}_i para $i = 1 \dots k$. Finalmente construyamos \mathbf{t} el árbol asociado a la sucesión como sigue:

$$\mathbf{t} = \{\emptyset\} \bigcup_{i=1}^k \{iu : u \in \mathbf{t}_i\}$$

Es decir pegamos los árboles \mathbf{t}_i a una raíz. Sólo nos resta probar que la sucesión asociada a este árbol es la sucesión dada. Por la construcción de \mathbf{t} es obvio que $\mathbf{k}_\emptyset(\mathbf{t}) - 1 = k - 1$ Por otro lado debido a que:

$$\begin{aligned} x_1 + \dots + x_{r_1} &= -1 \\ &\vdots \\ x_{r_{k-1}+1} + \dots + x_n &= -1 \end{aligned}$$

Sumando todas estas ecuaciones y agregandoles x_0 tenemos que: $\sum_{i=0}^n x_i = x_0 - k$ y de la condición 4 se sigue que $x_0 = k - 1$. Por la construcción recursiva podemos hacer lo mismo para cada vértice u_i de \mathbf{t} y obtener $x_i = \mathbf{k}_{u_i}(\mathbf{t}) - 1$, es decir que la sucesión dada es la asociada al árbol \mathbf{t} .

2.1. Algoritmo para crecer el árbol binario uniforme de n hojas.

Habiendo dado una manera de codificar los árboles, sólo necesitamos ver a que equivale crecer un árbol en el espacio de las sucesiones recién definidas.

Por definición de la sucesión a cada x_n se le asocia un vértice u_n y a cada uno de éstos puede asociársele la arista $[[u_n, \pi(u_n)]]$ (recuérdese que por lo visto en la sección 1 el vacío también tiene un padre) y esta asignación es biunívoca.

Ahora bien la transformación ϕ^{-1} elige un árbol de n hojas, un arista de éste y un lado de la arista, una vez elegidos, se agrega una nueva arista en el lado de la arista elegida. Sea (x_i) la sucesión asociada al árbol y x_k, u_k el número y el vértice respectivamente asociados a la arista elegida, si el lado elegido fue el izquierdo entonces notemos que estamos agregando dos vértices antes de u_k (con respecto al orden lexicográfico) a saber el nodo en la arista $[[u_k, \pi(u_k)]]$ que será el nuevo padre de u_k y la nueva hoja que por definición de ϕ^{-1} será hermano de la imagen de u_k pero su última coordenada será 1 por lo tanto está antes que la imagen de u_k con respecto al orden lexicográfico. Por último como el nodo tiene 2 hijos por ser un árbol binario entonces le corresponde el número 1 mientras que a la hoja agregada le corresponde el número -1. Por lo tanto el vector (x_i) se manda en $(x_0, x_1, \dots, 1, -1, x_k, \dots, x_{2n-1})$. Obteniendo así la codificación del árbol de $n+1$ hojas obtenido por la función ϕ^{-1} . De manera similar si elegimos el lado derecho entonces se agregan 2 nuevos vértices: el padre de la imagen de u_k y su hermano, sin embargo este último tendrá como última coordenada 2 y por lo tanto se tiene que agregar a la sucesión sólo hasta haber recorrido toda la progenie de u_k . Ahora bien como habíamos dicho antes si r es el primer número tal que

$$\sum_{m=k}^r x_m = -1$$

Entonces u_r es el último miembro de la progenie de u_k ; por lo tanto debo de agregar un '-1' en después del número x_r para poder codificar el árbol obtenido mediante la transformación. Por lo tanto (x_i) se transforma en

$$(x_0, \dots, 1, x_k, \dots, x_r, -1, \dots, x_{2n-1})$$

Ahora que sabemos cómo se ve nuestra transformación en el código, definimos una forma recursiva de obtener el árbol binario uniforme de n hojas. El árbol de una hoja es el (-1). El árbol de 2 hojas es el (1,-1,-1), es totalmente determinista pues sólo hay un árbol de 2 hojas.

Ahora supongamos que tenemos construido el árbol binario uniforme de n hojas, elegimos entonces el vector (x_i) con la distribución uniforme y uniformemente elegimos un valor de x_i (que corresponde a uniformemente elegir una arista del árbol), después elegimos uniformemente un número en $\{1, 2\}$ (que es elegir un lado de la arista), y finalmente si elegimos 1 modificamos el vector cómo se dijo anteriormente en el caso del lado izquierdo, si elegimos 2 entonces modificamos el vector cómo se dijo para el caso del lado derecho. En la siguiente figura se puede ver un ejemplo de un muestreo de 7 árboles de 5 hojas realizado con el algoritmo expuesto en esta sección.

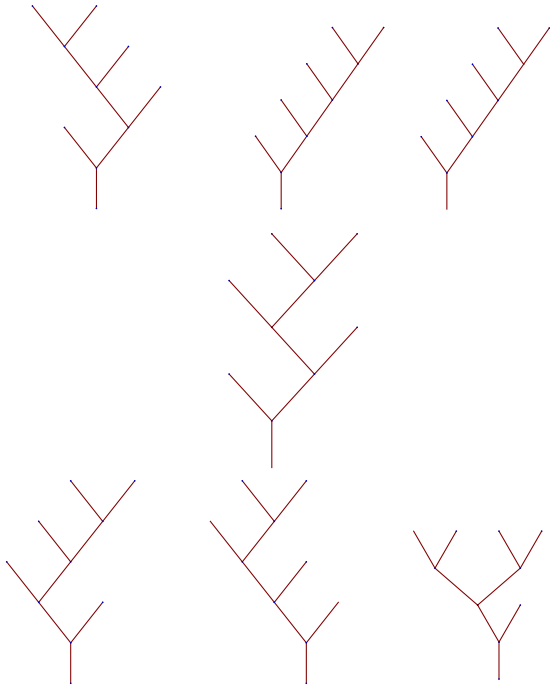


FIGURA 3. Muestreo de 7 árboles de 5 hojas distribuidos uniformemente

3. El árbol continuo de Aldous

Originalmente el CRT no fue construido como se explicó en el capítulo anterior, es decir, cómo un espacio cociente del intervalo $[0, 1]$; sino como un proceso límite en el espacio de sucesiones l_1 y la razón de esto fue dar una noción de convergencia en la métrica de Hausdorff inducida por la norma en l_1 , así como de convergencia débil en el espacio de medidas de l_1 ; sin embargo la visualización de este proceso no es tan clara cómo lo es usando la definición de la sección anterior.

Esta sección se dedicará a definir el CRT de la forma en que lo define Aldous [Ald91] y daremos un tipo de equivalencia con el CRT definido en el capítulo anterior.

Primero daremos una breve introducción de por que nos interesa trabajar con el espacio l_1 . Supongamos que tenemos un árbol discreto \mathcal{T} dotado de la métrica d inducida por la gráfica, y queremos dar una realización geométrica de éste, es decir, queremos tener una representación $r : \mathcal{T} \rightarrow X$ inyectiva, donde X es un espacio normado. Además pedimos la condición de que para cualesquiera dos puntos $x, y \in \mathcal{T}$ con realizaciones $r(x), r(y)$ y si $x = v_1, v_2, \dots, v_k = y$ son los vértices del único camino que une a x y a y entonces $d(x, y) = \sum_{i=1}^{k-1} \|v_{i+1} - v_i\|$, también para facilitar la visualización de estas realizaciones nos gustaría que las aristas fueran 'ortogonales' entre si, esto es sólo una manera de hablar, pues no necesariamente el espacio X tendrá que tener un producto interior; sin embargo esto no será un problema por que en l_1 tenemos una idea de 'ortogonalidad', a saber los vectores z_i los cuales tienen un 1 en su coordenada i -ésima y 0 en sus demás coordenadas, podemos pensarlos como 'ortogonales'.

Pero si pensamos en sumergir a \mathcal{T} en algún R^n las direcciones ortogonales 'se nos acaban' si \mathcal{T} tiene una cantidad infinita de ramas y por lo tanto l_1 resulta una buena opción. Ahora damos una manera de encajar un árbol en l_1 :

Etiquetamos los vértices de T con un conjunto $S = v, w, \dots \subset l_1$ de tal forma que la raíz sea 0 y $\|v - u\| = d(v, u)$ para ver que tal representación existe ordenemos arbitrariamente las áristas e_1, e_2, \dots y diremos que

$$v = \sum_{n=1}^{\infty} z_n J(e_n, v)$$

Donde $J(e_n, v) = 1$ si e_n pertenece al camino que va de v a la raíz. Por lo tanto $\|v - u\| = \sum_{n=0}^{\infty} |J(e_n, v) - J(e_n, u)| = d(v, u)$.

Definición 4.2. Sean (C_1, C_2, \dots) los tiempos de ocurrencia de eventos en un proceso Poisson con tasa $t dt$, definimos también $B_i = \xi_i C_i$ donde ξ_i se

distribuye uniforme en $[0, 1]$ y es independiente del proceso Poisson. Ahora definimos la función $\rho : \mathbb{R} \rightarrow l_1$ como $\rho(0) = 0$,

$$\rho(x) = \rho(B_i) + (x - C_i)z_i \text{ para } C_i < x \leq C_{i+1}$$

Sean:

$$\begin{aligned} \Gamma_t &= \rho([0, t]) \\ \Gamma_\infty &= \rho([0, \infty)) \\ \Gamma &= \overline{\Gamma_\infty} \end{aligned}$$

El CRT se define entonces como Γ .

Heurísticamente el proceso Γ_t definido en l_1 realiza lo siguiente: crecemos un tronco a velocidad unitaria iniciando en 0 y en dirección z_1 , después al tiempo C_1 elegimos uniformemente en el tronco un punto de ramificación y empezamos a crecer la rama en dirección z_2 a velocidad unitaria; al momento C_2 elegimos nuevamente un punto de ramificación pero en este caso será uniforme en todo el árbol y ahí crecemos una rama en dirección z_3 y así sucesivamente hasta llegar al tiempo t .

El hecho de que el CRT definido en esta sección y el definido en la sección anterior sean equivalentes claramente no es inmediato y de hecho en el sentido estricto no representan la misma medida en el conjunto de árboles reales cómo veremos más adelante, pues Γ será el árbol obtenido por $2\mathbf{e}$, es decir la función codificante para Γ es $2\mathbf{e}$; además las distribuciones finito dimensionales de Γ inducirán la medida uniforme en el espacio de árboles binarios etiquetados no ordenados mientras que las distribuciones finito dimensionales de \mathcal{T}_e indujeron la medida uniforme en los árboles binarios ordenados no etiquetados. A pesar de estas pequeñas diferencias veremos que tras hacer unos cuantos cambios podemos decir que de cierta forma estos 2 árboles aleatorios si son equivalentes.

Recordemos que el lema 3.4 del capítulo anterior nos dice que las distribuciones finito dimensionales de T_e están dadas por:

$$\mathbb{P}(\theta(e, U_1, \dots, U_n) \in d\tau) = 2^{n+1} n! L(\tau) \exp(-2L(\tau)^2)$$

con respecto a la medida uniforme en \mathbf{A}_p^{bin} , esto quiere decir que si integramos con respecto a las alturas obtenemos la medida uniforme en \mathbf{A}_p^{bin} , o dicho de otra manera, que las distribuciones finito dimensionales están determinadas por elegir una estructura combinatoria uniformemente y después dotarla de alturas. Supongamos ahora que el espacio de estructuras combinatorias no es

\mathbf{A}_p^{bin} sino el conjunto de árboles binarios ordenados y etiquetados, como a cada árbol ordenado le corresponden $n!$ maneras de etiquetarlo tendríamos que las distribuciones finito dimensionales de \mathcal{T}_e serían:

$$2^{n+1}L(\tau) \exp(-2L(\tau)^2)$$

Si olvidamos el orden, tomando en cuenta que hay 2^{n-1} árboles ordenados que al etiquetarlos representan el mismo árbol desordenado entonces las distribuciones se vuelven:

$$2^{2n}L(\tau) \exp(-2L(\tau)^2)$$

Si reemplazamos \mathbf{e} por $2\mathbf{e}$, es decir, multiplicamos las alturas aleatorias por 2, entonces resulta que las distribuciones finito dimensionales de \mathcal{T}_{2e} están dadas por:

$$(17) \quad L(\tau) \exp\left(\frac{-L(\tau)^2}{2}\right)$$

Con respecto a la medida uniforme en el conjunto de árboles etiquetados no ordenados.

Probemos ahora que las distribuciones finito dimensionales de Γ están dadas precisamente por 17.

Proposición 4.2. *Las distribuciones de los árboles Γ_{C_n} están dadas por la ecuación 17. Donde τ es cualquier árbol binario etiquetado no ordenado de n hojas.*

DEMOSTRACIÓN. Es intuitivamente claro que podemos codificar los árboles C_n mediante un árbol etiquetado discreto donde las etiquetas corresponden al orden en el cual se van agregando las hojas y un conjunto de alturas $(x_1, x_2, \dots, x_{2n-1})$, denotemos entonces por $f(\tau, x_1, \dots, x_{2n-1})$ a la función de densidad correspondiente a éste árbol, sea τ^* el árbol obtenido de τ al quitar la n -ésima hoja. Sabemos que el nodo de donde crece la n -ésima hoja se distribuye uniformemente en todo el árbol; por lo tanto el árbol de $n - 1$ hojas se codifica por: $\tau^* \text{ y } (x_{*1}, \dots, x_{2n-3*})$ donde cada x_{*i} es igual a una x_j y sólo para una se tiene que $x_{i*} = x_r + x_s$ y es la altura que representa la rama a la cual pertenece el nuevo nodo. Por lo tanto la densidad debe estar dada por:

$$f(\tau, x_1, \dots, x_{2n-1}) = \frac{1}{L(\tau^*)} L(\tau) \exp\left(-\frac{L(\tau)^2}{2} + \frac{L(\tau^*)^2}{2}\right) f(\tau^*, x_1^*, \dots, x_{2n-3}^*)$$

Ya que $\frac{1}{L(\tau^*)}$ es la densidad de elegir un punto uniformemente en el árbol de $n - 1$ hojas y $L(\tau) \exp\left(-\frac{L(\tau)^2}{2} + \frac{L(\tau^*)^2}{2}\right)$ es la densidad de la longitud de $C_n - C_{n-1}$ dado que $C_{n-1} = L(\tau^*)$. Además la independencia estocástica de $(\tau^*, x_1^*, \dots, x_{2n-3}^*)$ y lo dicho anteriormente se debe a que el proceso Poisson visto después de C_n es independiente de lo que ocurrió antes.

Por inducción y utilizando el hecho de que el árbol de una sola hoja de altura x tiene densidad $\exp\left(-\frac{x^2}{2}\right)$ llegamos a que la densidad está dada por:

$$f(\tau, x_1, \dots, x_{2n-1}) = L(\tau) \exp\left(-\frac{L(\tau)^2}{2}\right)$$

Que es precisamente la ecuación 17. □

Una vez probado esto sabemos por lo dicho en el capítulo anterior que en distribución

$$\lim_{n \rightarrow \infty} \Gamma_{C_n} = \mathcal{T}_{2e}$$

pero además es claro que por construcción $\lim_{n \rightarrow \infty} \Gamma_{C_n} = \Gamma$ por lo tanto $\mathcal{T}_{2e} = \Gamma$ en distribución.

Analicemos un poco más la densidad anterior. Primero notemos que la función en x_1, \dots, x_n es simétrica, por lo tanto las variables aleatorias que representan la longitud de las ramas son intercambiables; pero también la longitud de las ramas están representadas por $U_{(i)} - U_{(i-1)}$, donde $U_{(i)}$ es un punto aleatorio en el segmento $[0, C_n]$. Intuitivamente esperamos que los puntos $U_{(i)}$ sean los estadísticos de orden de variables aleatorias uniformes en $[0, C_n]$; de lo contrario las distancias no serían intercambiables pues los puntos se acumularían en donde la medida se acumula. Formalmente tenemos lo siguiente:

Proposición 4.3. *El árbol Γ_{C_n} está caracterizado por:*

1. *Una forma $S(n)$ elegida uniformemente en el conjunto de árboles etiquetados.*
2. *Un conjunto de alturas $H(n)$ intercambiables, donde las alturas están distribuidas como la longitud de los segmentos obtenidos al cortar el segmento C_n en $2n - 2$ puntos uniformes.*

3. $S(n)$ y $H(n)$ son independientes entre sí.

DEMOSTRACIÓN. Calculemos la densidad de C_n .

$$\begin{aligned} \mathbb{P}(C_n > t) &= \mathbb{P}(N_t < n) \\ &= \sum_{k=0}^{n-1} \exp\left(-\frac{t^2}{2}\right) \frac{(t^2/2)^k}{k!} \end{aligned}$$

Derivando esto obtenemos:

$$\begin{aligned} \mathbb{P}(C_n \in dt) &= \sum_{k=0}^{n-1} t \exp\left(-\frac{t^2}{2}\right) \frac{(t^2/2)^k}{k!} - \sum_{k=0}^{n-2} t \exp\left(-\frac{t^2}{2}\right) \frac{(t^2/2)^k}{k!} \\ &= t \exp\left(-\frac{t^2}{2}\right) \frac{(t^2/2)^{n-1}}{(n-1)!} \end{aligned}$$

Ahora sean U_1, \dots, U_{2n-2} variables uniformes en C_n , entonces:

$$\mathbb{P}(U_{(1)} \in dx_1, \dots, U_{(2n-2)} \in dx_{2n-2} | C_n = t) = \frac{(2n-2)!}{t^{2n-2}}$$

Por lo tanto la densidad conjunta está dada por:

$$\mathbb{P}(U_{(1)} \in dx_1, \dots, U_{(2n-2)} \in dx_{2n-2}, C_n \in dt) = \frac{(2n-2)!}{2^{n-1}(n-1)!} t \exp\left(-\frac{t^2}{2}\right).$$

Si hacemos un cambio de variable para calcular la densidad conjunta de las distancias de los segmentos se tiene la densidad:

$$\frac{(2n-2)!}{2^{n-1}(n-1)!} (x_1 + \dots + x_{2n-1}) \exp\left(-\frac{(x_1 + \dots + x_{2n-1})^2}{2}\right)$$

Si multiplicamos esta densidad por $\frac{2^{n-1}(n-1)!}{(2n-2)!}$ obtenemos precisamente la ecuación 17. Pero además sabemos que la cantidad de árboles binarios etiquetados no ordenados es precisamente $\frac{(2n-2)!}{2^{n-1}(n-1)!}$ por lo tanto formar a Γ_{C_n} equivale a lo dicho en la proposición y con esto termina la demostración. \square

Este resultado es particularmente útil para realizar simulaciones del CRT pues en la sección anterior dimos una forma de simular el árbol en n hojas uniforme y en esta sección dimos una manera de simular sus distancias por lo tanto combinando ambas y utilizando el resultado anterior que nos dice que las distancias y la forma son independientes podemos simular muestreos finitos del CRT.

Cuadrangulaciones

Este capítulo intentará dar una idea general de lo que son las cuadrangulaciones, su definición y algunas de sus propiedades. La necesidad de estudiar las cuadrangulaciones viene de la física donde se utilizan bajo el nombre de geometría euclidiana cuántica discretizada (!?). Más importante será la relación que hay entre las cuadrangulaciones y 2 clases de árboles: Árboles bien etiquetados y móviles bien etiquetados. La importancia de estas relaciones se encuentra en que actualmente no se sabe mucho sobre la convergencia de cuadrangulaciones aleatorias, mientras que en árboles se conoce el celebrado teorema de Aldous sobre límites de árboles al CRT [Ald93]. De esta manera actualmente se busca tratar de relacionar límites de cuadrangulaciones con lo que se podría llamar el mapa continuo aleatorio.

1. Mapas, cuadrangulaciones, árboles y móviles

Esta sección constará únicamente de las definiciones y ejemplos necesarios de los objetos que serán utilizados más adelante en este capítulo.

Un mapa es un encaje propio (sin autointersecciones) de una gráfica conexa en el plano. En principio un mapa puede tener aristas dobles y lazos. Nos interesaran en particular los mapas enraizados, es decir, cuando elegimos una arista direccionada en la cara infinita como raíz. Al vértice del cual parte la raíz se le conoce como vértice raíz. Dos mapas se consideran el mismo si existe un homeomorfismo del plano en si mismo que mande un mapa en el otro incluyendo las raíces.

La diferencia entre un mapa y una gráfica es que en el mapa nos interesa el orden cíclico de las aristas como se puede ver en la figura 1. Los mapas también pueden definirse en la esfera, si el encaje se pide en ésta en lugar de en el plano. Podemos utilizar entonces la fórmula de Euler: si f es el número de caras, n el número de aristas y v el número de vértices entonces $f + v = n + 2$.

Definición 5.1. Una cuadrangulación es un mapa en el cual todas sus caras son de orden 4, esto quiere decir que cada cara tiene 4 aristas incidentes.

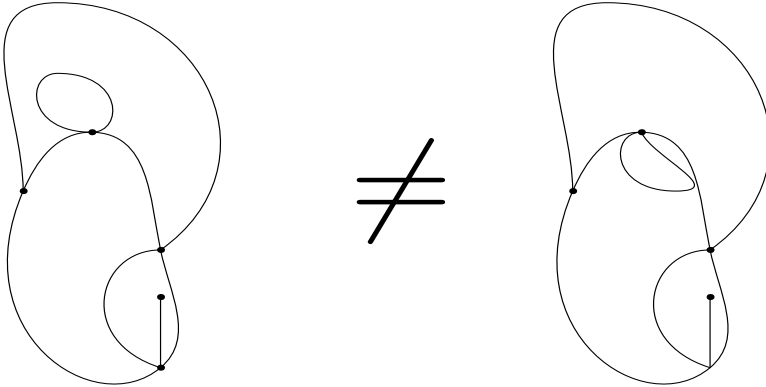


FIGURA 1.

Una cuadrangulación con n caras tendrá $2n$ aristas por la restricción del orden de las caras del mapa, y por la fórmula de Euler sabemos que tendrá $n + 2$ vértices. En la figura 2 pueden verse ejemplos de cuadrangulaciones en el plano y en la esfera. Todas las cuadrangulaciones son claramente bipartitas, es decir, podemos pintar los vértices de 2 colores (blanco y negro) de tal manera que cualesquiera dos vértices adyacentes estén pintados de distinto color.

Sea Q_n el conjunto de todas las cuadrangulaciones con n caras, un resultado muy conocido obtenido por *W.T. Tutte* es el siguiente:

$$|Q_n| = \frac{2}{n+2} \frac{3^n}{n+1} \binom{2n}{n}$$

Sean x, y dos vértices en un mapa, definimos su distancia $d(x, y)$ como el mínimo número de aristas que necesitan recorrerse para ir de x a y , visto de otra manera es como si todas las aristas tuvieran longitud 1. Una vez dotando de una métrica al conjunto de vértices del mapa podemos definir su perfil.

Definición 5.2. Sea M un mapa enraizado definimos el perfil del mapa como la sucesión $(H_k)_{k \geq 1}$. Donde H_k es el número de vértices que se encuentran a distancia k del vértice raíz v_0 .

Claramente el soporte del perfil es un intervalo, es decir, $\{k | H_k > 0\} = \{1, \dots, r\}$. A r se le conoce como el radio del mapa.

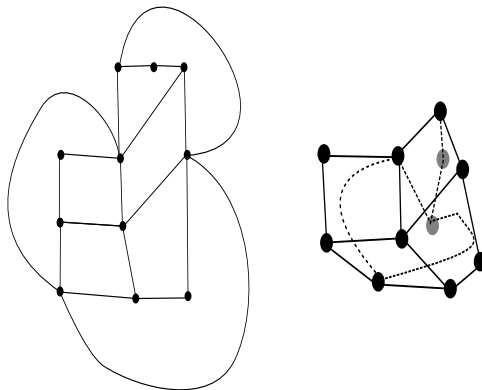


FIGURA 2.

Pasaremos ahora a definir el concepto de árbol plano, el cuál se puede pensar como una realización de un árbol real. Un árbol plano es por definición un mapa con una sola cara. Análogamente puede definirse un árbol plano recursivamente como sigue:

1. El árbol más pequeño es un punto.
2. Todos los otros árboles son una sucesión no vacía de subárboles pegados a una raíz.

En otras palabras, todos los vértices tienen una sucesión (posiblemente vacía) de hijos y todo vértice excepto por el vértice raíz tiene un padre.

Se dice que un árbol es bien etiquetado si todos los vértices del árbol son numerados con las dos condiciones siguientes: El vértice raíz tiene etiqueta 1 y las etiquetas de dos vértices adyacentes difieren en a lo más 1. En la figura 3 puede verse un ejemplo de un árbol bien etiquetado.

La distribución de etiquetas de un árbol bien etiquetado es una sucesión (λ_k) que denota el número de vértices del árbol etiquetados con el número k .

Para concluir esta sección definiremos lo que es un móvil, para en la siguiente sección ver como se relacionan todas estas definiciones.

Definición 5.3. *Un móvil etiquetado es un árbol plano que cumple las siguientes condiciones:*

1. *Sus vértices se dividen en dos clases: vértices no etiquetados y vértices etiquetados con números enteros positivos.*

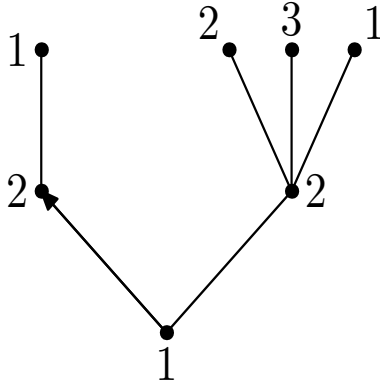


FIGURA 3.

2. Todas las aristas conectan un vértice etiquetado con un vértice no etiquetado.
3. Para todo vértice no etiquetado v las etiquetas m, n de dos vértices adyacentes a v y consecutivos en el sentido de las manecillas del reloj debe cumplirse que $m \geq n - 1$.

Si el móvil además cumple con la condición de que exista al menos un vértice con etiqueta 1, se dice que el móvil está bien etiquetado. En la figura 1 puede verse el ejemplo de un móvil.

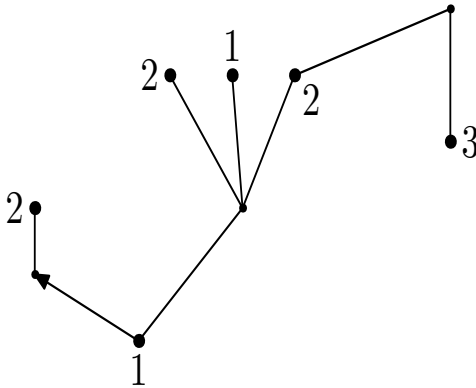


FIGURA 4

2. Una biyección entre cuadrangulaciones y árboles bien etiquetados

En esta sección daremos un método constructivo para crear árboles bien etiquetados a través de mapas y viceversa, estableciendo así una biyección entre ambos conjuntos; sin embargo no daremos la demostración de que las construcciones explicadas son en efecto una biyección pues la prueba es muy técnica y la notación un poco engorrosa, más bien el objetivo será dejar clara la relación entre los árboles y los mapas.

Sea Q una cuadrangulación enraizada y denotemos por v_0 a su vértice raíz. Etiquetamos los vértices de Q con la función ϕ definida como sigue $\phi(x) = d(x, v_0)$ donde $d(x, y)$ denota la distancia en Q . Notemos que $H_k = \{x | \phi(x) = k\}$. Esta manera de etiquetar los vértices tiene las siguientes sencillas y evidentes propiedades:

1. Si x, y están unidos por una arista entonces $|\phi(x) - \phi(y)| = 1$. Pues por ser un árbol bipartita $|\phi(x) - \phi(y)|$ es impar y sería una contradicción que la diferencia fuera mayor a 1 pues el camino de x a y sólo necesita de una arista.
2. Alrededor de una cara aparecen 2 vértices negros x_1, x_2 y 2 vértices blancos y_1, y_2 . Estos vértices satisfacen al menos una de las siguientes igualdades $\phi(x_1) = \phi(x_2)$ o $\phi(y_1) = \phi(y_2)$.

Una cara se llamará simple si sólo ocurre una de las igualdades anteriores y compuesta si ocurren ambas. Debe remarcarse que puede ocurrir que $x_1 = x_2$ o $y_1 = y_2$.

Ahora construiremos un árbol bien etiquetado \mathcal{T} a partir de una cuadrangulación Q . La construcción es como sigue: formamos un nuevo mapa Q' dividiendo cada cara compuesta en 2 caras triangulares uniendo los 2 vértices que están etiquetados con el mayor valor. Ahora construimos un subconjunto $\mathcal{T}(Q)$ de aristas de Q' con las siguientes reglas:

1. En cada cara compuesta de Q seleccionamos la arista que se agregó para formar Q' .
2. Para cada cara simple f de Q , sea v el vértice etiquetado con el mayor valor, entonces la arista que sale de v manteniendo a f a la izquierda se selecciona.

En principio esto es sólo un conjunto de vértices y aristas; la demostración de que es en efecto un árbol se puede ver en [CS04]. Por último elegimos la arista raíz del árbol tomando a la primera arista en el sentido opuesto a las manecillas del reloj que tiene como base al término de la raíz de Q y

1. Se coloca un vértice v_0 en la cara del árbol, se le da la etiqueta 0 y se une con una arista a las l esquinas que tienen etiqueta 1. Al terminar este paso se tendrá claramente un mapa enraizado \mathcal{T}_0 con l caras, donde la raíz del mapa es la arista que va de v_0 al vértice raíz v'_0 del árbol y que se encuentra antes de la raíz del árbol cuando se recorren las aristas incidentes a v'_0 en el sentido opuesto a las manecillas del reloj. El siguiente paso se hará en cada una de las caras así que tomaremos una cara genérica F de \mathcal{T}_0 . Sea k el grado de F por construcción $k \geq 3$; en F existe una sola esquina incidente a v_0 con etiqueta 0. Numeramos las esquinas de la cara de 1 a k en el sentido horario de tal forma que la esquina con etiqueta 0 le corresponda el número k . Más aún sea e_i la etiqueta de la esquina i (entonces $e_1 = e_{k-1} = 1$ y $e_k = 0$).
2. Definimos la función sucesor $s(i)$ para todas las esquinas $1, \dots, k-1$ como sigue : $s(i) := \inf\{j > i \mid e_j = e_i - 1\}$ es decir $s(i)$ es el vértice más próximo a i en el sentido horario, cuya etiqueta es 1 menor que e_i . Para toda esquina i tal que $s(i) \neq i+1$ trazamos una arista que una a i con $s(i)$. Una vez hecho esto se obtiene un mapa \mathcal{T}' .
3. Al final eliminamos de \mathcal{T}' todas las aristas de la forma (e, e) y obtenemos un mapa Q .

Esta construcción se ve ejemplificada en la figura 5. Las construcciones que se acaban de dar son inversas y esto se puede ver gráficamente en la figura.

3. Una biyección entre cuadrangulaciones y móviles

En esta sección daremos una construcción que relaciona mapas con la característica de que sus caras sean de orden par y los móviles que son una generalización de los árboles bien etiquetados. Como las cuadrangulaciones caen claramente en la categoría apenas descrita, entonces podemos aplicar esta construcción.

Nuevamente etiquetamos los vértices del mapa como en la sección anterior. De esta forma sabemos que las etiquetas entre cualesquiera 2 vértices adyacentes difiere en uno.

Comenzaremos la construcción de un móvil a partir de un mapa \mathcal{M} . En cada cara de \mathcal{M} realizamos la siguiente operación: colocamos un vértice dentro de la cara y conectamos este vértice con cada vértice de la cara que tenga la propiedad de que al recorrer la cara en el sentido de las manecillas del reloj, el siguiente vértice tenga etiqueta menor. Como se ve en la figura 3. Finalmente eliminamos todas las aristas originales del mapa y nos quedamos tan sólo con las que provienen de los vértices colocados en el centro de la cara.

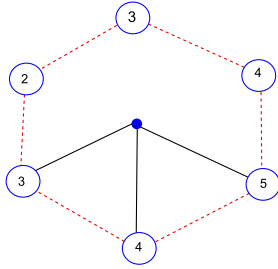


FIGURA 6. Se muestra la configuración de las etiquetas de una cara genérica de \mathcal{M} , así como la construcción descrita.

Para la construcción inversa se realiza algo parecido a la construcción de la sección anterior, etiquetamos cada esquina con el número de su vértice correspondiente y después unimos a cada esquina con su sucesor, donde el sucesor es la función definida en la sección pasada sólo que ahora utilizamos las esquinas de todo el árbol. Después de haber dibujado todas estas aristas agregamos un nuevo vértice con etiqueta 0 y unimos todos los vértices de etiqueta 1 con éste. Por último eliminamos todas las aristas que tienen un vértice no numerado como extremo. De esta forma obtenemos un mapa \mathcal{M} . Esta construcción es inversa a la anterior y la demostración de este hecho puede encontrarse en [BDFG04]

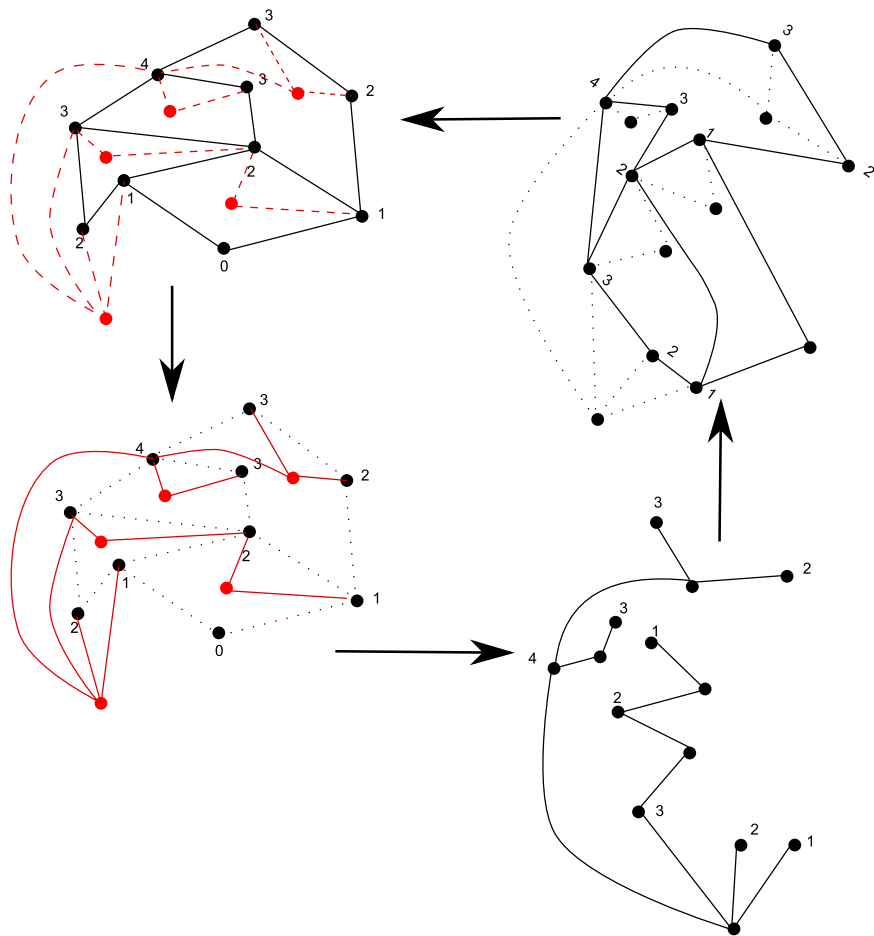


FIGURA 7. Esta figura muestra las 2 construcciones descritas en esta sección como ejemplo de que éstas son inversas entre si

Conclusiones

A lo largo de este trabajo hemos logrado reformular las demostraciones de varios resultados ya publicados y mostrar ciertas equivalencias, a mi parecer la más importante es la relación entre el CRT construido por Aldous y la construcción del CRT ofrecida por Le Gall. Así mismo creo que la tesis ofrece una explicación, que si no es más intuitiva al menos es técnicamente menos demandante, de lo que es el CRT, ¿a que me refiero con esto?: al hecho de entender por que las distribuciones finito dimensionales tienen esa expresión y de poder calcular estas distribuciones con elementos básicos de probabilidad. El objetivo de este trabajo era presentar de una manera más accesible conceptos como la excursión browniana normalizada, árboles reales y árboles aleatorios, es decir, sin elementos teóricos tan fuertes como la teoría de excursiones, la convergencia débil en l_1 , etc., y pienso que logró su objetivo.

¿Qué es lo que sigue? En el capítulo 4 de la tesis se encuentran distintas maneras de crecer árboles binarios aleatorios, ya sea cómo una simple estructura combinatoria o vistos cómo un espacio métrico, mientras que en el siguiente capítulo se vieron relaciones entre cuadrangulaciones y árboles, estas resultan muy útiles por que sabemos muy poco sobre que es lo que pasa con las cuadrangulaciones uniformes, entonces lo siguiente será tratar de iniciar un estudio sobre como crecer cuadrangulaciones uniformes por un lado y por otro tratar de estudiar los límites de éstas bajo reescalamientos. La conjetura es que si cierto tipo de árboles convergen a un árbol aleatorio continuo entonces sus cuadrangulaciones respectivas deberían de converger al objeto correspondiente al CRT, se ha demostrado que si las cuadrangulaciones en efecto convergen entonces deben de hacerlo a algo que es casi seguramente una esfera. Lo cual nos plantea la idea de que si logramos extender alguna de las biyecciones dadas en esta tesis al caso de un árbol infinito entonces al CRT le correspondería una esfera!

Bibliografía

- [Ald91] David Aldous, *The continuum random tree. I*, Ann. Probab. **19** (1991), no. 1, 1–28. MR MR1085326 (91i:60024)
- [Ald93] ———, *The continuum random tree. III*, Ann. Probab. **21** (1993), no. 1, 248–289. MR MR1207226 (94c:60015)
- [BDFG04] J. Bouttier, P. Di Francesco, and E. Guitter, *Planar maps as labeled mobiles*, Electron. J. Combin. **11** (2004), no. 1, Research Paper 69, 27 pp. (electronic). MR MR2097335 (2005i:05087)
- [Chu76] Kai Lai Chung, *Excursions in Brownian motion*, Ark. Mat. **14** (1976), no. 2, 155–177. MR MR0467948 (57 #7791)
- [CS04] Philippe Chassaing and Gilles Schaeffer, *Random planar lattices and integrated superBrownian excursion*, Probab. Theory Related Fields **128** (2004), no. 2, 161–212. MR MR2031225 (2004k:60016)
- [Duq08] Thomas Duquesne, *The coding of compact real trees by real valued functions*, <http://arxiv.org/abs/math/0604106>, 2008.
- [EPW06] Steven N. Evans, Jim Pitman, and Anita Winter, *Rayleigh processes, real trees, and root growth with re-grafting*, Probab. Theory Related Fields **134** (2006), no. 1, 81–126. MR MR2221786 (2007d:60003)
- [Gro99] Mikhail Gromov, *Metric structures for riemannian and non-riemannian spaces*, Birkhäuser, 1999.
- [LG05] Jean-François Le Gall, *Random trees and applications*, Probab. Surv. **2** (2005), 245–311 (electronic). MR MR2203728 (2007h:60078)
- [RY91] Daniel Revuz and Marc Yor, *Continuous martingales and brownian motion*, Springer, 1991.
- [Tud02] Constantine Tudor, *Procesos estocásticos*, Sociedad Matemática Mexicana, 2002.