



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

ÁRBOLES ALEATORIOS Y
PROPIEDADES FRACTALES DEL
CONTINUUM RANDOM TREE

T E S I S

QUE PARA OBTENER EL TÍTULO DE :

M A T E M Á T I C O

P R E S E N T A :

DANIEL ANTONIO MARQUEZ VAZQUEZ



DIRECTOR DE TESIS :
DR. JUAN CARLOS PARDO MILLAN

2013



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

1. Datos del alumno

Márquez
Vázquez
Daniel Antonio
Tel: 55 15 32 69 54
Universidad Nacional Autónoma de
México
Facultad de Ciencias
Matemáticas
306708301

2. Datos del tutor

Dr.
Juan Carlos
Pardo
Millán

3. Datos del sinodal 1

Dr.
Gerónimo Uribe
Uribe
Bravo

4. Datos del sinodal 2

Dr.
Arno
Siri
Jegousse

5. Datos del sinodal 3

Dr.
Víctor Manuel
Rivero
Mercado

6. Datos del sinodal 4

Dr.
Ricardo
Gómez
Aiza

7. Datos del trabajo escrito.

Árboles Aleatorios y Propiedades
Fractales del Continuum Random
Tree
179 p.
2013

*A mis padres,
María de Lourdes y Rafael.*

Mi más franco agradecimiento:

Para mi amigo, tutor y mentor

Juan Carlos,

por su total disposición y apoyo.

Mis hermanos,

Pablo y Fallo

por la motivación y experiencia que me proporcionaron

a lo largo de mi carrera.

A los miembros de mi honorable jurado:

Dr. Gerónimo Uribe Bravo

Dr. Arno Siri-Jegousse

Dr. Victor Manuel Rivero Mercado

Dr. Ricardo Gómez Aiza

Por sus comentarios, sugerencias

y el apoyo brindado.

A todas las personas

que me han respaldado, motivado y/o alentado,

ya que gracias a esas personas, ha sido posible

finalizar el presente trabajo.

*“La tormenta ya ha pasado,
revolvió mis sentimientos ...”*

ÁRBOLES ALEATORIOS Y PROPIEDADES
FRACTALES DEL CONTINUUM RANDOM TREE

Daniel Antonio Márquez Vázquez

Abril 2013

Índice general

Prefacio	v
1. Proceso de Ramificación de Galton-Watson	1
1.1. Hipótesis	1
1.2. Descripción matemática	2
1.3. Los Momentos de Z_n	5
1.4. La Probabilidad de Extinción	7
1.5. Ejemplos	14
2. Convergencia Débil en Espacios Métricos	17
2.1. Convergencia Débil en Espacios Métricos	17
2.2. El Teorema de Pormanteau	20
2.3. Otros criterios de convergencia débil.	26
2.4. Espacio Producto	28
2.5. Convergencia en Distribución	30
2.6. Convergencia en Probabilidad	32
2.7. Funciones Continuas	34

2.8. El Teorema de Prohorov	36
3. El Espacio $C([0, 1], \mathbb{R})$	43
3.1. Convergencia Débil y Tensión en C	45
3.2. Funciones Aleatorias	48
3.3. Medida de Wiener y Movimiento Browniano	52
3.4. El Teorema de Donsker	62
4. Árboles Aleatorios	67
4.1. Árboles Discretos	67
4.2. Árboles Galton-Watson	74
4.3. Convergencia al Movimiento Browniano	80
4.4. Algunas Aplicaciones	89
4.4.1. Tiempo Local y Excursión Browniana	89
4.4.2. Convergencia a la Excursión Browniana	91
4.4.3. Convergencia a un Proceso de Difusión de Feller	97
4.5. Árboles Galton-Watson con Progenie Fija	98
4.6. Convergencia de Funciones Contorno	102
5. Árboles Reales y su Codificación por Excursiones Brownianas.	107
5.1. Árboles Reales	107
5.2. Codificando Árboles Reales	111
5.3. El Continuum Random Tree (CRT)	124
5.4. La Medida de Excursión de Itô.	126

5.5. Marginales Finito Dimensionales bajo la medida de Itô	132
5.6. Marginales Finito Dimensionales del CRT	141
6. Medida de Hausdorff del CRT	147
6.1. Introducción	147
6.2. Medidas del Continuum Random Tree	150
6.3. Resultados de Comparación para Medidas de Hausdorff	151
6.4. Prueba de Teoremas	155

Prefacio

Los primeros capítulos, a riesgo de ser obsoletos para el lector experimentado en los temas, tienen la intención de dar una introducción rápida y lo más general posible a conceptos básicos de la teoría de procesos, probabilidad y el espacio de funciones continuas. Más adelante, discutimos varias conexiones entre cierto tipo de árboles aleatorios discretos denominados árboles Galton-Watson y un árbol aleatorio continuo denominado el Continuum Random Tree (CRT); en el ámbito discreto nos enfocamos a estudiar estos árboles condicionados a ser grandes respecto de alguna de las propiedades que los caracterizan y que a partir del Capítulo 5 se definen, como: su altura o el número de vértices que lo conforman. Introducimos dos funciones que codifican y caracterizan a un árbol discreto, la función de altura y de contorno. Se enuncian teoremas que nos dan la convergencia en distribución del proceso de altura (contorno) de una sucesión de árboles Galton-Watson hacia un movimiento browniano reflejado, donde dicho proceso es concatenación de la función altura (contorno) de cada uno de los árboles de la sucesión, cabe destacar que la prueba presentada es parcial para el caso de distribuciones con momentos exponenciales finitos; posteriormente se utilizan estos resultados para obtener nuevamente la convergencia en distribución del proceso de altura, solo que ahora de árboles Galton-Watson condicionados, hacia una excursión browniana. Discutimos brevemente ejemplos de aplicaciones para cierto tipo de árboles combinatorios. En la parte continua, utilizamos el formalismo de objetos llamados árboles reales, los cuales viven en un hiperespacio métrico; esto da como resultado una elegante formulación de la convergencia de árboles discretos reescalados hacia objetos continuos en términos de la métrica de Gromov-Hausdorff. Se explica como codificar árboles reales a través de funciones continuas, y como es que esto es el análogo continuo al proceso de codificación de árboles discretos por su función de contorno. Ponemos particular atención a la codificación de árboles reales por excursiones brownianas, y se calculan lo que llamamos distribuciones marginales del CRT haciendo uso de resultados de la teoría de excursiones. Enunciamos varias propiedades de la medida

de Itô de excursiones positivas del movimiento browniano. Posteriormente, damos un recorrido por propiedades básicas del CRT como la de ramificación. Dando ciertos argumentos y referencias, asumimos la existencia casi segura de una medida uniforme y local sobre el CRT y lo que denominaremos como un conjunto de nivel del CRT; finalmente se obtiene que estas coinciden con dos medidas de Hausdorff las cuales permiten obtener el tamaño fractal respectivamente del árbol y su conjunto de nivel. Esto último desafortunadamente ya no fue posible desarrollarlo en el presente trabajo.

Capítulo 1

Proceso de Ramificación de Galton-Watson

Los procesos de ramificación permiten modelar la evolución de poblaciones tales como células, neutrones en un reactor o una epidemia en una población. En particular para este modelo, es necesario imaginar que tenemos inicialmente un conjunto de individuos que puedan generar nuevos individuos del mismo tipo como bacterias reproduciéndose o los neutrones en una reacción en cadena. Comenzamos con un conjunto inicial de individuos, los cuales conforman la generación cero, si estos individuos tienen descendientes, estos nuevos integrantes conformarán a la primera generación y los descendientes de esta primera generación conformarán la segunda generación y así sucesivamente. Se seguirá con atención el tamaño de las generaciones sucesivas, no el tiempo al cual cada individuo nace; así como tampoco la relación familiar de cada individuo.

1.1. Hipótesis

Sean Z_0, Z_1, Z_2, \dots , variables aleatorias que denotan el número de individuos de la generación cero, primera generación, \dots y consideremos las siguientes dos hipótesis:

- Si el tamaño de la n -ésima generación es conocido entonces la ley de probabilidades que gobierna a las generaciones siguientes, no depende del tamaño de generaciones previas a la n -ésima; el proceso Z es una cadena de Markov.

- La cadena de Markov Z a considerarse, tiene una propiedad especial, la cual consiste en que diferentes individuos no interfieren unos con otros; esto es que el número de descendientes de un individuo no depende de cuantos otros individuos están presentes.

1.2. Descripción matemática

Sean $Z_0, \xi_{k,n}$ para $k, n \geq 1$ variables aleatorias donde las $\xi_{k,n}$ son independientes e idénticamente distribuidas y las cuales representan la población inicial y el número de descendientes que puede tener el k -ésimo individuo de la n -ésima generación respectivamente. Si ξ es una variable con la misma distribución que las variables $\xi_{k,n}$ para $k, n \geq 1$, tendremos que:

$$\mathbb{P}(\xi = k) = p_k \quad \text{para toda } k, n = 0, 1, \dots, \quad \text{y} \quad \sum_{k=0}^{\infty} p_k = 1, \quad (1.1)$$

donde p_k es la probabilidad de que un individuo tenga k hijos. La distribución de ξ es llamada la “distribución de hijos” o de “progenie” del proceso.

El *proceso de ramificación de Galton-Watson* $Z = (Z_n, n \geq 0)$ que comienza con N individuos, se define recursivamente por

$$Z_0 = M, \quad Z_{n+1} = \sum_{i=1}^{Z_n} X_{i,n} \quad \text{y donde} \quad Z_n = 0 \Rightarrow Z_{n+1} = 0.$$

En lo sucesivo supondremos que el proceso inicia con un solo individuo ($Z_0 = 1$). En conclusión, nuestra interpretación será la siguiente:

1. La variable aleatoria Z_n es el número de individuos en la n -ésima generación de Z .
2. Cualquier individuo k de la n -ésima generación da nacimiento independientemente de cualquier otro individuo a r individuos de la $(n+1)$ -ésima generación, $r = 0, 1, \dots$, inmediatamente después el individuo k muere.
3. El número de hijos de un individuo es una variable aleatoria con la misma distribución que en (1.1).

Es posible obtener probabilidades de transición para este proceso que denotaremos por P_{ij} y que representan la probabilidad de que en la $(n+1)$ -ésima generación haya j individuos dado que en la n -ésima generación la población total era de i individuos. La probabilidad de transición para toda $j, i, n \in \{0, 1, 2, 3, \dots\}$ esta dada por:

$$P_{ij} = \mathbb{P}(Z_{n+1} = j | Z_n = i) = \mathbb{P}\left\{ \sum_{k=0}^i \xi_{k,n} = j \right\}$$

Recordemos que $\xi \stackrel{(d)}{=} \xi_{k,n}$ para toda $k, n \geq 1$.

Definimos la función generadora para la variable ξ por:

$$f(s) = \mathbb{E}[s^\xi] = \sum_{k=0}^{\infty} p_k s^k \quad \text{donde } s \in [-1, 1].$$

Para $n \geq 1$, la función generadora para la variable Z_n la definimos por:

$$f_n(s) = \mathbb{E}[s^{Z_n}] = \sum_{k=0}^{\infty} \mathbb{P}\{Z_n = k\} s^k.$$

Como se tiene que $\mathbb{P}\{Z_0 = 1\} = 1$, tendremos que

$$f_0(s) = \sum_{k=0}^{\infty} \mathbb{P}\{Z_0 = k\} s^k = s.$$

Observemos que el número de individuos de nuestra primera generación, es exactamente el número de hijos del único individuo con el cual se comienza el proceso. Concluimos que la variable aleatoria Z_1 tiene la misma distribución que la variable aleatoria ξ , por lo que

$$f_1(s) = \sum_{k=0}^{\infty} \mathbb{P}\{Z_1 = k\} s^k = f(s).$$

Teorema 1. (Funciones Generadoras)

Sean X_1, X_2, X_3, \dots una sucesión de variables aleatorias independientes e idénticamente distribuidas con valores en los enteros y función generadora en común f , y sea N una variable aleatoria independiente de X_k para $k \geq 1$ con valores en los enteros positivos y función generadora f_N , entonces la variable aleatoria $Y = X_1 + X_2 + \dots + X_N$ tiene función generadora dada por

$$f_Y(s) = f_N(f(s)).$$

Demostración. Utilizando esperanza condicional obtenemos que,

$$\begin{aligned}
 f_Y(s) &= \mathbb{E}[s^Y] = \mathbb{E}[\mathbb{E}[s^Y | N]] = \sum_n \mathbb{E}[s^Y | N = n] \mathbb{P}(N = n) \\
 &= \sum_n \mathbb{E}[s^{X_1 + \dots + X_n}] \mathbb{P}(N = n) \\
 &= \sum_n \mathbb{E}[s^{X_1}] \dots \mathbb{E}[s^{X_n}] \mathbb{P}(N = n) \quad \text{por independencia} \\
 &= \sum_n [f(s)]^n \mathbb{P}(N = n) = f_N(f(s)).
 \end{aligned}$$

□

Recordemos que $\xi_{k,n}$ son v.a.i.i.d. para $k, n \geq 1$ y además

$$Z_n = \sum_{k=1}^{Z_{n-1}} \xi_{k,n-1}.$$

Es posible entonces aplicar el teorema anterior para obtener la función generadora del proceso Galton-Watson en la n -ésima generación, obteniendo así que

$$f_n(s) = f_{n-1}(f_{\xi_{k,n}}(s)) = f_{n-1}(f(s)). \quad (1.2)$$

Iterando esta función obtenemos lo siguiente:

$$\begin{aligned}
 f_{n+1}(s) = f_n(f(s)) &= f_{n-1}(f(f(s))) \\
 &= f_{n-1}(f_2(s)) = f_{n-2}(f(f_2(s))) = f_{n-2}(f_3(s)) = \dots
 \end{aligned}$$

Siguiendo inductivamente este procedimiento obtenemos que

$$f_{n+1}(s) = f_{n-k}(f_{k+1}(s)) \quad \text{para } k \in \{0, 1, 2, \dots, n\}.$$

En particular para $k = n - 1$ tendremos que

$$f_{n+1}(s) = f_1(f_n(s)) = f(f_n(s)), \quad (1.3)$$

y por lo tanto para toda $n \geq 1$

$$f_n(f(s)) = f(f_n(s)).$$

1.3. Los Momentos de Z_n

Utilizaremos ahora los resultados de la sección anterior. Recordemos que $Z_0 = 1$ y $\xi \stackrel{(d)}{=} \xi_{k,n}$ para $k, n \geq 1$; tenemos como consecuencia el siguiente lema.

Lema 2. Sean $\mu = \mathbb{E}[\xi]$, $\sigma^2 = \text{Var}[\xi]$ y suponemos que son finitas. Entonces

$$\mathbb{E}[Z_n] = \mu^n, \quad n \geq 1$$

y

$$\text{Var}[Z_n] = \begin{cases} n\sigma^2 & \text{si } \mu = 1 \\ \sigma^2 \mu^{n-1} \frac{\mu^n - 1}{\mu - 1} & \text{si } \mu \neq 1 \end{cases}$$

Demostración. Observemos que

$$f'(1) = \sum_{k=0}^{\infty} p_k k = \mu \quad \text{y} \quad f''(1) + \mu - \mu^2 = \sum_{k=0}^{\infty} p_k (k^2 - k) + \mu - \mu^2 = \sigma^2.$$

Derivando la expresión en (1.2) obtenemos que

$$f'_{n+1}(s) = f'_n(f(s))f'(s),$$

y evaluando en $s = 1$ llegamos a que

$$f'_{n+1}(1) = f'_n(1)f'(1).$$

Iterando esta última relación observamos que

$$f'_{n+1}(1) = f'_n(1)f'(1) = f'_{n-1}(1)[f'(1)]^2 = f'_{n-2}(1)[f'(1)]^3 = \dots$$

y así hasta llegar a que

$$f'_{n+1}(1) = f'_1(1)[f'(1)]^n = [f'(1)]^{n+1}.$$

Finalmente, por la observación hecha al inicio de la demostración se tiene que

$$\mathbb{E}[Z_{n+1}] = f'_{n+1}(1) = [f'_1(1)]^{n+1} = \mu^{n+1}$$

Ahora bien, para obtener la $\text{Var}[Z_{n+1}]$ observemos primero que

$$f''_n(1) = \sum_{k=2}^{\infty} k(k-1)\mathbb{P}(Z_n = k) = \mathbb{E}[Z_n^2] - \mathbb{E}[Z_n] = \mathbb{E}[Z_n^2] - f'_n(1)$$

de manera que

$$\text{Var}[Z_n] = f_n''(1) + f_n'(1) - (f_n'(1))^2.$$

Si ahora derivamos dos veces (1.3), obtenemos la siguiente relación:

$$f_{n+1}''(s) = f''(f_n(s)) [f_n'(s)]^2 + f'(f_n(s)) f_n''(s).$$

Evalutando en $s = 1$ obtenemos

$$f_{n+1}''(1) = f''(1) [f_n'(1)]^2 + f'(1) f_n''(1).$$

Sustituyendo observemos que

$$\begin{aligned} f_{n+1}''(1) &= (\sigma^2 + \mu^2 - \mu) \mu^{2n} + \mu f_n''(1) \\ &= (\sigma^2 + \mu^2 - \mu) \mu^{2n} + \mu [(\sigma^2 + \mu^2 + \mu) \mu^{2(n-1)} + \mu f_{n-1}''(1)] \\ &= (\sigma^2 + \mu^2 - \mu) (\mu^{2n} + \mu^{2n-1}) + \mu^2 f_{n-1}''(1) \\ &= (\sigma^2 + \mu^2 - \mu) (\mu^{2n} + \mu^{2n-1}) + \mu^2 [(\sigma^2 + \mu^2 - \mu) \mu^{2(n-2)} + \mu f_{n-2}''(1)] \\ &= (\sigma^2 + \mu^2 - \mu) (\mu^{2n} + \mu^{2n-1} + \mu^{2n-2}) + \mu^3 f_{n-2}''(1) = \dots, \end{aligned}$$

por lo que siguiendo inductivamente la iteración llegamos a que

$$f_{n+1}''(1) = (\sigma^2 + \mu^2 - \mu) (\mu^{2n} + \mu^{2n-1} + \mu^{2n-2} + \dots + \mu^n).$$

Utilizando el desarrollo anterior vamos a obtener la varianza de Z_{n+1} . Notemos que

$$\begin{aligned} \text{Var}[Z_{n+1}] &= (\sigma^2 + \mu^2 - \mu) (\mu^{2n} + \mu^{2n-1} + \mu^{2n-2} + \dots + \mu^n) + \mu^{n+1} - \mu^{2n+2} \\ &= \sigma^2 (\mu^{2n} + \mu^{2n-1} + \mu^{2n-2} + \dots + \mu^n) \\ &\quad + (\mu^{2n+2} + \mu^{2n+1} + \mu^{2n} + \dots + \mu^{n+2}) \\ &\quad - (\mu^{2n+1} + \mu^{2n} + \mu^{2n-1} + \dots + \mu^{n+1}) + \mu^{2n+1} - \mu^{2n+2} \\ &= \sigma^2 (\mu^{2n} + \mu^{2n-1} + \mu^{2n-2} + \dots + \mu^n) \\ &= \sigma^2 \mu^n (\mu^n + \mu^{n-1} + \mu^{n-2} + \dots + 1). \end{aligned}$$

De esto último se deduce fácilmente que

$$\text{Var}[Z_{n+1}] = \begin{cases} (n+1) \sigma^2 & \text{si } \mu = 1 \\ \sigma^2 \mu^n \frac{\mu^{n+1} - 1}{\mu - 1} & \text{si } \mu \neq 1. \end{cases}$$

Con esto finalizamos la prueba. □

OBSERVACION. La varianza crece o decrece geoméricamente según si $\mu > 1$ ó $\mu < 1$, y linealmente si $\mu = 1$.

1.4. La Probabilidad de Extinción

Ahora veremos la probabilidad de extinción del proceso Galton-Watson.

Definición 3. (Extinción del Proceso)

Por extinción vamos a entender como el evento en que la sucesión $(Z_n, n \geq 0)$ de variables aleatorias consiste en ceros para toda $n \geq n^*$ para algún n^* finito.

Como Z_n es un valor entero, la extinción es el evento tal que

$$\lim_{n \rightarrow \infty} Z_n = 0$$

Además, decimos que el proceso de Galton-Watson se extingue en el momento n si

$$Z_{n-1} > 0 \quad \text{y} \quad Z_n = 0$$

Definición 4. (Probabilidad de Extinción)

Tomaremos a η como la probabilidad de extinción de una familia, esto es,

$$\eta = \mathbb{P} \left\{ \bigcup_{n=0}^{\infty} \{\omega | Z_n(\omega) = 0\} \right\}$$

Teorema 5. *Tenemos que*

$$\lim_{n \rightarrow \infty} \mathbb{P}\{Z_n = 0\} = \eta = \mathbb{P}\{\text{extinción en un momento finito}\}$$

donde η es la raíz no negativa más pequeña de la ecuación $s = f_1(s)$. Además

$$\begin{array}{llll} \eta = 1 & \text{si} & \mu < 1 & \text{Caso Subcrítico} \\ \eta = 1 & \text{si} & \mu = 1 & \text{Caso Crítico} \\ \eta < 1 & \text{si} & \mu > 1 & \text{Caso Supercrítico} \end{array}$$

siempre y cuando $\mathbb{P}\{Z_1 = 1\} < 1$.

Demostración. Definimos los conjuntos A y A_n como

$$A = \bigcup_{n=1}^{\infty} A_n \quad \text{donde} \quad A_n = \{\omega | Z_n(\omega) = 0\}$$

y a la probabilidad del evento A_n como $\eta_n = \mathbb{P}\{A_n\}$. Es claro que dicha sucesión de conjuntos es creciente, es decir, $A_n \subseteq A_{n+1}$ para toda $n \geq 1$, por ser el 0 un estado absorbente del proceso. Se tiene entonces que

$$\lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n = A$$

y en consecuencia

$$\eta = \mathbb{P}\{A\} = \mathbb{P}\left\{\bigcup_{n=1}^{\infty} A_n\right\} = \mathbb{P}\left\{\lim_{n \rightarrow \infty} A_n\right\} = \lim_{n \rightarrow \infty} \mathbb{P}\{A_n\} = \lim_{n \rightarrow \infty} \eta_n.$$

Notando que $f_n(0) = \mathbb{P}\{Z_n = 0\} = \eta_n$ y tomando límites obtenemos que

$$\lim_{n \rightarrow \infty} f_n(0) = \lim_{n \rightarrow \infty} \mathbb{P}\{Z_n = 0\} = \lim_{n \rightarrow \infty} \eta_n.$$

Recordemos que $f_n(s) = f(f_{n-1}(s))$, donde f es la función generadora de Z_1 . Haciendo $s = 0$ se tiene que

$$\eta_n = f_n(0) = f(f_{n-1}(0)) = f(\eta_{n-1}) \tag{1.4}$$

y sabemos que $\eta = \lim_{n \rightarrow \infty} \eta_n$, por lo cual

$$\begin{aligned} \eta &= \lim_{n \rightarrow \infty} \eta_n = \lim_{n \rightarrow \infty} f(\eta_{n-1}) = \lim_{n \rightarrow \infty} \mathbb{E}[(\eta_{n-1})^{Z_1}] \\ &= \mathbb{E}\left[\lim_{n \rightarrow \infty} (\eta_{n-1})^{Z_1}\right] = \mathbb{E}[\eta^{Z_1}] = f(\eta). \end{aligned}$$

donde en la cuarta igualdad se utilizó el Teorema de la Convergencia Dominada.

Ahora consideramos la ecuación $s = f(s)$. Mostraremos que si e es cualquier raíz no negativa de la ecuación dada entonces $\eta \leq e$. De (1.4) y ya que f es una función no decreciente en $[0, 1]$ tenemos que

$$\begin{aligned} \eta_1 &= f(0) \leq f(e) = e, \\ \eta_2 &= f(\eta_1) \leq f(e) = e, \dots \end{aligned}$$

de manera que continuando con esta iteración observamos que para toda $n \in \mathbb{N}$, $\eta_n \leq e$, por lo que tomando límites concluimos que $\eta \leq e$. Entonces η es la raíz más pequeña no negativa de la ecuación $s - f(s) = 0$.

Recordemos que $\mathbb{P}\{Z_1 = 1\} < 1$. Observemos que la función f es estrictamente convexa sobre el intervalo $[0, 1]$, para esto necesitamos mostrar que su primera derivada es monótona no decreciente sobre dicho intervalo, para lo que basta ver que su segunda derivada es no negativa, lo cuál es fácil de verificar ya que

$$f''(s) = \mathbb{E}[Z_1(Z_1 - 1)s^{(Z_1-2)}] = \sum_{k=0}^{\infty} p_k k(k-1)s^{k-2} > 0 \quad \text{si } s \geq 0.$$

Por lo tanto f es estrictamente convexa y creciente en $[0, 1]$ con $f(1) = 1$, además es posible verificar que las curvas $y = f(s)$ y $y = s$ generalmente se intersectan un par de veces en dicho intervalo y esto ocurre cuando $s = \eta$ y $s = 1$.

Finalmente, si $\mu \leq 1$, probaremos que la ecuación $f(s) = s$ no tiene raíces en $[0, 1)$ y si $\mu > 1$, probaremos que la ecuación $f(s) - s = 0$ tiene una única raíz en $[0, 1)$.

Notemos que si $f'(s) < 1$ para toda $s \in [0, 1)$; en consecuencia

$$\frac{d}{ds}(f(s) - s) < 0 \quad \text{para } 0 \leq s < 1.$$

Esto nos dice que $f(s) - s$ es una función estrictamente decreciente en $[0, 1)$ y como $f(1) - 1 = 0$ entonces $f(s) - s > 0$ para toda $0 \leq s < 1$, concluyendo de esto último que la ecuación $f(s) - s = 0$ no tiene raíces en el intervalo $[0, 1)$.

Veamos entonces que efectivamente $f'(s) < 1$ para toda $s \in [0, 1)$ cuando $\mu \leq 1$.

Caso 1. Si $\mu < 1$ podemos afirmar que $f'(1) = \mu < 1$; como $f'(s)$ es no decreciente sobre $[0, 1]$, tendremos que $f'(s) < 1$ para toda $s \in [0, 1)$.

Caso 2. Si $\mu = 1$ entonces $\mathbb{P}\{Z_1 = 1\} < 1$ implica que $\mathbb{P}\{Z_1 = n\} > 0$ para algún $n \geq 2$. Como $f'(1) = \mu = 1$, ocurre que $f'(s) < 1$ para toda $s \in [0, 1)$, razón de esto es que $\lim_{s \uparrow 1} f'(s) = 1$ y f' es estrictamente creciente.

Con esto hemos mostrado que si $\mu < 1$ la única raíz de la ecuación $f(s) - s = 0$ es cuando $\eta = 1$. En el caso $\mu = 1$, es necesario distinguir entre los casos en que $\sigma^2 = 0$ y $\sigma^2 > 0$. El primer caso obliga a que $\mathbb{P}\{Z_1 = 1\} = 1$, por lo que $f(s) = s$ y en consecuencia $\eta = 0$, y en el segundo $f(s) > s$ para $s \in [0, 1)$ y por lo tanto $\eta = 1$.

Caso 3. Si $\mu > 1$ tendremos que $\lim_{s \uparrow 1} f'(s) = f'(1) > 1$. Luego, por continuidad de $f'(s)$ existe un $s_0 \in (0, 1)$ tal que para toda $s \in (s_0, 1)$, $f'(s) > 1$. Por el Teorema del Valor Medio para derivadas se tiene que

$$f'(\xi) = \frac{f(1) - f(s_0)}{1 - s_0} > 1 \quad \text{para alguna } \xi \in (s_0, 1)$$

y como $f(1) = 1$ entonces $f(s_0) - s_0 < 0$.

Ahora bien, $f(s) - s$ es continua en $s \in [0, s_0]$ y no negativa en $s = 0$, entonces por el Teorema del Valor Intermedio tendremos un cero de la función en $[0, s_0)$. Por lo tanto afirmamos que tenemos una raíz en η .

Supongamos que existe otra raíz $\eta_1 \in (0, 1)$, entonces $f(s) - s$ se anula en η , η_1 , y 1, y tendríamos que $0 \leq \eta < \eta_1 < 1$ pues ya vimos que η es la raíz más pequeña; luego, como f es estrictamente convexa, se tiene que

$$f(\eta(1-t) + t) < (1-t)f(\eta) + tf(1) \quad \text{para toda } t \in (0, 1),$$

y debido a que los intervalos son convexos, tenemos que para todo $s \in (\eta, 1)$, s es combinación convexa de η y 1, es decir $s = \eta(1-t) + t$ para alguna $t \in (0, 1)$, por lo que

$$f(s) = f(\eta(1-t) + t) < (1-t)f(\eta) + tf(1) = (1-t)\eta + t = s \quad \text{para toda } s \in (\eta, 1),$$

esto es $f(s) - s < 0$ para toda $s \in (\eta, 1)$; dicho de otra manera, $f(s) - s \neq 0$ para toda $s \in (\eta, 1)$, lo cual es una contradicción ya que $\eta_1 \in (\eta, 1)$ y es raíz. Concluimos que existe una única raíz que es menor a uno. \square

Con la teoría general es posible resumir la evolución a través del tiempo de un proceso de ramificación. Una de las dificultades, es que el proceso puede comportarse relativamente diferente dependiendo del momento de la extinción.

Una manera de abordar este problema es analizando el comportamiento del proceso condicionado a la ocurrencia de algún evento como la extinción o el valor de una variable aleatoria como el número total de progenie. Sean $g(k)$ y $f_1(s)$ la función de densidad y la función generadora de Z_1 respectivamente, es decir :

$$g(k) = \mathbb{P}\{Z_1 = k\} \quad f_1(s) = \mathbb{E}[s^{Z_1}].$$

Sea T el tiempo en que ocurre la extinción, definido como

$$T = \begin{cases} \inf \{n : Z_n = 0\} & \text{si la extinción ocurre} \\ \infty & \text{en otro caso} \end{cases}$$

Hablando de manera estricta si $T = \infty$ entonces el proceso crecerá más allá de toda cota, en cambio si $T < \infty$ se tendrá que el tamaño del proceso nunca llegará a ser muy grande y por lo tanto se irá a cero. Como ya vimos $\mathbb{P}\{T < \infty\}$ es la más pequeña raíz no negativa de la ecuación $s = f_1(s)$.

Definimos $E_n = \{n < T < \infty\}$ el evento: la extinción ocurre en algún momento posterior a n .

Debemos estudiar la distribución de Z_n condicionado a la ocurrencia de E_n , para lo cual definiremos una probabilidad de transición $P_{0j}^n = \mathbb{P}(Z_n = j | E_n)$, la cual es la probabilidad condicional de que en la n -ésima generación haya j individuos dada la futura extinción del proceso. Estamos interesados en encontrar el límite de dicha probabilidad, esto es:

$$\pi_j^0 = \lim_{n \rightarrow \infty} P_{0j}^n$$

si es que dicho límite existe.

En lo que prosigue y para evadir algunos casos triviales, supondremos que

$$0 < g(0) + g(1) < 1 \quad \text{y} \quad g(0) > 0,$$

estas condiciones implican que $0 < \mathbb{P}(E_n) < 1$ y $0 < \eta = \mathbb{P}\{T < \infty\} \leq 1$.

Teorema 6. Si $\mathbb{E}[Z_1] < \infty$ entonces $\lim_{n \rightarrow \infty} P_{0j}^n = \pi_j^0$ existe.

La función generadora $f^\pi(s) = \sum_{j=0}^{\infty} \pi_j^0 s^j$ satisface la ecuación

$$f^\pi(\eta^{-1} f_1(s\eta)) = \lambda f^\pi(s) + 1 - \lambda \quad (1.5)$$

donde η es la probabilidad de la extinción final y $\lambda = f'(\eta)$.

Notando que si $\mathbb{E}[Z_1] \leq 1$ entonces $\eta = 1$ y $\mu = \lambda$; tendremos que la ecuación (1.5) se reduce a

$$f^\pi(f_1(s)) = \mu f^\pi(s) + 1 - \mu.$$

Cualquiera que sea el valor de μ tenemos que $f_1'(\eta) \leq 1$ dándose la igualdad si y sólo si $\mu = 1$.

Demostración. Para $s \in [0, 1)$ sea

$$f_n^\pi(s) = \mathbb{E}[s^{Z_1} | E_n]$$

$$f_n^\pi(s) = \sum_{j=0}^{\infty} s^j P_{0j}^n = \sum_{j=0}^{\infty} s^j \frac{\mathbb{P}\{Z_n = j, E_n\}}{\mathbb{P}(E_n)}$$

dónde

$$\mathbb{P}(E_n) = \mathbb{P}\{n < T < \infty\} = \mathbb{P}\{T < \infty\} - \mathbb{P}\{T \leq n\} = \eta - f_n(0)$$

y

$$\begin{aligned}
\mathbb{P}\{Z_n = j, E_n\} &= \mathbb{P}\{Z_n = j, \text{todas las líneas siguientes mueren}\} \\
&= \mathbb{P}\{n < T < \infty | Z_n = j\} \mathbb{P}\{Z_n = j\} \\
&= \mathbb{P}\{T < \infty | Z_0 = j\} \mathbb{P}\{Z_n = j\} \\
&= \eta^j \mathbb{P}\{Z_n = j\} \quad j \geq 1.
\end{aligned}$$

Por otro lado, tenemos que $\mathbb{P}\{Z_n = 0, E_n\} = \mathbb{P}\{Z_n = 0\}$, por lo que

$$\begin{aligned}
f_n^\pi(s) &= \sum_{j=0}^{\infty} s^j \frac{\mathbb{P}\{Z_n = j, E_n\}}{\mathbb{P}(E_n)} \\
&= \frac{1}{\mathbb{P}(E_n)} \sum_{j=0}^{\infty} s^j \mathbb{P}\{Z_n = j\} \eta^j \\
&= \frac{1}{\eta - f_n(0)} \sum_{j=0}^{\infty} (s\eta)^j \mathbb{P}\{Z_n = j\} \\
&= \frac{f_n(s\eta) - f_n(0)}{\eta - f_n(0)}.
\end{aligned}$$

A continuación definiremos dos funciones que nos auxiliarán en la demostración del teorema de la siguiente manera:

$$H_n(s) = \frac{\eta - f_n(s)}{\eta - f_n(0)} \quad \text{y} \quad h(s) = \frac{\eta - f_1(s)}{\eta - s} \quad \text{donde } 0 \leq s < \eta$$

De esta manera es posible reescribir a $f_n^\pi(s)$ como:

$$f_n^\pi(s) = 1 - H_n(s\eta) \tag{1.6}$$

Nótese que H_n tiene dominio en el intervalo $[0, \eta)$ y $f_n^\pi(s)$ en el intervalo $[0, 1)$. Ahora proponemos la siguiente relación

$$\frac{H_n(s)}{H_{n-1}(s)} = \frac{h(f_{n-1}(s))}{h(f_{n-1}(0))}$$

esto ya que

$$\frac{H_n(s)}{H_{n-1}(s)} = \frac{\frac{\eta - f_n(s)}{\eta - f_n(0)}}{\frac{\eta - f_{n-1}(s)}{\eta - f_{n-1}(0)}} = \frac{(\eta - f_n(s))(\eta - f_{n-1}(0))}{(\eta - f_{n-1}(s))(\eta - f_n(0))}.$$

Por otro lado

$$\begin{aligned} \frac{h(f_{n-1}(s))}{h(f_{n-1}(0))} &= \frac{\frac{\eta - f_1(f_{n-1}(s))}{\eta - f_{n-1}(s)}}{\frac{\eta - f_1(f_{n-1}(0))}{\eta - f_{n-1}(0)}} = \frac{(\eta - f_1(f_{n-1}(s)))(\eta - f_{n-1}(0))}{(\eta - f_{n-1}(s))(\eta - f_1(f_{n-1}(0)))} \\ &= \frac{(\eta - f_n(s))(\eta - f_{n-1}(0))}{(\eta - f_{n-1}(s))(\eta - f_n(0))}. \end{aligned}$$

En cualquier caso $f_{n-1}(s)$ es no decreciente y $h(s)$ es no decreciente ya que f es convexa en $[0, \eta]$.

Por lo tanto tenemos que para $s < \eta$

$$H_n(s) \geq H_{n-1}(s)$$

entonces por la ecuación (1.6) tenemos que

$$f^\pi(s) = 1 - H(s\eta) \quad 0 \leq s < 1$$

si se satisfacen los siguientes límites para $s \in [0, 1]$

$$\lim_{n \rightarrow \infty} H_n(s\eta) = H(s\eta) \quad \lim_{n \rightarrow \infty} f_n^\pi(s) = f^\pi(s)$$

pero claramente esto ocurre.

Entonces el coeficiente π_j^0 de s^j en $f^\pi(s)$ existe para toda j como se deseaba. Aún más, si $0 \leq s < \eta$ tenemos

$$H_n(f(s)) = \frac{\eta - f_n(f_1(s))}{\eta - f_n(0)} = \frac{\eta - f_1(f_n(0))}{\eta - f_n(0)} \frac{\eta - f_{n+1}(s)}{\eta - f_{n+1}(0)} = h(f_n(0))H_{n+1}(s) \quad (1.7)$$

Luego dejando que $n \rightarrow \infty$ tenemos que $f_n(0) \rightarrow \eta$ y entonces podemos decir que

$$\lim_{n \rightarrow \infty} h(f_n(0)) = \lim_{s \rightarrow \eta} \frac{\eta - f_1(s)}{\eta - s} = f'_1(\eta)$$

entonces, cuando $n \rightarrow \infty$ en (1.7) obtenemos

$$H(f(s)) = f'_1(\eta)H(s) \quad \text{para } 0 \leq s < \eta \quad \text{y} \quad f^\pi(s) = 1 - H(s\eta).$$

Finalmente, con esto último podemos apreciar que

$$\begin{aligned} f^\pi(\eta^{-1}f_1(s\eta)) &= 1 - H(\eta^{-1}\eta f_1(s\eta)) = 1 - H(f_1(s\eta)) \\ &= 1 - f'_1(\eta)H(s\eta) = 1 - \lambda H(s\eta) \\ &= 1 - \lambda(1 - f^\pi(s)) = 1 - \lambda + f^\pi(s)\lambda. \end{aligned}$$

Con esto concluimos la prueba. □

1.5. Ejemplos

Ejemplo 1 (Cálculo de Esperanzas)

Sea un proceso de ramificación Z con $\mathbb{E}[Z_1] = \mu < \infty$. Para calcular $\mathbb{E}[Z_m Z_n]$ para $m \leq n$ primero tenemos que mostrar que $\mathbb{E}[Z_n | Z_m] = Z_m \cdot \mu^{n-m}$, lo cual se hará por inducción. Para demostrar que es válido para $n = m + 1$ primero veremos que

$$\mathbb{E}[Z_{m+1} | Z_m = j] = \mathbb{E}\left[\sum_{i=1}^j \xi_i \middle| Z_m = j\right] = \sum_{i=1}^j \mathbb{E}[\xi_i] = \sum_{i=1}^j \mu = j\mu$$

De esta manera tenemos que $\mathbb{E}[Z_{m+1} | Z_m] = \mu Z_m$. Ahora supondremos que es válido para un $n > m + 1$ y lo probaremos para $n + 1$.

$$\mathbb{E}[Z_{n+1} | Z_m] = \mathbb{E}[\mathbb{E}[Z_{n+1} | Z_n] | Z_m] = \mathbb{E}[\mu Z_n | Z_m] = \mu Z_m \mu^{n-m} = \mu^{n+1-m} Z_m$$

entonces, con este resultado es posible calcular lo siguiente

$$\mathbb{E}[Z_n Z_m | Z_m] = Z_m \mathbb{E}[Z_n | Z_m] = Z_m^2 \mu^{n-m}$$

obteniendo de esta manera que

$$\mathbb{E}[Z_n Z_m] = \mathbb{E}[\mathbb{E}[Z_n Z_m | Z_m]] = \mathbb{E}[Z_m^2 \mu^{n-m}] = \mu^{n-m} \mathbb{E}[Z_m^2].$$

Ejemplo 2 (Probabilidad de un ancestro común).

Consideremos un proceso de ramificación Z en el cuál $Z_0 = 1$ y $\mathbb{P}\{Z_1 = 0\} = 0$. Tomamos dos individuos aleatoriamente (con reemplazo) de la n -ésima generación y denotamos por L al índice de la generación que contiene su más reciente ancestro común. Queremos probar que

$$\mathbb{P}\{L = r\} = \mathbb{E}[Z_r^{-1}] - \mathbb{E}[Z_{r+1}^{-1}] \quad \text{para } 0 \leq r < n.$$

Supongamos $0 \leq r \leq n$, y que todo es conocido en el proceso hasta la generación r . Condicionando con esta información, un individuo escogido aleatoriamente de la n -ésima generación tiene probabilidad $1/Z_r$ de tener como ancestro de la r -ésima generación a cualquier miembro de la r -ésima generación. La probabilidad de que dos individuos de la n -ésima generación, tomados aleatoriamente e independientemente uno del otro, tengan el mismo ancestro en la r -ésima generación es entonces $1/Z_r$.

Luego,

$$\begin{aligned}
\mathbb{P}\{L < r\} &= \sum_{i=1}^{\infty} \mathbb{P}\{L < r, Z_r = i\} \\
&= \sum_{i=1}^{\infty} \mathbb{P}\{L < r | Z_r = i\} \mathbb{P}\{Z_r = i\} \\
&= \mathbb{E}[\mathbb{P}\{L < r | Z_r\}] = \mathbb{E}[1 - Z_r^{-1}]
\end{aligned}$$

de esta manera para $0 \leq r < n$

$$\begin{aligned}
\mathbb{P}\{L = r\} &= \mathbb{P}\{L < r + 1\} - \mathbb{P}\{L < r\} \\
&= \mathbb{E}[1 - Z_{r+1}^{-1}] - \mathbb{E}[1 - Z_r^{-1}] \\
&= \mathbb{E}[Z_r^{-1}] - \mathbb{E}[Z_{r+1}^{-1}].
\end{aligned}$$

Si ahora tenemos que $0 < \mathbb{P}\{Z_1 = 0\} < 1$ entonces

$$\begin{aligned}
\mathbb{P}\{L < r | Z_n > 0\} &= \sum_{i=1}^{\infty} \mathbb{P}\{L < r, Z_r = i | Z_n > 0\} \\
&= \sum_{i=1}^{\infty} \mathbb{P}\{L < r | Z_n > 0, Z_r = i\} \mathbb{P}\{Z_r = i | Z_n > 0\} \\
&= \mathbb{E}[\mathbb{P}\{L < r | Z_n > 0, Z_r\} | Z_n > 0] \\
&= \mathbb{E}[1 - Z_r^{-1} | Z_n > 0] = 1 - \mathbb{E}[Z_r^{-1} | Z_n > 0]
\end{aligned}$$

con este resultado obtenemos que

$$\mathbb{P}\{L = r | Z_n > 0\} = \mathbb{E}[Z_r^{-1} | Z_n > 0] - \mathbb{E}[Z_{r+1}^{-1} | Z_n > 0].$$

Ejemplo 3 (Ramificación con Migración).

Cada generación de un proceso de ramificación (con un solo progenitor), aumenta por un número aleatorio de inmigrantes los cuales son indistinguibles de los otros miembros de la población. Supongamos que el número de inmigrantes en diferentes generaciones son independientes del número de descendientes naturales de cada uno de los individuos en la historia inmediata del proceso, dado tal número de individuos en la generación previa, el número de inmigrantes tendrá función generadora $H(s)$. Sea Z_n el número de miembros de la n -ésima generación. La $n + 1$ -ésima generación tiene tamaño $C_{n+1} + I_{n+1}$ donde C_{n+1} es el número de descendientes naturales de la

generación previa e I_{n+1} es el número de inmigrantes.

Entonces por independencia tenemos

$$\begin{aligned}\mathbb{E} [s^{Z_{n+1}} | Z_n] &= \mathbb{E} [s^{C_{n+1} + I_{n+1}} | Z_n] \\ &= \mathbb{E} [s^{C_{n+1}} | Z_n] \mathbb{E} [s^{I_{n+1}} | Z_n] \\ &= \mathbb{E} [s^{C_{n+1}} | Z_n] H(s) = f(s)^{Z_n} H(s)\end{aligned}$$

y de esta manera obtenemos que $f_{n+1}(s)$ queda determinada por

$$f_{n+1}(s) = \mathbb{E} [s^{Z_{n+1}}] = \mathbb{E} [\mathbb{E} [s^{Z_{n+1}} | Z_n]] = \mathbb{E} [f(s)^{Z_n} H(s)] = f_n(f(s)) H(s).$$

Ejemplo 4 (Martingala).

Consideramos un proceso de ramificación $\{Z_n\}_{n=0}^{\infty}$ con $\mathbb{E}[Z_1] = \mu$. Sabemos que

$$\mathbb{E}[Z_{n+r} | Z_n] = Z_n \mu^r$$

donde $r, n \in \mathbb{N}$. Si definimos a $W_n = Z_n / \mu^n$ se puede ver que $\{W_n\}_{n=0}^{\infty}$ es una martingala, verificando que

$$\mathbb{E}[W_{n+r} | W_n] = \frac{1}{\mu^{n+r}} \mathbb{E}[Z_{n+r} | Z_n] = \frac{1}{\mu^{n+r}} Z_n \mu^r = W_n.$$

Capítulo 2

Convergencia Débil en Espacios Métricos

2.1. Convergencia Débil en Espacios Métricos

Sea (E, d) un espacio métrico y \mathcal{B}_E su σ -álgebra de Borel generada por los conjuntos abiertos.

Definición 7. (Convergencia Débil)

Dada una colección de medidas de probabilidad \mathbb{P}_n, \mathbb{P} en (E, \mathcal{B}_E) tal que satisfacen

$$\lim_{n \rightarrow \infty} \int_E f \, d\mathbb{P}_n = \int_E f \, d\mathbb{P}$$

para toda función real acotada y continua f , decimos que la sucesión \mathbb{P}_n converge débilmente a \mathbb{P} y escribimos

$$\mathbb{P}_n \xrightarrow{w} \mathbb{P}.$$

Con el siguiente teorema mostramos que toda medida de probabilidad queda determinada por los valores de $\mathbb{P}(F)$ para todo F cerrado.

Teorema 8.

Toda medida de probabilidad sobre un espacio medible (E, \mathcal{B}_E) es regular, es decir, si $A \in \mathcal{B}_E$ y $\epsilon > 0$, entonces existe un conjunto cerrado F y un abierto G tal que $F \subseteq A \subseteq G$ y $\mathbb{P}(G - F) < \epsilon$.

Demostración. Sea d la métrica en E y denotamos por $d(x, A)$ a la distancia de x al conjunto A . Si A es cerrado, hacemos $F = A$ y definimos a los conjuntos abiertos $G_\delta = \{x : \rho(x, A) < \delta\}$ con $\delta > 0$. Es claro que $G_\delta \downarrow F$ cuando $\delta \downarrow 0$, por lo que se tiene que para toda $\epsilon > 0$ existe una $\delta^* > 0$, tal que para toda $0 < \delta < \delta^*$,

$$\mathbb{P}(G_\delta - F) = \mathbb{P}(G_\delta) - \mathbb{P}(F) < \epsilon,$$

por lo que claramente la medida de probabilidad cumple la propiedad deseada sobre la colección de los conjuntos cerrados.

Consideramos ahora a la clase

$$\mathcal{G} = \left\{ A \in \mathcal{B}_E : \begin{array}{l} \text{Para todo } \epsilon > 0, \text{ existe } F \text{ cerrado y } G \text{ abierto} \\ \text{tales que } F \subseteq A \subseteq G \text{ y } \mathbb{P}(G - F) < \epsilon. \end{array} \right\}.$$

Observemos que esta clase contiene a la clase de los conjuntos cerrados y está contenida en \mathcal{B}_E , por lo que basta ver que \mathcal{G} es una σ -álgebra.

Sea $\epsilon > 0$, dada una sucesión $A_n \in \mathcal{G}$ escogemos conjuntos cerrados F_n y abiertos G_n tales que

$$F_n \subseteq A_n \subseteq G_n \quad \text{y} \quad \mathbb{P}(G_n - F_n) < \frac{\epsilon}{2^{n+1}}.$$

Como $\bigcup_{k \leq n} F_k$ es una sucesión creciente a $\bigcup_{k=1}^{\infty} F_k$, se tiene que $\lim_{n \uparrow \infty} \mathbb{P}\left(\bigcup_{k \leq n} F_k\right) = \mathbb{P}\left(\bigcup_{k=1}^{\infty} F_k\right)$, por lo que para $\frac{\epsilon}{2} > 0$ existe un entero N_0 tal que para toda $n \geq N_0$,

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} F_k\right) - \mathbb{P}\left(\bigcup_{k \leq n} F_k\right) = \mathbb{P}\left(\bigcup_{k=1}^{\infty} F_k - \bigcup_{k \leq n} F_k\right) < \frac{\epsilon}{2}.$$

En particular si hacemos

$$F = \bigcup_{n \leq N_0} F_n \quad \text{y} \quad G = \bigcup_{n=1}^{\infty} G_n.$$

tendremos que

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} F_n - F\right) < \frac{\epsilon}{2}.$$

obteniendo finalmente que $F \subseteq \bigcup_n A_n \subseteq G$ y además,

$$\mathbb{P}(G - F) = \mathbb{P}\left(G - \bigcup_{n=1}^{\infty} F_n\right) + \mathbb{P}\left(\bigcup_{n=1}^{\infty} F_n - F\right) < \sum_{n=1}^{\infty} \frac{\epsilon}{2^{n+1}} + \frac{\epsilon}{2} = \epsilon.$$

Con esto hemos probado que la unión numerable de elementos de \mathcal{G} vuelve a estar en la clase; claramente \mathcal{G} es cerrada bajo complementos, concluyendo así que \mathcal{G} es una σ -álgebra. \square

Denotamos por $C_b(E, \mathbb{R})$ a la clase de funciones reales continuas y acotadas definidas en un espacio métrico arbitrario E .

A continuación veremos una manera de aproximar funciones indicadoras $\mathbf{1}_F$ de conjuntos cerrados F por elementos en $C_b(E, \mathbb{R})$. El siguiente resultado podría demostrarse argumentando que todo espacio métrico es regular en el sentido topológico, pero nos interesa la construcción de una función específica utilizada en la demostración que posteriormente utilizaremos.

Teorema 9.

Si F es cerrado y $\epsilon > 0$, entonces existe una función $f \in C_b(E, \mathbb{R})$ tal que

$$\begin{aligned} 0 \leq f(x) \leq 1 & \quad \text{para toda } x \in E, \\ f(x) = 1 & \quad \text{si } x \in F, \\ \text{y } f(x) = 0 & \quad \text{si } d(x, F) \geq \epsilon. \end{aligned}$$

La función f incluso puede ser tomada uniformemente continua.

Demostración. Definimos una función continua γ de variable real dada por

$$\gamma(t) = \begin{cases} 1 & \text{si } t \leq 0 \\ 1 - t & \text{si } 0 \leq t \leq 1, \\ 0 & \text{si } 1 \leq t. \end{cases} \quad (2.1)$$

Si

$$f(x) = \gamma\left(\frac{1}{\epsilon}d(x, F)\right) \quad (2.2)$$

entonces f tiene las propiedades requeridas e incluso es uniformemente continua. \square

Teorema 10.

Dos medidas de probabilidad \mathbb{P} y \mathbb{P}' sobre un espacio medible (E, \mathcal{B}_E) , coinciden si

$$\int_E f \, d\mathbb{P} = \int_E f \, d\mathbb{P}'$$

para cada $f \in C_b(E, \mathbb{R})$.

Demostración. Supongamos que F es cerrado. Utilizamos la función (2.1) y definimos para cada entero positivo m ,

$$\gamma_m(t) = \gamma(mt)$$

y

$$f_m(x) = \gamma_m(d(x, F)).$$

Entonces $\{f_m\}_{m \in \mathbb{N}}$ es una sucesión de elementos en $C_b(E, \mathbb{R})$ e uniformemente continuas que convergen puntualmente a $\mathbf{1}_F$. Por el Teorema de la Convergencia Dominada, obtenemos que

$$\mathbb{P}(F) = \lim_{m \rightarrow \infty} \int f_m \, d\mathbb{P} \quad \text{y} \quad \mathbb{P}'(F) = \lim_{m \rightarrow \infty} \int f_m \, d\mathbb{P}'.$$

de manera que por nuestra hipótesis tendremos que $\mathbb{P}(F) = \mathbb{P}'(F)$. Como \mathbb{P} , \mathbb{P}' coinciden en todo conjunto cerrado, se sigue del Teorema 8 que \mathbb{P} y \mathbb{P}' son idénticas. \square

Al integrar sobre el espacio completo omitimos la región de integración. Entonces, el teorema anterior nos dice, que los valores de $\int f \, d\mathbb{P}$ para $f \in C_b(E, \mathbb{R})$ uniformemente continua, completamente determinan los valores de $\mathbb{P}(A)$ para toda $A \in \mathcal{B}_E$.

2.2. El Teorema de Pormanteau

Notemos que por el Teorema 10 las integrales $\int f \, d\mathbb{P}$ determinan a \mathbb{P} , lo cual impide que la sucesión \mathbb{P}_n converja débilmente a dos límites distintos a la vez.

Es importante hacer notar que la convergencia débil depende únicamente de la topología con la cual se dota a E , además de que métricas equivalentes — que generan la misma topología — dan lugar a la misma σ -álgebra de Borel y al mismo espacio $C_b(E, \mathbb{R})$ y en consecuencia a la misma noción de convergencia débil.

Dicho lo anterior, dado un espacio métrico E , es de interés considerar en su totalidad al conjunto de medidas de probabilidad y así tener una noción más sensible sobre la convergencia débil.

Recordemos que $C_b(E, \mathbb{R})$ denota la clase de funciones reales continuas y acotadas en E , este espacio de Banach cuenta con un espacio adjunto $C_b(E, \mathbb{R})^*$ que también

es de Banach. Sea ρ la métrica en $C_b(E, \mathbb{R})$ – ya que toda norma induce una métrica – por definición,

$$C_b(E, \mathbb{R})^* = \{G : (C_b(E, \mathbb{R}), \tau_\rho) \rightarrow \mathbb{R} : G \text{ es lineal y continua}\}$$

Es decir, cada $G \in C_b(E, \mathbb{R})^*$ es continua respecto a la topología τ_ρ de $C_b(E, \mathbb{R})$ determinada por su métrica ρ , pero esta no siempre es la más pequeña. Por esta razón es común otorgarle la llamada *topología débil* τ_w , que por definición es la topología más pequeña ($\tau_w \subseteq \tau_\rho$) que hace a cada $G \in C_b(E, \mathbb{R})^*$ continua ($G : (C_b(E, \mathbb{R}), \tau_w) \rightarrow \mathbb{R}$).

Es claro que entonces τ_w deberá contener y más aún, podrá ser generada a partir de conjuntos de la forma

$$V(\epsilon, G_1, G_2, \dots, G_m)(g) = \bigcap_{i=1}^m \{f \in C_b(E, \mathbb{R}) : |G_i(f) - G_i(g)| < \epsilon\}$$

con $\epsilon > 0$ y $G_1, G_2, \dots, G_m \in C_b(E, \mathbb{R})^*$. De esto tendremos que una base local para $g \in C_b(E, \mathbb{R})$, está dada por:

$$\mathfrak{B}_g = \{V(\epsilon, G_1, G_2, \dots, G_m)(g) \mid \epsilon > 0 \text{ y } G_1, \dots, G_m \in C_b(E, \mathbb{R})^*\}_{m \in \mathbb{N}}.$$

La convergencia de sucesiones en esta topología equivale a lo siguiente, sea $(f_n)_{n \in \mathbb{N}}$ una sucesión en $C_b(E, \mathbb{R})$, entonces:

$$f_n \rightarrow f_0 \text{ relativo a } \tau_w \text{ si y solo si } G(f_n) \rightarrow G(f_0) \text{ para toda } G \in C_b(E, \mathbb{R})^*.$$

De manera análoga, es posible considerar la topología débil en $C_b(E, \mathbb{R})^*$ que hace a cada elemento de $C_b(E, \mathbb{R})^{**}$ –el doble dual– continua, así pues una base local para $G \in C_b(E, \mathbb{R})^*$ esta dada por:

$$\mathfrak{B}_G = \{V(\epsilon, F_1, F_2, \dots, F_m)(G) : \epsilon > 0 \text{ } F_1, \dots, F_m \in C_b(E, \mathbb{R})^{**}\}_{m \in \mathbb{N}}.$$

De igual forma la convergencia de una sucesión de elementos en $C_b(E, \mathbb{R})^*$ bajo esta topología esta dada por,

$$G_n \rightarrow G \text{ relativo a } \tau_w \text{ si y solo si } F(G_n) \rightarrow F(G) \text{ para toda } F \in C_b(E, \mathbb{R})^{**}.$$

Más aún, es un resultado conocido que el espacio $C_b(E, \mathbb{R})^*$ coincidirá con el espacio de las medidas localmente finitas o de Radon definidas en \mathcal{B}_E . Como $C_b(E, \mathbb{R})^*$

es un espacio dual, es posible definir otra topología más en $C_b(E, \mathbb{R})^*$ que llamaremos topología débil*, denotada por τ_{w^*} , la cual sólo tiene sentido en un espacio vectorial dual de un espacio de Banach. Ya que los elementos de $C_b(E, \mathbb{R})^*$ son funcionales sobre $C_b(E, \mathbb{R})$, consideraremos la funcionales definidas sobre $C_b(E, \mathbb{R})^*$ de la forma $\mathcal{L} \rightarrow \mathcal{L}[f]$ para alguna $f \in C_b(E, \mathbb{R})$ fija; resulta que éstas funcionales $\mathcal{L} \in C_b(E, \mathbb{R})^{**}$, en las que estaremos interesados para definir ésta nueva topología, son expresadas de manera única en forma de integral,

$$\mathcal{L}[f](\mu) = \int f \, d\mu,$$

donde f es una función que determina a la funcional. Efectivamente, la topología débil* τ_{w^*} se define como aquella que hace a cada funcional $\mathcal{L}[f] : C_b(E, \mathbb{R})^* \rightarrow \mathbb{R}$ continua para toda $f \in C_b(E, \mathbb{R})$ y es la menor con dicha propiedad. Si consideramos a $C_b(E, \mathbb{R})^*$ con la topología débil* (τ_{w^*}) y restringimos esta topología al subespacio $Z(E)$ de medidas de probabilidad, tendremos que $(Z(E), \tau_{w^*|Z(E)})$ será un espacio topológico cuyos abiertos básicos serán de la forma

$$V(\epsilon, \mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_m)(\mathbb{P}) = \bigcap_{i=1}^m \left\{ \mathbb{Q} \in Z(E) : \left| \int f_i \, d\mathbb{Q} - \int f_i \, d\mathbb{P} \right| < \epsilon \right\},$$

donde $f_1, f_2, \dots, f_m \in C_b(E, \mathbb{R})$, $\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_m \in C_b(E, \mathbb{R})^{**}$, $\mathcal{L}_i(\mathbb{Q}) = \int f_i \, d\mathbb{Q}$ y $\mathcal{L}_i(\mathbb{P}) = \int f_i \, d\mathbb{P}$.

A τ_{w^*} también se le conoce como la topología de la convergencia débil y por la estructura de los abiertos básicos de esta topología tendremos que para una sucesión de elementos \mathbb{P}_n en $Z(E)$,

$$\begin{array}{lll} \mathbb{P}_n \xrightarrow{w} \mathbb{P} & & \\ \text{si y solo si} & \lim_{n \rightsquigarrow \infty} \mathbb{P}_n = \mathbb{P} & \text{relativo a } \tau_{w^*|Z(E)} \\ \text{si y solo si} & \lim_{n \rightsquigarrow \infty} \mathcal{L}[f](\mathbb{P}_n) = \mathcal{L}[f](\mathbb{P}) & \text{para toda } f \in C_b(E, \mathbb{R}). \end{array}$$

es decir, tendremos que la convergencia débil de medidas de probabilidad se refiere simplemente a la convergencia puntual en el espacio topológico $(Z(E), \tau_{w^*|Z(E)})$.

El siguiente teorema nos da condiciones equivalentes para la convergencia débil; cualquiera de estas condiciones podría haber servido como definición.

Un conjunto $A \in \mathcal{B}_E$ es llamado conjunto \mathbb{P} -continuo, si su frontera ∂A satisface que $\mathbb{P}(\partial A) = 0$.

Teorema 11. (Pormanteau) Sean \mathbb{P}_n, \mathbb{P} medidas de probabilidad definidas sobre el espacio medible (E, \mathcal{B}_E) . Las siguientes cinco condiciones son equivalentes,

i) $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$

ii) $\lim_{n \rightsquigarrow \infty} \int f d\mathbb{P}_n = \int f d\mathbb{P}$ para toda $f \in C_b(E, \mathbb{R})$.

iii) $\limsup_{n \rightsquigarrow \infty} \mathbb{P}_n(F) \leq \mathbb{P}(F)$ para todo cerrado F .

iv) $\liminf_{n \rightsquigarrow \infty} \mathbb{P}_n(G) \geq \mathbb{P}(G)$ para todo abierto G .

v) $\lim_{n \rightsquigarrow \infty} \mathbb{P}_n(A) = \mathbb{P}(A)$ para todo conjunto \mathbb{P} -continuo A .

Demostración. $i) \Rightarrow ii)$] Esta implicación es clara por la definición dada.

$ii) \Rightarrow iii)$] Supongamos que (ii) es válido y que F es cerrado. Sea $\delta > 0$ arbitraria pero fija.

Gracias a la regularidad de la medida de probabilidad, es posible tomar una $\epsilon > 0$ suficientemente pequeña y definir el conjunto abierto $G = \{x \in E : d(x, F) < \epsilon\}$ de manera que satisfaga que

$$\mathbb{P}(G) < \mathbb{P}(F) + \delta,$$

ya que los conjuntos definidos como G decrecen hacia F conforme $\epsilon \downarrow 0$.

Ahora bien, para el cerrado F y la $\epsilon > 0$ con la que se definió a G , sabemos existe una función $f \in C(E, \mathbb{R})$ que tomaremos uniformemente continua tal que

$$f(x) = 1 \text{ en } F, \quad f(x) = 0 \text{ en } G^c, \quad f(E) \subseteq [0, 1].$$

Como (ii) es válido, tenemos la igualdad $\lim_{n \rightarrow \infty} \int f d\mathbb{P}_n = \int f d\mathbb{P}$, que junto a las siguientes relaciones:

$$\mathbb{P}_n(F) = \int_F f d\mathbb{P}_n \leq \int f d\mathbb{P}_n$$

y

$$\int f d\mathbb{P} = \int_G f d\mathbb{P} \leq \mathbb{P}(G) < \mathbb{P}(F) + \delta,$$

implican que

$$\limsup_{n \rightsquigarrow \infty} \mathbb{P}_n(F) \leq \lim_{n \rightsquigarrow \infty} \int f d\mathbb{P}_n = \int f d\mathbb{P} < \mathbb{P}(F) + \delta,$$

y ya que la δ fue arbitraria, se sigue (iii).

(iii) \Rightarrow i)] Supongamos que se cumple (iii) y que $f \in C_b(E, \mathbb{R})$. Primero mostraremos que

$$\limsup_{n \rightarrow \infty} \int f \, d\mathbb{P}_n \leq \int f \, d\mathbb{P} \quad (2.3)$$

Transformando f linealmente – con coeficiente positivo en el término de primer grado – es posible reducir la prueba al caso en que $0 < f(x) < 1$ para todo x . Para un entero positivo k , fijo por el momento, definimos los siguientes conjuntos cerrados:

$$F_i = \left\{ x : \frac{i}{k} \leq f(x) \right\}, \quad \text{para } i = 0, 1, 2, 3, \dots, k.$$

Ya que $0 < f(x) < 1$, tendremos que

$$\sum_{i=1}^k \frac{i-1}{k} \mathbb{P} \left\{ x : \frac{i-1}{k} \leq f(x) < \frac{i}{k} \right\} \leq \int f \, d\mathbb{P} < \sum_{i=1}^k \frac{i}{k} \mathbb{P} \left\{ x : \frac{i-1}{k} \leq f(x) < \frac{i}{k} \right\}.$$

Utilizando los conjuntos cerrados antes definidos, podemos reescribir la suma derecha como:

$$\sum_{i=1}^k \frac{i}{k} [\mathbb{P}(F_{i-1}) - \mathbb{P}(F_i)] = \frac{1}{k} + \frac{1}{k} \sum_{i=1}^k \mathbb{P}(F_i).$$

Análogamente para la suma del extremo izquierdo tendremos que:

$$\sum_{i=1}^k \frac{i-1}{k} [\mathbb{P}(F_{i-1}) - \mathbb{P}(F_i)] = \frac{1}{k} \sum_{i=1}^k \mathbb{P}(F_i).$$

Así pues, estas nuevas identidades nos dan como resultado que

$$\frac{1}{k} \sum_{i=1}^k \mathbb{P}(F_i) \leq \int f \, d\mathbb{P} < \frac{1}{k} + \frac{1}{k} \sum_{i=1}^k \mathbb{P}(F_i).$$

Como (iii) es válido, para cada $i = 1, 2, \dots, k$ $\limsup_{n \rightarrow \infty} \mathbb{P}(F_i) \leq \mathbb{P}(F_i)$, esto a la vez que aplicamos la desigualdad derecha de la relación anterior a cada \mathbb{P}_n y la desigualdad izquierda a \mathbb{P} nos da como resultado que

$$\limsup_{n \rightarrow \infty} \int f \, d\mathbb{P}_n \leq \frac{1}{k} + \int f \, d\mathbb{P},$$

por lo que cuando $k \rightarrow \infty$, se obtiene (2.3). Aplicando ahora (2.3) a $-f$ nos da como resultado

$$\liminf_{n \rightarrow \infty} \int f \, d\mathbb{P}_n \geq \int f \, d\mathbb{P},$$

que conjuntamente con (2.3), prueba la convergencia débil.

iii) \Leftrightarrow *iv*)] La equivalencia entre (*iii*) y (*iv*) se sigue fácilmente de tomar complementos respectivamente de los conjuntos cerrados y abiertos, y recordar la definición de límite superior e inferior para una sucesión de reales.

iii) \Rightarrow *v*)] Denotamos por $\text{Int}(A)$ y \bar{A} al interior y la cerradura respectivamente de un conjunto A . Ahora, si tenemos que se cumple (*iii*) entonces también es válido (*iv*), por lo que para cada $A \in \mathcal{B}_E$,

$$\begin{aligned} \mathbb{P}(\bar{A}) &\geq \limsup_{n \rightsquigarrow \infty} \mathbb{P}_n(\bar{A}) \geq \limsup_{n \rightsquigarrow \infty} \mathbb{P}_n(A) \\ &\geq \liminf_{n \rightsquigarrow \infty} \mathbb{P}_n(A) \geq \liminf_{n \rightsquigarrow \infty} \mathbb{P}_n(\text{Int}(A)) \geq \mathbb{P}(\text{Int}(A)) \end{aligned}$$

Sin embargo, si A es un conjunto \mathbb{P} -continuo, es decir $\mathbb{P}(\partial A) = 0$, tendremos que

$$\mathbb{P}(\text{Int}(A)) \leq \mathbb{P}(A) \leq \mathbb{P}(\bar{A}) = \mathbb{P}(\partial A \cup \text{int}(A)) \leq \mathbb{P}(\text{Int}(A))$$

Luego, $\mathbb{P}(\text{Int}(A)) = \mathbb{P}(A) = \mathbb{P}(\bar{A})$ y por lo tanto $\lim_{n \rightsquigarrow \infty} \mathbb{P}_n(A) = \mathbb{P}(A)$.

v) \Rightarrow *iii*)] Ya que $\partial\{x : d(x, F) \leq \delta\} \subseteq \{x : d(x, F) = \delta\}$ y además

$$\partial\{x : d(x, F) \leq \delta\} \cap \partial\{x : d(x, F) \leq \delta'\} = \emptyset \text{ si } \delta \neq \delta',$$

se tiene que la unión de a lo más una cantidad numerable de conjuntos frontera de este tipo puede tener medida \mathbb{P} positiva. Luego, dada una sucesión de reales positivos $(\delta_k)_{k \in \mathbb{N}}$ tales que $\delta_k \downarrow 0$ cuando $k \rightarrow \infty$, los conjuntos $F_k = \{x : d(x, F) \leq \delta_k\}$ son conjuntos \mathbb{P} -continuos.

Como (*v*) es válido tenemos que

$$\limsup_{n \rightsquigarrow \infty} \mathbb{P}_n(F) \leq \lim_{n \rightsquigarrow \infty} \mathbb{P}_n(F_k) = \mathbb{P}(F_k) \text{ para cada } k.$$

Luego si F es cerrado tendremos que $\bigcap_k F_k = F$, de manera que tomando el límite cuando $k \rightarrow \infty$, obtenemos que

$$\limsup_{n \rightsquigarrow \infty} \mathbb{P}_n(F) \leq \lim_{k \rightsquigarrow \infty} \mathbb{P}(F_k) = \mathbb{P}(F),$$

de donde se sigue (*iii*).

Con esto hemos finalizado la prueba del Teorema de Pormanteau. \square

2.3. Otros criterios de convergencia débil.

En ocasiones nos encontramos en situaciones en las que es conveniente probar la convergencia débil mostrando que $\mathbb{P}_n(A) \rightarrow \mathbb{P}(A)$ cuando $n \rightarrow \infty$ para alguna clase especial de conjuntos.

Teorema 12. *Sea \mathcal{U} una subclase de \mathcal{B}_E tal que*

- (i) *Es cerrada bajo la intersección finita de sus elementos.*
- (ii) *Cada conjunto abierto en \mathcal{B}_E es unión finita o numerable de elementos de \mathcal{U} .*

Si $\mathbb{P}_n(A) \rightarrow \mathbb{P}(A)$ para todo $A \in \mathcal{U}$ entonces $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$.

Demostración. Si A_1, \dots, A_m son elementos de \mathcal{U} , también lo serán sus intersecciones finitas; por tanto, tendremos que

$$\begin{aligned} \mathbb{P}_n\left(\bigcup_{i=1}^m A_i\right) &= \sum_i \mathbb{P}_n(A_i) - \sum_{i<j} \mathbb{P}_n(A_i \cap A_j) + \sum_{i<j<k} \mathbb{P}_n(A_i \cap A_j \cap A_k) - \dots \\ &\xrightarrow{n \rightsquigarrow \infty} \sum_i \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \sum_{i<j<k} \mathbb{P}(A_i \cap A_j \cap A_k) - \dots \\ &= \mathbb{P}\left(\bigcup_{i=1}^m A_i\right). \end{aligned}$$

Si G es un conjunto abierto, entonces existe $I \subset \mathbb{N}$ tal que $G = \bigcup_{i \in I} A_i$ para alguna sucesión $(A_i)_{i \in I}$ de elementos de \mathcal{U} . Entonces para $\epsilon > 0$, es posible elegir un entero positivo m , de manera que

$$\mathbb{P}(G) - \mathbb{P}\left(\bigcup_{i=1}^m A_i\right) < \epsilon.$$

Con el límite anterior tendremos que,

$$\liminf_{n \rightsquigarrow \infty} \mathbb{P}_n(G) \geq \lim_{n \rightsquigarrow \infty} \mathbb{P}\left(\bigcup_{i=1}^m A_i\right) = \mathbb{P}\left(\bigcup_{i=1}^m A_i\right) > \mathbb{P}(G) - \epsilon$$

Ya que la ϵ fue arbitraria, la condición (iv) del Teorema de Pormanteau se cumple. \square

Denotamos por $B_\epsilon(x)$ a la bola abierta de radio ϵ centrada en x .

Corolario 13. *Sea \mathcal{U} una subclase de \mathcal{B}_E tal que*

- (i) *Es cerrada bajo la intersección finita de sus elementos.*
- (ii) *Para todo $x \in E$ y toda $\epsilon > 0$, existe un conjunto $A \in \mathcal{U}$ con $x \in \text{Int}(A) \subset A \subset B_\epsilon(x)$.*

Si E es separable y $\mathbb{P}_n(A) \rightarrow \mathbb{P}(A)$ para todo $A \in \mathcal{U}$ entonces $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$.

Demostración. Sea G un abierto. Como E es separable existe una base numerable de abiertos $\{B_n\}_{n \in \mathbb{N}}$ que genera la topología de E , por lo que para todo $x \in G$, existe un abierto básico B_n tal que $x \in B_n \subset G$, de donde es claro que $G = \bigcup_{n \in I} B_n$, para algún $I \subset \mathbb{N}$. Esto último aunado a la condición (ii) de la hipótesis nos dice que para todo punto x de G es posible encontrar un elemento $A_x \in \mathcal{U}$ tal que $x \in B_n \subset \text{Int}(A_x) \subset A_x \subset G$, lo cual implica que existe una colección a lo más numerable $\{A_n\}_{n \in I}$ tal que $G = \bigcup_{n \in I} A_n$, donde $A_n = A_x$ para alguna x . Con esto la clase \mathcal{U} satisface la condición (ii) del teorema anterior. \square

Corolario 14. *Supongamos que para cada intersección finita de bolas abiertas A , se tiene que $\mathbb{P}_n(A) \rightarrow \mathbb{P}(A)$, dado que A sea un conjunto \mathbb{P} -continuo. Si E es separable, entonces $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$.*

Demostración. La frontera $\partial B_\epsilon(x)$ está contenida en el conjunto $\{x : d(x, F) = \epsilon\}$ y estos últimos son ajenos cuando hacemos variar la ϵ (fijando x), en consecuencia éstas tienen medida- \mathbb{P} nula, con excepción de a lo más una cantidad numerable de ellas. Ahora bien, ya que

$$\partial(A \cap B) \subset (\partial A) \cap (\partial B),$$

se sigue que las hipótesis del Corolario 13, se satisfacen para la clase \mathcal{U} de conjuntos \mathbb{P} -continuos que además son intersección finita de bolas abiertas, probando así el resultado deseado. \square

Recordemos que (E, d) es un espacio métrico y \mathcal{B}_E la σ -álgebra de Borel generada por los conjuntos abiertos de E . Introducimos las siguientes nociones para una subclase \mathcal{V} de \mathcal{B}_E ,

Definición 15. (Clase-Determinante)

Llamamos a una subclase \mathcal{V} de \mathcal{B}_E , Clase-Determinante, si dos medidas \mathbb{P} y \mathbb{Q} son idénticas siempre que coincidan en los elementos de dicha subclase.

Ejemplo. La clase de conjuntos cerrados y cualquier álgebra de conjuntos que genere a \mathcal{B}_E , son ejemplos de clases-determinantes.

Definición 16. (Clase que Determina la Convergencia)

Llamamos a una subclase \mathcal{V} de \mathcal{B}_E , Clase que Determina la Convergencia, si el hecho de que $\mathbb{P}_n(A) \rightarrow \mathbb{P}(A)$ para todo conjunto \mathbb{P} -continuo $A \in \mathcal{V}$, implica $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$.

A la luz de ésta definición, el Corolario 14 nos dice lo siguiente

Ejemplo. En un espacio separable las intersecciones finitas de bolas abiertas constituyen una clase que determina la convergencia.

Claramente, por la unicidad de los límites tendremos que

$$\text{Clase que Determina la Convergencia} \Rightarrow \text{Clase-Determinante}$$

Cerramos la sección con otra condición para convergencia débil, para esto recordemos que una sucesión de números reales $(a_n)_{n \in \mathbb{N}}$ converge a un límite a si y solo si toda subsucesión $(a_{n'})_{n \in \mathbb{N}}$ contiene otra subsucesión $(a_{n''})_{n \in \mathbb{N}}$ que converge a a . De este hecho se deduce su análogo para la convergencia débil.

Teorema 17. $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ si y solo si cada subsucesión $\{\mathbb{P}_{n'}\}_{n \in \mathbb{N}}$ contiene otra subsucesión $\{\mathbb{P}_{n''}\}_{n \in \mathbb{N}}$ tal que $\mathbb{P}_{n''} \xrightarrow{w} \mathbb{P}$.

2.4. Espacio Producto

Sea $E = E' \times E''$ donde (E', d') y (E'', d'') son espacios métricos y separables. Entonces E es un espacio métrico y separable, ya que el espacio topológico generado por el producto cartesiano de espacios separables y metrizable vuelve a ser separable y metrizable dotado con la topología producto (Tychonoff), que para el caso de un producto finito coincide con la topología de la caja. Por lo tanto, una base para la topología de E serán los rectángulos $A' \times A''$, donde A' y A'' son abiertos en E' y E'' respectivamente y por lo tanto son borelianos.

Luego si E es separable, las σ -álgebras $\mathcal{B}_E, \mathcal{B}_{E'}, \mathcal{B}_{E''}$ de conjuntos de Borel de los respectivos espacios estarán relacionadas de la siguiente manera

$$\mathcal{B}_E = \mathcal{B}_{E'} \otimes \mathcal{B}_{E''}$$

donde \mathcal{B}_E denota a la σ -álgebra generada por los rectángulos medibles, o bien, la topología producto.

Las dos distribuciones marginales de una medida de probabilidad \mathbb{P} , en el espacio medible (E, \mathcal{B}_E) , estarán definidas por:

$$\begin{aligned} \mathbb{P}'(A') &= \mathbb{P}(A' \times E''), & A' \in \mathcal{B}_{E'} \\ \mathbb{P}''(A'') &= \mathbb{P}(E' \times A''), & A'' \in \mathcal{B}_{E''}. \end{aligned}$$

Teorema 18.

Si E es un espacio separable, una condición necesaria y suficiente para que $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ es que

$$\mathbb{P}_n(A' \times A'') \xrightarrow{n \rightarrow \infty} \mathbb{P}(A' \times A'')$$

para cada conjunto \mathbb{P}' -continuo A' y cada conjunto \mathbb{P}'' -continuo A'' , donde $\mathbb{P}', \mathbb{P}''$ son las distribuciones marginales de \mathbb{P} .

Demostración. Ya que

$$\partial(A' \times A'') \subset ((\partial A') \times E'') \cup (E' \times (\partial A'')),$$

la condición es necesaria.

Para probar la suficiencia, aplicamos el Corolario 13 del Teorema 12 a la clase \mathcal{U} , donde

$$\mathcal{U} = \{A' \times A'' \mid A' \text{ es un conjunto } \mathbb{P}'\text{-continuo y } A'' \text{ es un conjunto } \mathbb{P}''\text{-continuo}\}.$$

La clase \mathcal{U} es cerrada bajo intersecciones, y por hipótesis $\mathbb{P}_n(A) \rightarrow \mathbb{P}(A)$ para todo $A \in \mathcal{U}$.

Sea $(a, b) \in E$ y $\epsilon > 0$, consideremos a los conjuntos:

$$A_\delta = \{x \in E' : d'(x, a) < \delta\} \times \{y \in E'' : d''(y, b) < \delta\}.$$

Observemos que las colecciones

$$(\partial\{x \in E' : d'(x, a) < \delta\})_{0 < \delta < \epsilon} \quad , \quad (\partial\{y \in E'' : d''(y, b) < \delta\})_{0 < \delta < \epsilon}$$

están conformadas respectivamente por conjuntos ajenos. En consecuencia, para alguna $\delta > 0$ con $0 < \delta < \epsilon$, se tiene que $A_\delta \in \mathcal{U}$. Si E es dotado de la métrica definida por:

$$d((a_1, b_1), (a_2, b_2)) = \text{máx}\{d'(a_1, a_2), d''(b_1, b_2)\},$$

entonces A_δ es únicamente la bola abierta de radio δ centrada en (a, b) . Por lo tanto, \mathcal{U} satisface las condiciones del Corolario 13 del Teorema 12 como se requería. \square

Nótese que la suficiencia de la condición del teorema anterior nos dice que los rectángulos medibles conforman una Clase que Determina la Convergencia.

Para medidas de probabilidad \mathbb{P}' y \mathbb{P}'' definidas sobre $(E', \mathcal{B}_{E'})$ y $(E'', \mathcal{B}_{E''})$ respectivamente, la medida producto $\mathbb{P}' \otimes \mathbb{P}''$ es una medida sobre $\mathcal{B}_{E'} \otimes \mathcal{B}_{E''}$, en consecuencia, si E es separable, será una medida sobre \mathcal{B}_E . El siguiente teorema, en el que $\mathbb{P}'_n, \mathbb{P}'$ son medidas de probabilidad sobre $(E', \mathcal{B}_{E'})$ y $\mathbb{P}''_n, \mathbb{P}''$ son medidas de probabilidad sobre $(E'', \mathcal{B}_{E''})$, es una consecuencia inmediata del teorema anterior.

Teorema 19.

Sea E un espacio separable, entonces

$$\mathbb{P}'_n \otimes \mathbb{P}''_n \xrightarrow{w} \mathbb{P}' \otimes \mathbb{P}''$$

si y solo si

$$\mathbb{P}'_n \xrightarrow{w} \mathbb{P}' \quad y \quad \mathbb{P}''_n \xrightarrow{w} \mathbb{P}''.$$

2.5. Convergencia en Distribución

Sea X una variable aleatoria definida sobre un espacio de probabilidad $(\Omega, \mathcal{B}, \mathbb{P})$ y con valores en un espacio métrico (E, d) . La distribución de X es la medida de probabilidad $\mathbb{P}_X = \mathbb{P} \circ X^{-1}$ sobre el espacio medible (E, \mathcal{B}_E) .

$$\mathbb{P}_X(A) = \mathbb{P}(X^{-1}(A)) = \mathbb{P}\{\omega : X(\omega) \in A\} = \mathbb{P}(X \in A), \quad A \in \mathcal{B}_E.$$

Obsérvenos que aquí, \mathbb{P} representa una medida de probabilidad sobre un espacio de naturaleza arbitraria Ω , mientras que \mathbb{P}_X se encuentra siempre definida sobre el espacio métrico E .

Así pues, sea $h : E \rightarrow \mathbb{R}$ una función \mathcal{B}_E -medible ($h^{-1}(\mathcal{B}_{\mathbb{R}}) \subseteq \mathcal{B}_E$), entonces, la función real $h \circ X : \Omega \rightarrow \mathbb{R}$, es \mathcal{B} -medible. Bajo estas circunstancias, por la fórmula

de cambio de variable, se tiene que h es integrable respecto a $\mathbb{P} \circ X^{-1}$ si y solo si $h(X)$ es integrable respecto a \mathbb{P} , en cuyo caso se tiene que

$$\int_{X^{-1}(A)} h(X(\omega)) \, d\mathbb{P}(d\omega) = \int_A h(\omega') \, d\mathbb{P} \circ X^{-1}(d\omega') \quad \text{para cada } A \in \mathcal{B}_E.$$

En particular, tendremos que

$$\int h(X) \, d\mathbb{P} = \int h \, d\mathbb{P}_X = \mathbb{E}[h(X)].$$

Definición 20. (Convergencia en Distribución)

Decimos que una sucesión de variables aleatorias X_1, X_2, \dots , converge en distribución a la variable aleatoria X , denotado por

$$X_n \xrightarrow{\mathcal{D}} X,$$

si la sucesión de distribuciones \mathbb{P}_n de las X_n convergen débilmente a la distribución \mathbb{P} de X ,

$$\mathbb{P}_n \xrightarrow{w} \mathbb{P}.$$

La definición de convergencia en distribución no tiene sentido cuando el espacio imagen (rango) E y su topología no es la misma para cada una de las variables aleatorias X, X_1, X_2, \dots , sin embargo, es posible que tengan distintos espacios de probabilidad como dominio de definición. Ordinariamente no se hace mención de éstos espacios ya que su estructura solo entra en discusión por la vía de las distribuciones que inducen sobre E . Cabe aclarar el abuso de notación, ya que usualmente se escribe $\mathbb{P}(X_n \in A)$ cuando realmente deberíamos escribir $\mathbb{P}_n(X_n \in A)$ y escribimos $\mathbb{E}[f(X_n)]$ cuando deberíamos escribir $\int f(X_n) \, d\mathbb{P}_n$ o bien tal vez $\mathbb{E}_n[f(X_n)]$.

Por el momento y sin abuso de notación, tenemos que si \mathbb{P}, \mathbb{P}_n son las medidas del espacio de probabilidad donde se encuentran definidas respectivamente X, X_n , entonces por el teorema de cambio de variable

$$\begin{aligned} \int_E f(x) \, d\mathbb{P} \circ X^{-1}(dx) &= \int_{\Omega} f(X) \, d\mathbb{P} = \mathbb{E}[f(X)], \\ \int_E f(x) \, d\mathbb{P}_n \circ X_n^{-1}(dx) &= \int_{\Omega} f(X_n) \, d\mathbb{P}_n = \mathbb{E}_n[f(X_n)], \end{aligned}$$

por lo que es razonable pedir que $X_n \xrightarrow{\mathcal{D}} X$ si y solo si $\lim_{n \rightsquigarrow \infty} \mathbb{E}_n[f(X_n)] = \mathbb{E}[f(X)]$ para toda $f \in C_b(E, \mathbb{R})$.

Un conjunto A en \mathcal{B}_E es llamado un conjunto X -continuo si $\mathbb{P}(X \in \partial A) = 0$. El siguiente teorema es consecuencia inmediata del Teorema de Pormanteau 11.

Teorema 21. *Los siguientes cinco enunciados son equivalentes*

- i) $X_n \xrightarrow{\mathcal{D}} X$
- ii) $\lim_{n \rightarrow \infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)]$ para toda $f \in C_b(E, \mathbb{R})$.
- iii) $\limsup_{n \rightarrow \infty} \mathbb{P}\{X_n \in F\} \leq \mathbb{P}\{X \in F\}$ para todo cerrado F .
- iv) $\liminf_{n \rightarrow \infty} \mathbb{P}\{X_n \in G\} \geq \mathbb{P}\{X \in G\}$ para todo abierto G .
- v) $\lim_{n \rightarrow \infty} \mathbb{P}\{X_n \in A\} = \mathbb{P}\{X \in A\}$ para todo conjunto X -continuo A .

Cada teorema sobre convergencia débil, puede ser similarmente reenunciado para variables aleatorias.

2.6. Convergencia en Probabilidad

De nueva cuenta, consideramos a (E, d) un espacio métrico.

Definición 22. (Convergencia en Probabilidad)

Si para toda ϵ positiva y un elemento $\mathbf{b} \in E$,

$$\mathbb{P}(\{\omega : d(X_n(\omega), \mathbf{b}) \geq \epsilon\}) \rightarrow 0, \text{ cuando } n \rightarrow \infty,$$

decimos que X_n converge en probabilidad a \mathbf{b} y escribimos

$$X_n \xrightarrow{\mathbb{P}} \mathbf{b}.$$

Si pensamos en \mathbf{b} como un elemento aleatorio constante, tendremos que

$$X_n \xrightarrow{\mathbb{P}} \mathbf{b} \text{ si y solo si } X_n \xrightarrow{\mathcal{D}} \mathbf{b}$$

Visto de otra manera tendremos que

$$X_n \xrightarrow{\mathbb{P}} \mathbf{b} \text{ si y solo si } \mathbb{P}_n \xrightarrow{w} \mathbb{Q}$$

donde \mathbb{P}_n es la distribución de X_n y \mathbb{Q} es una medida de probabilidad tal que $\mathbb{Q}\{\mathbf{b}\} = 1$.

Las variables aleatorias X_n , como ya se ha comentado, pueden estar definidas sobre distintos espacios de probabilidad, pero deben tener un contradominio común.

Si tenemos variables aleatorias X_n, Y_n con el mismo dominio, tiene sentido hablar de la distancia $\rho(X_n, Y_n)$ — una función con valor $\rho(X_n(\omega), Y_n(\omega))$ en ω . Es bien sabido que si E es un espacio separable, tendremos que $\rho(X_n, Y_n)$ es una variable aleatoria.

Para el siguiente teorema, asumimos que el espacio E es separable y que para cada n , X_n y Y_n tienen un dominio común.

Teorema 23. Si $X_n \xrightarrow{\mathcal{D}} X$ y $\rho(X_n, Y_n) \xrightarrow{\mathbb{P}} 0$, entonces $Y_n \xrightarrow{\mathcal{D}} X$.

Demostración. Para cada $\epsilon > 0$ definimos los conjuntos cerrados

$$F_\epsilon = \{x : \rho(x, F) \leq \epsilon\}.$$

Observemos que

$$\begin{aligned} \{Y_n \in F\} &= \{Y_n \in F\} \cap (\{X_n \in F_\epsilon\} \cup \{X_n \notin F_\epsilon\}) \\ &\subseteq \{X_n \in F_\epsilon\} \cup \{Y_n \in F, X_n \notin F_\epsilon\} \\ &\subseteq \{X_n \in F_\epsilon\} \cup \{\rho(Y_n, X_n) \geq \epsilon\} \end{aligned}$$

aclarado esto, es fácil ver que

$$\mathbb{P}\{Y_n \in F\} \leq \mathbb{P}\{\rho(Y_n, X_n) \geq \epsilon\} + \mathbb{P}\{X_n \in F_\epsilon\}$$

Como F_ϵ es cerrado, la hipótesis nos dice que

$$\limsup_{n \rightarrow \infty} \mathbb{P}\{Y_n \in F\} \leq \limsup_{n \rightarrow \infty} \mathbb{P}\{X_n \in F_\epsilon\} \leq \mathbb{P}\{X \in F_\epsilon\}$$

Si F es cerrado, $F_\epsilon \downarrow F$ si $\epsilon \downarrow 0$. Así el resultado se sigue por el Teorema 21. \square

Cerramos la presente sección con un teorema de vital importancia, el cual nos permitirá asegurar que para una sucesión de medidas de probabilidad que converja débilmente podemos encontrar una sucesión de variables aleatorias cuyas distribuciones en su respectivo espacio coincidan con dichas medidas.

Teorema 24. (Teorema de representación de Skorohod.)

Sea $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ una sucesión de medidas de probabilidad sobre un espacio medible (E, \mathcal{E}) donde E es un espacio métrico completo y \mathcal{E} la σ -álgebra de conjuntos de Borel en E ; entonces

$$\mathbb{P}_n \xrightarrow{w} \mathbb{P}_0,$$

si y sólo si existe un proceso E -valuado $(X_n)_{n=0}^\infty$, definido sobre algún espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, de manera que

$$\mathbb{P}X_n^{-1} = \mathbb{P}_n, \quad n = 0, 1, \dots$$

y además

$$X_n \xrightarrow{c.s.} X_0.$$

Véase [7] Teorema 3.1.8. para una prueba de este teorema.

2.7. Funciones Continuas

Sea $h : (E, d) \rightarrow (E', d')$ una función medible entre espacios métricos con respectivas σ -álgebras de Borel $\mathcal{B}_E, \mathcal{B}_{E'}$. Cada medida de probabilidad \mathbb{P} sobre (E, \mathcal{B}_E) induce en $(E', \mathcal{B}_{E'})$ una única medida de probabilidad $\mathbb{P} \circ h^{-1}$, definida por

$$\mathbb{P} \circ h^{-1}(A) = \mathbb{P}(h^{-1}(A)) \quad \text{para toda } A \in \mathcal{B}_{E'}.$$

Nuestro objetivo en esta sección es ver algunos resultados de nuestro interés sobre algunas condiciones bajo las cuales $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ implica $\mathbb{P}_n \circ h^{-1} \xrightarrow{w} \mathbb{P} \circ h^{-1}$.

Una condición fácil de ver, surge cuando h es una función continua, ya que si $f \in C_b(E', \mathbb{R})$ se tendrá que la composición $f \circ h \in C_b(E, \mathbb{R})$, es decir $f \circ h$ será nuevamente una función acotada y continua definida en E . Así las cosas, tendremos que $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ implica

$$\int_E f(h(x)) d\mathbb{P}_n(dx) \rightarrow \int_E f(h(x)) d\mathbb{P}(dx),$$

que por la fórmula de cambio de variable, esto se transforma en

$$\int_{E'} f(y) d\mathbb{P}_n \circ h^{-1}(dy) \rightarrow \int_{E'} f(y) d\mathbb{P} \circ h^{-1}(dy), \quad \text{para toda } f \in C_b(E', \mathbb{R})$$

con lo que finalmente se tendría que las medidas inducidas $\mathbb{P}_n \circ h^{-1}$ convergen débilmente.

La condición de continuidad puede ser debilitada. Supongamos que $h : E \rightarrow E'$ únicamente es medible y sea

$$D_h = \{\text{Discontinuidades de } h\}.$$

Entonces $D_h \in \mathcal{B}_E$.

Teorema 25. Si $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ y $\mathbb{P}(D_h) = 0$, entonces $\mathbb{P}_n \circ h^{-1} \xrightarrow{w} \mathbb{P} \circ h^{-1}$

Demostración. Sea F un subconjunto cerrado de E' . Como $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$, tenemos que

$$\limsup_{n \rightarrow \infty} \mathbb{P}_n(h^{-1}(F)) \leq \limsup_{n \rightarrow \infty} \mathbb{P}_n(\overline{h^{-1}(F)}) \leq \mathbb{P}(\overline{h^{-1}(F)})$$

por lo que basta ver que $\mathbb{P}(\overline{h^{-1}(F)}) = \mathbb{P}(h^{-1}(F))$, que se sigue del hecho de que $\mathbb{P}(D_h) = 0$ y que $\overline{h^{-1}(F)} \subseteq D_h \cup (h^{-1}(F))$ \square

Le siguen dos corolarios inmediatos. Para un elemento aleatorio X de E , $h(X)$ es un elemento aleatorio de E' (asumiendo aún que h es medible).

Corolario 26. Si $X_n \xrightarrow{\mathfrak{D}} X$ y $\mathbb{P}\{X \in D_h\} = 0$, entonces $h(X_n) \xrightarrow{\mathfrak{D}} h(X)$

Corolario 27. Si $X_n \xrightarrow{\mathbb{P}} \mathbf{b}$ y h es continua en \mathbf{b} , entonces $h(X_n) \xrightarrow{\mathbb{P}} h(\mathbf{b})$

Ejemplo Importante. Para variables aleatorias reales, la convergencia $(X_n, Y_n) \xrightarrow{\mathfrak{D}} (X, Y)$ implica la convergencia $X_n + Y_n \xrightarrow{\mathfrak{D}} X + Y$.

Particular interés concede el Teorema 25 cuando $E' = \mathbb{R}$, en cuyo caso $h : E \rightarrow \mathbb{R}$ es una función real medible.

Teorema 28. Se tiene que:

- (i) Si $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$, entonces $\mathbb{P}_n \circ h^{-1} \xrightarrow{w} \mathbb{P} \circ h^{-1}$ para toda función real medible h tal que $\mathbb{P}(D_h) = 0$.
- (ii) Si $\mathbb{P}_n \circ h^{-1} \xrightarrow{w} \mathbb{P} \circ h^{-1}$ para toda $h \in C_b(E, \mathbb{R})$ uniformemente continua, entonces $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$.
- (iii) Si $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ y h es una función real acotada y medible con $\mathbb{P}(D_h) = 0$, entonces

$$\int h d\mathbb{P}_n \rightarrow \int h d\mathbb{P} \quad \text{cuando } n \rightarrow \infty.$$

Demostración. (i) Se sigue del Teorema 25.

(ii) Si $\mathbb{P}_n \circ h^{-1} \xrightarrow{w} \mathbb{P} \circ h^{-1}$, por cambio de variable y la definición de convergencia débil

$$\int f(h(x)) d\mathbb{P}_n(dx) \rightarrow \int f(h(x)) d\mathbb{P}(dx) \quad \text{para todo } f \in C_b(\mathbb{R}).$$

(iii) Es consecuencia de (i), ya que si M es la constante que acota a h , haciendo

$$f(t) = \begin{cases} -M & \text{si } t \leq -M, \\ t & \text{si } -M \leq t \leq M, \\ M & \text{si } t \geq M, \end{cases}$$

tendremos que

$$\int h d\mathbb{P}_n \rightarrow \int h d\mathbb{P}. \quad \square$$

2.8. El Teorema de Prohorov

En esta sección se introduce la noción de tensión para medidas de probabilidad, definidas sobre —como hasta el momento hemos considerado— un espacio métrico (E, d) y su σ -álgebra de Borel \mathcal{B}_E . De esta manera, cimentamos el camino para presentar el Teorema de Prohorov, el cual es una herramienta muy poderosa y útil; en particular, será de nuestro interés a la hora de querer probar la convergencia débil de medidas sobre el espacio de funciones continuas $C([0, 1], \mathbb{R})$.

Definición 29. (Compacidad Relativa)

Una familia de medidas de probabilidad \mathfrak{T} sobre el espacio medible (E, \mathcal{B}_E) , es relativamente compacta, si toda sucesión tiene una subsucesión que converge débilmente; esto es, para toda sucesión $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ en \mathfrak{T} existe una subsucesión $\{\mathbb{P}_{n'}\}_{n' \in \mathbb{N}}$ y una medida de probabilidad \mathbb{Q} —definida sobre el espacio medible (E, \mathcal{B}_E) , pero no necesariamente un elemento de \mathfrak{T} — tal que $\mathbb{P}_{n'} \xrightarrow{w} \mathbb{Q}$.

Es de interés observar, que si consideramos a \mathfrak{T} como subespacio del espacio $Z(E)$ de medidas de probabilidad dotado con la topología débil* y bajo la cual, la convergencia débil es simplemente la convergencia en el espacio topológico $(Z(E), \tau_w^*)$, tendremos que esta definición realmente se refiere a la noción de ser compacto por sucesiones. Además, recordemos que para un espacio métrico las nociones de compacidad, compacto por sucesiones y que todo conjunto infinito tiene un punto límite son equivalentes.

Definición 30. (Medida de Probabilidad Tensa)

Una medida de probabilidad \mathbb{P} sobre un espacio de medida (E, \mathcal{B}_E) , es tensa, si para cada $\epsilon > 0$, existe un conjunto compacto K tal que $\mathbb{P}(K) > 1 - \epsilon$.

Por la regularidad de una medida de probabilidad definida sobre un espacio métrico (Teorema 8), tendremos la siguiente caracterización de una medida de probabilidad tensa sobre un espacio de medida (E, \mathcal{B}_E) .

Lema 31. Una medida de probabilidad \mathbb{P} , es tensa, si y solo si

$$\mathbb{P}(A) = \sup \{ \mathbb{P}(K) : K \subset A, K \text{ compacto} \} \quad \text{para todo } A \in \mathcal{B}_E.$$

Demostración. \Leftarrow] Sea $\epsilon > 0$. Si $A \in \mathcal{B}_E$, consideramos su complemento A^c ; entonces como

$$\begin{aligned} \mathbb{P}(A) &= \sup \{ \mathbb{P}(K) : K \subset A, K \text{ compacto} \} \\ \mathbb{P}(A^c) &= \sup \{ \mathbb{P}(K) : K \subset A^c, K \text{ compacto} \} \end{aligned}$$

tendremos que existen compactos K_1, K_2 subconjuntos de A y A^c respectivamente tales que

$$\begin{aligned} \mathbb{P}(A) - \frac{\epsilon}{2} &< \mathbb{P}(K_1) \\ \mathbb{P}(A^c) - \frac{\epsilon}{2} &< \mathbb{P}(K_2) \end{aligned}$$

de donde obtenemos por la aditividad de la medida que

$$1 - \epsilon < \mathbb{P}(K_1 \cup K_2) = \mathbb{P}(K)$$

y claramente K es compacto por ser unión finita de compactos.

\Rightarrow] Si la medida es tensa, sea $A \in \mathcal{S}$ y consideremos la clase

$$\mathcal{G} = \{ \mathbb{P}(K) : K \subset A, K \text{ compacto} \}.$$

La clase \mathcal{G} es no vacía ya que todo evento consistente de un solo punto de A es compacto en S , además está acotada superiormente por ser una medida de probabilidad. Ahora bien, es claro que si α es el supremo de dicha clase tendremos que $\alpha \leq \mathbb{P}(A)$. Por otro lado, supongamos que $\mathbb{P}(A) < \alpha$, entonces existe K compacto subconjunto de A tal que $\mathbb{P}(A) < \mathbb{P}(K) < \alpha$, lo cual no es posible. Por lo tanto $\mathbb{P}(A) = \alpha$. \square

El siguiente resultado es bastante útil, ya que nos permite afirmar que en todo espacio métrico, completo y separable, toda medida de probabilidad es tensa.

Teorema 32. *Sea E un espacio métrico, completo y separable, entonces cada medida de probabilidad definida sobre (E, \mathcal{B}_E) es tensa.*

Demostración. Como E es separable, entonces existe una base numerable que genera su topología, luego para cada $n \in \mathbb{N}$ consideramos una sucesión de bolas abiertas (abiertos básicos) de radio $\frac{i}{n}$ que denotaremos $B_{n,i}$ con $i = 1, 2, \dots$, las cuales son una cubierta para E . Para cada $n \in \mathbb{N}$, escogemos i_n de manera que

$$\mathbb{P}\left(\bigcup_{i \leq i_n} B_{n,i}\right) > 1 - \frac{\epsilon}{2^n}$$

Por la hipótesis de completitud el conjunto

$$\bigcap_{n \geq 1} \bigcup_{i \leq i_n} B_{n,i}$$

que es totalmente acotado – es decir lo podemos cubrir con una cantidad finita de abiertos (básicos) de diámetro fijo – tiene cerradura compacta K . Como claramente $\mathbb{P}(K) > 1 - \epsilon$, con esto concluimos la prueba. \square

La definición dada al inicio, de una medida de probabilidad tensa, puede extenderse a una familia de medidas de probabilidad de una manera bastante natural.

Definición 33. (Familia Tensa de Medidas de Probabilidad)

Una familia \mathfrak{T} de medidas de probabilidad sobre un espacio métrico E es tensa, si para cada $\epsilon > 0$ existe un compacto K tal que $\mathbb{P}(K) > 1 - \epsilon$ para todo $\mathbb{P} \in \mathfrak{T}$.

Nuestro objetivo al introducir la noción de tensión de una colección de medidas de probabilidad, es para que a partir de esta propiedad, podamos decir con certeza cuando dicha colección será relativamente compacta y viceversa. El Teorema de Prohorov, extiende la suficiencia –de que una familia de medidas de probabilidad sea tensa– para espacios métricos arbitrarios y la necesidad –de la compacidad relativa– para espacios que además sean completos y separables.

Teorema 34. (Teorema de Prohorov)

Sea \mathfrak{T} una colección de medidas de probabilidad sobre (E, d) , un espacio métrico arbitrario. Si la colección \mathfrak{T} es tensa, entonces es relativamente compacta. Si además E es separable y completo, también tendremos que si \mathfrak{T} es relativamente compacta, entonces es tensa.

Nos referimos a la ida del Teorema de Prohorov como el *Teorema Directo* y al regreso -el cual requiere hipótesis adicionales- como el *Teorema Recíproco*.

El Teorema Directo

Antes de comenzar la prueba consideremos a (E, d) un espacio métrico, τ_d la topología inducida a partir de la métrica d y \mathcal{B}_E la σ -álgebra de Borel generada a partir de la topología τ_d .

Supongamos que $E_0 \in \mathcal{B}_E$, entonces $E_0 \subseteq E$ y es posible considerar a E_0 como subespacio topológico de (E, τ_d) con su respectiva topología relativa $\tau_{d|E_0}$. Se sigue que

$$\mathcal{B}_{E_0} = \{A : A \subset E_0, A \in \mathcal{B}_E\},$$

y en particular tendremos que $\mathcal{B}_{E_0} \subset \mathcal{B}_E$.

Si \mathbb{P} es una medida de probabilidad sobre el espacio medible (E, \mathcal{B}_E) y $\mathbb{P}(E_0) = 1$, denotamos por \mathbb{P}^r a la medida de probabilidad sobre el espacio medible (E_0, \mathcal{B}_{E_0}) obtenida por la restricción de \mathcal{B}_E a \mathcal{B}_{E_0} .

Si \mathbb{P} es una medida de probabilidad sobre el espacio medible (E_0, \mathcal{B}_{E_0}) , denotamos por \mathbb{P}^e a la medida de probabilidad sobre el espacio medible (E, \mathcal{B}_E) obtenida por la extensión de la medida \mathbb{P} a \mathcal{B}_E , definida por

$$\mathbb{P}^e(A) = \mathbb{P}(A \cap E_0) = \mathbb{P}(A|E_0) \quad \text{para toda } A \in \mathcal{B}_E.$$

Observemos que el restringir e inmediatamente extender una medida de probabilidad o viceversa, en cierta manera es llevar a cabo operaciones inversas. Es decir:

Si \mathbb{P} es una medida de probabilidad sobre (E, \mathcal{B}_E) con $\mathbb{P}(E_0) = 1$, entonces

$$(\mathbb{P}^r)^e = \mathbb{P},$$

y si \mathbb{P} es una medida de probabilidad sobre (E_0, \mathcal{B}_{E_0}) , entonces

$$(\mathbb{P}^e)^r = \mathbb{P}.$$

Presentamos los siguientes resultados que utilizaremos en la prueba.

Lema 35.

Si \mathfrak{T} es una familia tensa sobre (E, \mathcal{B}_E) y $h : E \rightarrow E'$ es una función continua, entonces $\{\mathbb{P} \circ h^{-1} : \mathbb{P} \in \mathfrak{T}\}$ es una familia tensa sobre $(E', \mathcal{B}_{E'})$.

Demostración. Dada una $\epsilon > 0$, escogemos en E un conjunto compacto K tal que $\mathbb{P}(K) > 1 - \epsilon$ para toda $\mathbb{P} \in \mathfrak{T}$. Sea $K' = h(K)$, entonces K' es compacto y $h^{-1}(K') \supset K$, por lo que $\mathbb{P} \circ h^{-1}(K') > 1 - \epsilon$ para toda $\mathbb{P} \in \mathfrak{T}$. \square

Si ahora consideramos a la función $h : E_0 \rightarrow E$ como la función identidad, obtenemos directamente del lema anterior y el Teorema (25) el siguiente resultado.

Lema 36.

Si \mathfrak{T} es una familia tensa sobre (E_0, \mathcal{B}_{E_0}) y $h \equiv Id$, entonces $\mathfrak{T}^e = \{\mathbb{P}^e : \mathbb{P} \in \mathfrak{T}\}$ es una familia tensa sobre (E, \mathcal{B}_E) .

Además, si $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ sobre (E_0, \mathcal{B}_{E_0}) entonces $\mathbb{P}_n^e \xrightarrow{w} \mathbb{P}^e$ sobre (E, \mathcal{B}_E) .

Lema 37.

Si $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ sobre (E, \mathcal{B}_E) y $\mathbb{P}(E_0) = \mathbb{P}_n(E_0) = 1$, entonces $\mathbb{P}_n^r \xrightarrow{w} \mathbb{P}^r$ sobre (E_0, \mathcal{B}_{E_0}) .

Demostración. La topología relativa de E_0 como subconjunto de E , nos define la estructura de los abiertos en E_0 , por lo que todo abierto en E_0 es posible expresarlo como $G_0 = G \cap E_0$ donde G un abierto en E . Ya que $\mathbb{P}_n^r(G_0) = \mathbb{P}_n(G)$ y $\mathbb{P}^r(G_0) = \mathbb{P}(G)$, $\liminf_{n \rightsquigarrow \infty} \mathbb{P}_n(G) \geq \mathbb{P}(G)$ implica

$$\liminf_{n \rightsquigarrow \infty} \mathbb{P}_n^r(G_0) \geq \mathbb{P}^r(G_0).$$

□

Un soporte de una medida de probabilidad \mathbb{P} sobre (E, \mathcal{B}_E) , es un conjunto $E_0 \in \mathcal{B}_E$ con $\mathbb{P}(E_0) = 1$; Un conjunto es σ -compacto si se puede representar como unión numerable de conjuntos compactos.

Ahora podemos dar comienzo con la prueba del Teorema Directo. Asumiremos la veracidad del Teorema de Prohorov para el caso $E = \mathbb{R}^\infty$, (Véase [1] pág. 38).

Demostración. (Teorema Directo)

Sea E un espacio métrico arbitrario, si $E_0 = \bigcup_m K_m$, donde K_m es un conjunto compacto en E tal que $\mathbb{P}(K_m) > 1 - 1/m$ para todo $\mathbb{P} \in \mathfrak{T}$, entonces E_0 es soporte de cada elemento de \mathfrak{T} y $\mathfrak{T}^r = \{\mathbb{P}^r : \mathbb{P} \in \mathfrak{T}\}$ es tensa en (E_0, \mathcal{B}_{E_0}) . Como tenemos que E_0 es además un espacio σ -compacto, entonces es separable y puede ser encajado homeomórficamente en un subconjunto de \mathbb{R}^∞ . Sea h dicho homeomorfismo, por el Lema 35 la familia $\mathfrak{T}_h^r = \{\mathbb{P} \circ h^{-1} : \mathbb{P} \in \mathfrak{T}^r\}$ es tensa en \mathbb{R}^∞ y la imagen de E_0 bajo dicho homeomorfismo h será de igual manera σ -compacta; en particular esta imagen es un boreliano de \mathbb{R}^∞ . Por el Teorema 25 la convergencia débil se preserva bajo homeomorfismos, y por lo tanto la compacidad relativa de \mathfrak{T}_h^r . Entonces, para cada sucesión $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ en \mathfrak{T} , la correspondiente sucesión $\{\mathbb{P}_n^r\}_{n \in \mathbb{N}}$ en \mathfrak{T}^r , contiene una subsucesión $\{\mathbb{P}_{n'}^r\}_{n' \in \mathbb{N}}$ que converge débilmente en el espacio (E_0, \mathcal{B}_{E_0}) hacia una medida \mathbb{Q} . Por el Lema 36 y ya que al extender una medida restringida obtenemos la

medida original, se sigue que $\mathbb{P}_{n'} \xrightarrow{w} \mathbb{Q}^e$. Por lo tanto \mathfrak{T} es relativamente compacta. \square

El Teorema Recíproco.

Notemos que para el caso en el que la familia consta de una sola medida, la prueba se reduce al Teorema 32; de hecho, la prueba es un refinamiento de esta última.

Demostración. (Teorema Recíproco)

Supongamos que para cada real positivo ϵ y δ existe una colección finita A_1, \dots, A_n de bolas abiertas de radio δ tal que:

$$\mathbb{P}\left(\bigcup_{i \leq n} A_i\right) > 1 - \epsilon \text{ para toda } \mathbb{P} \in \mathfrak{T}.$$

El siguiente argumento muestra que la familia \mathfrak{T} es tensa. Dada $\epsilon > 0$, escogemos, para cada entero k , una cantidad finita de bolas abiertas de radio $1/k$, $A_{k,1}, \dots, A_{k,n_k}$ tal que:

$$\mathbb{P}\left(\bigcup_{i \leq n_k} A_{k,i}\right) > 1 - \frac{\epsilon}{2^k} \text{ para toda } \mathbb{P} \in \mathfrak{T}.$$

Si K es la cerradura del conjunto $\bigcap_{k \geq 1} \bigcup_{i \leq n_k} A_{k,i}$, entonces $\mathbb{P}(K) > 1 - \epsilon$, y ya que E es completo, K es compacto.

Probamos ahora el teorema mostrando que, si la condición expuesta en el párrafo precedente falla, entonces \mathfrak{T} no es relativamente compacta. Es decir probaremos la negación de la implicación del Teorema Recíproco. Supongamos que para cada real positivo ϵ y δ existe una colección finita A_1, \dots, A_n de bolas abiertas de radio δ tal que:

$$\mathbb{P}\left(\bigcup_{i \leq n} A_i\right) \leq 1 - \epsilon \text{ para alguna } \mathbb{P} \in \mathfrak{T}.$$

Como E es separable, es unión numerable de bolas abiertas A_1, A_2, A_3, \dots de radio δ . Sea $B_n = \bigcup_{i \leq n} A_i$ y escojamos $\mathbb{P}_n \in \mathfrak{T}$ de manera que $\mathbb{P}_n(B_n) < 1 - \epsilon$. Supongamos que para esta sucesión $\{\mathbb{P}_n\}$, existe alguna subsucesión $\{\mathbb{P}_{n'}\}$ que converge débilmente a un límite \mathbb{Q} . Ya que B_m es abierto tendríamos que $\mathbb{P}(B_m) \leq \liminf_{n' \rightsquigarrow \infty} \mathbb{P}_{n'}(B_m)$ para cada m fija. Pero entonces, como $B_m \subset B_{n'}$ para n' suficientemente grande, $\mathbb{P}(B_m) \leq \liminf_{n' \rightsquigarrow \infty} \mathbb{P}_{n'}(B_{n'}) \leq 1 - \epsilon$ se seguiría; como B_m es creciente hacia el espacio E , esto no es posible, completando con esto la prueba. \square

Podemos parafrasear los resultados de la presente sección a una colección de elementos aleatorios de un espacio métrico E .

Definición 38. (Sucesión Tensa de Elementos Aleatorios)

Si X_1, X_2, \dots son elementos aleatorios de un espacio métrico (E, d) , decimos que $\{X_n\}_{n \in \mathbb{N}}$ es tensa cuando la sucesión $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ es tensa, donde \mathbb{P}_n es la distribución de X_n .

Además, en el caso $E = \mathbb{R}^\infty$ ó $C([0, 1], \mathbb{R})$, identificamos las distribuciones finito dimensionales de X_n con aquellas de \mathbb{P}_n .

Capítulo 3

El Espacio $C([0, 1], \mathbb{R})$

A lo largo de este capítulo, denotamos $C = C([0, 1], \mathbb{R})$, donde $C([0, 1], \mathbb{R})$ es el espacio de las funciones reales, continuas y acotadas definidas sobre el intervalo $[0, 1]$. Es bien sabido que las funciones continuas definidas sobre un compacto son uniformemente continuas y que la imagen continua de un espacio compacto es nuevamente compacto y por lo tanto acotada. Por lo que bien podríamos haber dicho que C consta de todas las funciones reales acotadas e uniformemente continuas.

Consideramos a este espacio con la topología de la convergencia uniforme. Para una mejor comprensión de este espacio topológico, damos una breve presentación sobre el origen de esta topología.

Consideremos al espacio métrico (\mathbb{R}, e) donde e denota a la métrica Euclidiana en \mathbb{R} , definida para $x, y \in \mathbb{R}$ como:

$$e(x, y) = |x - y|,$$

y denotamos por $\mathbb{R}^{[0,1]}$ al conjunto de todas las posibles funciones de $[0, 1]$ en \mathbb{R} . Para $f \in \mathbb{R}^{[0,1]}$ y cada $\epsilon > 0$, sea $V(f, \epsilon)$ la colección de funciones $g \in \mathbb{R}^{[0,1]}$ cuya gráfica G_g definida por:

$$G_g = \{(x, g(x)) : x \in [0, 1]\} \subset [0, 1] \times \mathbb{R},$$

no se separa de la gráfica de f en más de una cantidad $0 < \delta < \epsilon$; es decir,

$$V(f, \epsilon) = \{g \in C : \exists \delta < \epsilon, |f(x) - g(x)| < \delta \text{ para toda } x \in [0, 1]\}.$$

La colección $\mathcal{B} = \{V(f, \epsilon) : f \in \mathbb{R}^{[0,1]}, \epsilon > 0\}$ forma una base para una topología en $\mathbb{R}^{[0,1]}$. A esta topología en $\mathbb{R}^{[0,1]}$ y a su restricción en C , la denotaremos indistintamente por τ_e y le llamamos la topología de la convergencia uniforme (con respecto a la métrica e).

Como es usual, diremos que dos métricas en el espacio \mathbb{R} son equivalentes si inducen la misma topología; y una métrica d en el espacio (\mathbb{R}, τ) es compatible si la topología generada por d en \mathbb{R} coincide con la topología τ . Una métrica d en \mathbb{R} es acotada si existe un número natural N , tal que

$$d(x, y) < N \quad \text{para cualesquiera dos puntos } x, y \in \mathbb{R}.$$

En particular, para la métrica Euclidiana e en \mathbb{R} , la métrica $e^* : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ definida por

$$e^*(x, y) = \min\{1, e(x, y)\},$$

es una métrica acotada en \mathbb{R} y equivalente a la métrica e .

Como e^* es acotada, para el espacio métrico (\mathbb{R}, e^*) , es posible definir la función

$$\hat{e}^* : \mathbb{R}^{[0,1]} \times \mathbb{R}^{[0,1]} \rightarrow \mathbb{R}_+$$

como

$$\hat{e}^*(f, g) = \sup_{x \in [0,1]} \{e^*(f(x), g(x))\}.$$

Además, tendremos que la métrica \hat{e}^* en $\mathbb{R}^{[0,1]}$ es compatible con τ_e . Más aún, gracias a un resultado topológico conocido, tenemos que como el intervalo $[0, 1]$ es compacto y las métricas e^* y e son equivalentes en \mathbb{R} , entonces \hat{e}^* y \hat{e} generan ambas a la topología uniforme τ_e en C . Así pues, escribimos ρ para denotar a la métrica \hat{e}^* en el espacio C y bajo la cual, la distancia entre dos elementos $X = (X(t), t \in [0, 1])$ y $Y = (Y(t), t \in [0, 1])$ de C se define como:

$$\rho(X, Y) = \sup_{t \in [0,1]} |X(t) - Y(t)|.$$

De la discusión anterior concluimos, que gracias a que (\mathbb{R}, e) es un espacio métrico, entonces $(\mathbb{R}^{[0,1]}, \tau_e)$ es metrizable y por consiguiente (C, ρ) es un espacio métrico.

El espacio (C, ρ) es separable; un subconjunto denso numerable es aquel que consiste de todas las funciones poligonales lineales sobre los intervalos $[(i-1)/n, i/n]$ con $1 \leq i \leq n$ para algún $n \in \mathbb{N}$, y que toma valores racionales en i/n donde $0 \leq i \leq n$.

El espacio (C, ρ) es completo; si $\{X_n\}_{n \in \mathbb{N}}$ es una sucesión de Cauchy en C , entonces,

$$\rho(X_n, X_m) < \epsilon \quad \text{si } n, m \geq N \quad (3.1)$$

por lo que para cada $0 \leq s \leq 1$ fija, se tiene que

$$|X_n(s) - X_m(s)| < \sup_t |X_n(t) - X_m(t)| < \epsilon \quad \text{si } n, m \geq N$$

esto es, que para cada $0 \leq t \leq 1$ la sucesión de reales $\{X_n(t)\}_{n \in \mathbb{N}}$ es de Cauchy en \mathbb{R} y por completitud de \mathbb{R} es convergente a X donde $X(t) = \lim_{n \rightarrow \infty} X_n(t)$ para toda $t \in [0, 1]$.

Claramente, se tiene que para cada $0 \leq t \leq 1$,

$$||X_n(t) - X(t)| - |X_m(t) - X(t)|| \leq \rho(X_n - X, X_m - X).$$

Por (3.1) y haciendo $m \rightarrow \infty$ tendremos que

$$|X_n(t) - X(t)| \leq \sup_t |X_n(t) - X(t)| \leq \epsilon \quad \text{para toda } n \geq N$$

Esto muestra que la sucesión $\{X_n(t)\}_{n \in \mathbb{N}}$ converge uniformemente a $X(t)$ en $[0, 1]$. Luego, como para toda $n \in \mathbb{N}$ las funciones X_n son continuas en $[0, 1]$ y la convergencia es uniforme, la función límite X es continua en $[0, 1]$. Así $X \in C$ y $\lim_{n \rightarrow \infty} X_n = X$.

Esto último nos permite obtener el siguiente resultado, el cual justifica el nombre de topología uniforme en C .

Proposición 39. (Convergencia Uniforme en C)

Una sucesión $\{X_n\}_{n \in \mathbb{N}}$ en C converge uniformemente a X si y solo si $\{X_n\}_{n \in \mathbb{N}}$ es ρ -Cauchy y converge puntualmente a X .

Por lo tanto, la métrica ρ describe la convergencia uniforme sobre $[0, 1]$, es por esto que es también conocida como la métrica uniforme.

En conclusión, tendremos que (C, ρ) es un espacio métrico, completo y separable.

3.1. Convergencia Débil y Tensión en C

Para cada entero no-negativo k , denotamos por \mathcal{R}^k a la σ -álgebra de Borel de \mathbb{R}^k y por \mathcal{C} a la σ -álgebra de Borel del espacio (C, ρ) . Para un entero k y puntos

t_1, t_2, \dots, t_k en $[0, 1]$, consideremos las proyecciones

$$\begin{aligned} \pi_{t_1 \dots t_k} &: C \longmapsto \mathbb{R}^k \\ x &\longmapsto (x(t_1), x(t_2), \dots, x(t_k)). \end{aligned}$$

Note que las proyecciones son funciones continuas. Además, una función $h : E \rightarrow C$ es continua si y solo si la composición $\pi_{t_1 \dots t_k} \circ h : E \rightarrow \mathbb{R}^k$ lo es para cada k .

Definición 40. (Cilindros)

Los cilindros o conjuntos finito-dimensionales se definen en C como

$$\pi_{t_1 \dots t_k}^{-1}(H) \quad \text{donde } H \in \mathcal{R}^k.$$

Denotaremos por \mathcal{F} a la clase de conjuntos finito-dimensionales o cilindros en C

Ya que cada proyección es continua, es medible y los cilindros recaen sobre \mathcal{C} ($\mathcal{F} \subset \mathcal{C}$). Por otro lado, para $\epsilon > 0$ y un elemento $x \in C$, la cerradura de las bolas abiertas son de la siguiente forma:

$$\bar{B}_\epsilon(x) = \{y : \rho(x, y) \leq \epsilon\} = \bigcap_{n=1}^{\infty} \{y : |x^{(i/n)} - y^{(i/n)}| \leq \epsilon, i = 1, 2, \dots, n\};$$

donde cada uno de los interseccionados es un conjunto finito-dimensional.

Recordemos la definición de Clase Determinante y Clase que Determina la Convergencia dadas con anterioridad.

Ejemplo 1. Como C es separable, cada abierto es unión contable de bolas abiertas y por lo tanto de bolas cerradas, de manera que la clase de los cilindros es una clase que genera a \mathcal{C} . Además, ya que \mathcal{F} es un álgebra de conjuntos, se tiene que los cilindros conforman una Clase Determinante.

Ejemplo 2. La clase \mathcal{F} de cilindros en C , no es una clase que determina la convergencia.

Definición 41. (Distribuciones Finito Dimensionales en C)

Las distribuciones finito dimensionales de una medida de probabilidad \mathbb{P} en (C, \mathcal{C}) , se definen como las medidas $\mathbb{P} \circ \pi_{t_1 \dots t_k}^{-1}$, dónde $\pi_{t_1 \dots t_k}$ son las proyecciones antes mencionadas.

Por la continuidad de las proyecciones, la convergencia débil de las medidas de probabilidad \mathbb{P}_n en (C, \mathcal{C}) , implica la convergencia de las respectivas distribuciones

finito dimensionales $\mathbb{P}_n \circ \pi_{t_1 \dots t_k}^{-1}$; el regreso de esta afirmación es falso ya que \mathcal{F} no es una clase que determina la convergencia. Por otro lado, ya que \mathcal{F} es una clase-determinante, tenemos que una medida de probabilidad en (C, \mathcal{C}) queda únicamente determinada por sus distribuciones finito dimensionales.

Debido a que el espacio C es un espacio completo y separable, del Teorema de Prohorov se deduce el siguiente resultado:

Proposición 42.

Sean \mathbb{P}_n medidas de probabilidad sobre el espacio medible (C, \mathcal{C}) . Si la colección $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ es relativamente compacta y las distribuciones finito dimensionales de \mathbb{P}_n convergen a las de una medida de probabilidad \mathbb{P} , entonces $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$.

Demostración. Como la colección $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ es relativamente compacta, entonces toda subsucesión $\{\mathbb{P}_{n'}\}_{n' \in \mathbb{N}}$ contiene otra subsucesión $\{\mathbb{P}_{n''}\}_{n'' \in \mathbb{N}}$ que converge débilmente a un límite \mathbb{Q} . Las distribuciones finito dimensionales de \mathbb{Q} deben ser entonces límite débil de las de $\{\mathbb{P}_{n''}\}_{n'' \in \mathbb{N}}$ y por lo tanto coincidir con las distribuciones finito dimensionales de \mathbb{P} . Es decir, para cada entero no negativo k y puntos $(t_1, t_2, \dots, t_k) \in [0, 1]^k$,

$$\mathbb{P}_{n''} \circ \pi_{t_1 \dots t_k}^{-1} \xrightarrow{w} \mathbb{Q} \circ \pi_{t_1 \dots t_k}^{-1} \quad \text{y} \quad \mathbb{P}_n \circ \pi_{t_1 \dots t_k}^{-1} \xrightarrow{w} \mathbb{P} \circ \pi_{t_1 \dots t_k}^{-1},$$

por lo que $\mathbb{P} \circ \pi_{t_1 \dots t_k}^{-1} = \mathbb{Q} \circ \pi_{t_1 \dots t_k}^{-1}$. Pero como toda medida de probabilidad en C queda completamente determinada por sus distribuciones finito dimensionales, \mathbb{P} y \mathbb{Q} deben de coincidir. Entonces cada subsucesión $\{\mathbb{P}_{n'}\}_{n' \in \mathbb{N}}$ contiene otra subsucesión $\{\mathbb{P}_{n''}\}_{n'' \in \mathbb{N}}$ que converge débilmente a \mathbb{P} , y por el Teorema 17 la sucesión completa $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ converge débilmente a \mathbb{P} . \square

No está de mas observar que si de inicio $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ converge débilmente a \mathbb{P} , entonces es relativamente compacta, de manera que no es un requerimiento excesivo para pedir la convergencia.

Supongamos que $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ es relativamente compacto y que para cada vector (t_1, t_2, \dots, t_k) , $\mathbb{P}_n \circ \pi_{t_1 \dots t_k}^{-1}$ converge débilmente a alguna medida de probabilidad $\mathbb{Q}_{t_1 \dots t_k}^{-1}$ sobre $(\mathbb{R}^k, \mathcal{R}^k)$ —el punto por el momento es no asumir que $\mathbb{Q}_{t_1 \dots t_k}^{-1}$ es una distribución finito dimensional de una medida de probabilidad en (C, \mathcal{C}) . Se seguirá de igual manera que para cada subsucesión $\{\mathbb{P}_{n'}\}_{n' \in \mathbb{N}}$ existe otra subsucesión $\{\mathbb{P}_{n''}\}_{n'' \in \mathbb{N}}$ que converge débilmente a algún límite. Ya que este límite debe tener a las $\mathbb{Q}_{t_1 \dots t_k}^{-1}$ como sus distribuciones finito dimensionales, esta medida queda únicamente determinada. Por lo que $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ converge débilmente a alguna medida \mathbb{P} .

Estas ideas proveen de una técnica poderosa para probar convergencia débil sobre (C, \mathcal{C}) . Uno primero prueba que las distribuciones finito dimensionales convergen débilmente y posteriormente que la sucesión es relativamente compacta.

Además, gracias al Teorema de Prohorov, vimos que un criterio efectivo para probar la compacidad relativa, es que dicha sucesión sea tensa; de hecho, la compacidad relativa de un familia de medidas de probabilidad en el espacio (C, \mathcal{C}) es equivalente a la tensión de dicha familia.

Recordemos que si X_n son elementos aleatorios del espacio C , decimos que la colección $\{X_n\}_{n \in \mathbb{N}}$ es tensa si la colección $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ de sus correspondientes distribuciones lo es. Además identificamos a las distribuciones finito dimensionales de X_n con las de \mathbb{P}_n ; con esto la proposición anterior puede interpretarse de la siguiente manera:

Proposición 43.

Sean X_1, X_2, X_3, \dots elementos aleatorios del espacio medible (C, \mathcal{C}) . Si la colección $\{X_n\}_{n \in \mathbb{N}}$ es tensa y las distribuciones finito dimensionales de X_n convergen a las de X , entonces $X_n \xrightarrow{\mathcal{D}} X$.

3.2. Funciones Aleatorias

Sea X una función de un espacio de probabilidad $(\Omega, \mathcal{B}, \mathbb{P})$ en C . Por el momento no asumimos que X sea medible ($X^{-1}(\mathcal{C}) \subseteq \mathcal{B}$). Para cada $\omega \in \Omega$, $X(\omega)$ es un elemento en C – una función continua en $[0, 1]$ cuyo valor en un punto t de dicho intervalo lo denotamos por $X_t(\omega) = X(t, \omega)$. Dada una t fija ($0 \leq t \leq 1$), denotamos por X_t a la función real sobre Ω con valor $X_t(\omega)$ en ω

$$\begin{aligned} X_t &: \Omega \longmapsto \mathbb{R} \\ &\omega \longmapsto X_t(\omega). \end{aligned}$$

Notemos que X_t es simplemente la composición $\pi_t \circ X$. Con lo anterior, obtenemos el siguiente diagrama conmutativo

$$\begin{array}{ccc} \Omega & \xrightarrow{X} & C \\ X_t \searrow & & \swarrow \pi_t \\ & \mathbb{R} & \end{array}$$

De manera similar, denotamos $(X_{t_1}, X_{t_2}, \dots, X_{t_k})$ a la función

$$\begin{aligned} (X_{t_1}, X_{t_2}, \dots, X_{t_k}) &: \Omega \mapsto \mathbb{R}^k \\ \omega &\mapsto (X_{t_1}(\omega), X_{t_2}(\omega), \dots, X_{t_k}(\omega)) \end{aligned}$$

Si tenemos que $A = \{x \in C : x(t) \leq \alpha\}$, entonces

$$A \in \mathcal{C} \quad \text{y} \quad X^{-1}(A) = \{\omega : X_t(\omega) \leq \alpha\}.$$

Se sigue que si X es un elemento aleatorio de C ($X^{-1}(C) \subseteq \mathcal{B}$) entonces cada X_t es una variable aleatoria ($X_t^{-1}(\mathcal{B}_{\mathbb{R}}) \subseteq \mathcal{B}$) y en consecuencia cada $(X_{t_1}, X_{t_2}, \dots, X_{t_k})$ es un vector aleatorio.

Por otro lado, supongamos que cada X_t es una variable aleatoria. Denotamos por $\bar{B}_\delta(X(\omega_0))$ a la bola cerrada en C de radio $\delta > 0$ y centrada en $X(\omega_0) = x \in C$. Se tiene entonces que:

$$\begin{aligned} X^{-1}(\bar{B}_\delta(x)) &= \{\omega : X(\omega) \in \bar{B}_\delta(x)\} \\ &= \{\omega : \rho(X(\omega), x) \leq \delta\} \\ &= \{\omega : x(t) - \delta \leq X(t, \omega) \leq \delta + x(t) \text{ para toda } t\} \\ &= \bigcap_{t \in [0,1] \cap \mathbb{Q}} \{\omega : x(t) - \delta \leq X(t, \omega) \leq \delta + x(t)\}, \end{aligned}$$

de manera que $X^{-1}(\bar{B}_\delta(x)) \in \mathcal{B}$. Como C es separable, se sigue que $X^{-1}(C) \subseteq \mathcal{B}$. Por lo tanto, X es un elemento aleatorio de C . Con esto hemos mostrado que X es una función aleatoria (elemento aleatorio de C) si y solo si cada X_t es una variable aleatoria.

En lo que resta que del capítulo tendremos principal interés en funciones aleatorias construidas de la siguiente manera. Sean ξ_1, ξ_2, \dots , variables aleatorias definidas sobre un espacio de probabilidad $(\Omega, \mathcal{B}, \mathbb{P})$. Por el momento no es necesario que este proceso tenga alguna propiedad especial como el ser estacionario o independencia de las variables. Definimos

$$S_n = \sum_{i=1}^n \xi_i \quad \text{con} \quad S_0 = 0$$

y construimos a X_n a partir de las sumas parciales S_0, S_1, \dots, S_n ; para puntos $\frac{i}{n} \in [0, 1]$ con $0 \leq i \leq n$, hacemos

$$X_n\left(\frac{i}{n}, \omega\right) = \frac{1}{\sigma\sqrt{n}} S_i(\omega).$$

Nos concentraremos en factores que normalizen la expresión de la forma $\sigma\sqrt{n}$.

Para los puntos restantes $t \in [0, 1]$, definimos a $X_n(t, \omega)$ por interpolación lineal: Si $t \in \left[\frac{(i-1)}{n}, \frac{i}{n}\right]$, entonces

$$\begin{aligned} X_n(t, \omega) &= \frac{\frac{i}{n} - t}{\frac{1}{n}} X_n\left(\frac{i-1}{n}, \omega\right) + \frac{t - \frac{(i-1)}{n}}{\frac{1}{n}} X_n\left(\frac{i}{n}, \omega\right) \\ &= \frac{1}{\sigma\sqrt{n}} S_{i-1}(\omega) + n\left(t - \frac{i-1}{n}\right) \frac{1}{\sigma\sqrt{n}} \xi_i(\omega). \end{aligned}$$

como $i-1 = [nt]$ si $t \in \left[\frac{(i-1)}{n}, \frac{i}{n}\right]$, podemos definir mas concisamente la función por

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) + (nt - [nt]) \frac{1}{\sigma\sqrt{n}} \xi_{[nt]+1}(\omega). \quad (3.2)$$

Como las ξ_i y S_i son variables aleatorias, tenemos por (3.2) que $X_n(t)$ es una variable aleatoria para cada t . Por lo tanto, las X_n son funciones aleatorias.

Una pregunta natural es la siguiente: ¿Cuándo este tipo de funciones conforman una sucesión tensa? En primera instancia tenemos el siguiente resultado cuya prueba puede ser encontrada en [1] Teorema 8.3.

Teorema 44. *Una sucesión de medidas $\{\mathbb{P}_n\}_{n \in \mathbb{N}}$ es tensa si las siguientes condiciones se satisfacen:*

(i) *Para cada $\eta > 0$, existe a tal que*

$$\mathbb{P}_n\{x : |x(0)| > a\} \leq \eta, \quad n \geq 1.$$

(ii) *Para cada $\epsilon, \eta > 0$, existen $\delta \in (0, 1)$ y un entero N_0 tal que para toda $t \in [0, 1]$,*

$$\frac{1}{\delta} \mathbb{P}_n \left\{ x : \sup_{t \leq s \leq \min\{t+\delta, 1\}} |x(s) - x(t)| \geq \epsilon \right\} \leq \eta, \quad n \geq N_0.$$

Sin embargo, al momento de considerar una sucesión de funciones aleatorias $\{X_n\}_{n \in \mathbb{N}}$ sabemos que esta sucesión es por definición tensa cuando la sucesión conformada por sus respectivas distribuciones es tensa. Luego es posible enunciar dicho teorema en términos de la sucesión $\{X_n\}_{n \in \mathbb{N}}$ de la siguiente manera:

La sucesión $\{X_n\}_{n \in \mathbb{N}}$ es tensa si $\{X_n(0)\}_{n \in \mathbb{N}}$ es tensa y para cada $\epsilon, \eta > 0$, existe $\delta \in (0, 1)$ y un entero N_0 tal que para $n \geq N_0$ y $t \in [0, 1]$

$$\frac{1}{\delta} \mathbb{P} \left\{ \sup_{t \leq s \leq \min\{t+\delta, 1\}} |X_n(s) - X_n(t)| \geq \epsilon \right\} \leq \eta, \quad n \geq N_0.$$

Así las cosas, utilizando la definición de las funciones dadas por (3.2) es posible traducir este último resultado restringiendo nuestra atención sobre la manera en que se comportan las fluctuaciones de las sumas parciales S_i . Si $t = k/n$ y $t + \delta = j/n$, con $k, j \in \mathbb{N}$, entonces la condición anterior se reduce a

$$\frac{1}{\delta} \mathbb{P} \left\{ \max_{i \leq n\delta} \frac{1}{\sigma\sqrt{n}} |S_{k+i} - S_k| \geq \epsilon \right\} \leq \eta.$$

Teorema 45.

Sea $\{X_n\}_{n \in \mathbb{N}}$ una sucesión de funciones aleatorias definidas como en (3.2). La sucesión $\{X_n\}_{n \in \mathbb{N}}$ es tensa si para cada $\epsilon > 0$ existe un $\lambda > 1$ y un entero positivo N_0 , tal que si $n \geq N_0$ y $k \in \mathbb{N}$, entonces

$$\mathbb{P} \left\{ \max_{i \leq n} |S_{k+i} - S_k| \geq \lambda\sigma\sqrt{n} \right\} \leq \frac{\epsilon}{\lambda^2}.$$

Requerimos que $\lambda > 1$. Obsérvese que el resultado es trivial si $\lambda \leq \sqrt{\epsilon}$.

Demostración. Dadas ϵ y η , encontraremos $\delta \in (0, 1)$ y un entero positivo N_0 para el cual, si $n \geq N_0$ y $k \in \mathbb{N}$

$$\frac{1}{\delta} \mathbb{P} \left\{ \max_{i \leq n\delta} \frac{1}{\sigma\sqrt{n}} |S_{k+i} - S_k| \geq \epsilon \right\} \leq \eta. \quad (3.3)$$

Ya que (3.3) se vuelve una desigualdad más estricta conforme decrecen ϵ y η -debido a que heurísticamente estamos pidiendo que conforme el respectivo evento crece su probabilidad decrece- sin pérdida de generalidad supondremos que $\epsilon, \eta \in (0, 1)$.

Por la hipótesis, reemplazando ϵ por $\eta\epsilon^2$, tenemos que existe $\lambda > 1$ y un entero positivo N_1 , tal que si $n \geq N_1$ y $k \geq 1$,

$$\mathbb{P} \left\{ \max_{i \leq n} |S_{k+i} - S_k| \geq \lambda\sigma\sqrt{n} \right\} \leq \frac{\eta\epsilon^2}{\lambda^2}. \quad (3.4)$$

Proponemos $\delta = \epsilon^2/\lambda^2$; como $\lambda > 1 > \epsilon$ tenemos que $0 < \delta < 1$. Sea N_0 un entero por arriba de N_1/δ .

Si $n \geq N_0$, entonces $[n\delta] \geq N_1$, y se sigue de (3.4) que

$$\mathbb{P}\left\{\max_{i \leq [n\delta]} |S_{k+i} - S_k| \geq \lambda\sigma\sqrt{[n\delta]}\right\} \leq \frac{\eta\epsilon^2}{\lambda^2}.$$

Como $\lambda\sqrt{[n\delta]} \leq \epsilon\sqrt{n}$ y también $\eta\frac{\epsilon^2}{\lambda^2} = \eta\delta$, (3.3) es válido para todo k si $n \geq N_0$. \square

Si $\{S_n\}_{n \in \mathbb{N}}$ es un proceso estacionario, el teorema recién probado se reduce a

$$\mathbb{P}\left\{\max_{i \leq n} |S_i| \geq \lambda\sigma\sqrt{n}\right\} \leq \frac{\epsilon}{\lambda^2}.$$

Podemos absorber σ en λ y pedir

$$\mathbb{P}\left\{\max_{i \leq n} |S_i| \geq \lambda\sqrt{n}\right\} \leq \frac{\epsilon}{\lambda^2} \quad \text{con } \lambda > \sigma.$$

3.3. Medida de Wiener y Movimiento Browniano

En la presente sección recordaremos algunas propiedades básicas e importantes del movimiento Browniano. Se enuncian teoremas y proposiciones esenciales que forman parte de una teoría extensa cuyo objetivo difiere del presente trabajo.

Recordemos que una variable aleatoria Normal con parámetros (μ, σ^2) , donde $\mu \in \mathbb{R}$ y $\sigma^2 > 0$ son respectivamente su esperanza y su varianza, es una variable aleatoria cuya densidad está dada por:

$$f_X(t) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(t-\mu)^2/2\sigma^2}, \quad -\infty < t < \infty.$$

Tal distribución se denota usualmente como $N(\mu, \sigma^2)$. Debido al trabajo fundamental del matemático C. F. Gauss, las variables aleatorias normales son también comúnmente llamadas *gaussianas*.

Decimos que X se distribuye como una variable aleatoria gaussiana con media μ y varianza $\sigma^2 > 0$ si su transformada de Fourier (función característica) satisface:

$$\varphi_X(u) := \mathbb{E}[e^{iuX}] = e^{iu\mu - \frac{\sigma^2 u^2}{2}} \quad \text{para toda } u \in \mathbb{R}_+.$$

Teorema 46. Sea $(X_n)_{n \in \mathbb{N}}$ una sucesión de variables aleatorias gaussianas tal que para toda n

$$X_n \stackrel{(d)}{=} N(m_n, \sigma_n^2) \quad \text{donde } m_n \in \mathbb{R}, \sigma_n \geq 0$$

i.e se distribuyen como una normal con media m_n y varianza σ_n^2 .

(i) Si $X_n \xrightarrow{\mathfrak{D}} X$, entonces $X \stackrel{(d)}{=} N(m, \sigma^2)$ donde $m = \lim_{n \rightarrow \infty} m_n$ y $\sigma^2 = \lim_{n \rightarrow \infty} \sigma_n^2$.

(ii) Si $X_n \xrightarrow{\mathbb{P}} X$, entonces para toda $p \geq 1$, $X_n \xrightarrow{\mathcal{L}^p} X$.

Nótese que en el caso degenerado $\delta_\mu = N(\mu, 0)$ (también conocida como la medida de Dirac en el punto μ). Cuando $\mu = 0$ y $\sigma^2 = 1$, decimos que $N(0, 1)$ es la distribución Normal estándar.

Demostración. (i) Como $X_n \xrightarrow{\mathfrak{D}} X$, entonces

$$\varphi_{X_n}(\lambda) = \exp \left\{ im_n \lambda - \lambda^2 \frac{\sigma_n^2}{2} \right\} \rightarrow \varphi_X(\lambda),$$

en particular tendremos que por continuidad del módulo

$$|\varphi_{X_n}(\lambda)| = \exp \left\{ -\lambda^2 \frac{\sigma_n^2}{2} \right\} \rightarrow |\varphi_X(\lambda)|, \quad \text{para toda } \lambda \in \mathbb{R},$$

lo cual implica que $\sigma_n^2 \rightarrow \sigma^2 \in [0, \infty)$. Ahora veamos que $(m_n)_{n \geq 1}$ es acotada. Supongamos que existe una subsucesión $(m_{n_k})_{k \geq 1}$ tal que $m_{n_k} \rightarrow \infty$ y para toda $a \in \mathbb{R}$ consideramos $b > a$ un punto de continuidad de F_X , de función de distribución de X . Claramente

$$F_X(a) \leq F_X(b) = \lim_{k \rightarrow \infty} F_{X_{n_k}}(b) \leq \limsup_{k \rightarrow \infty} F_{X_{n_k}}(m_{n_k}) = \frac{1}{2},$$

lo cual es una contradicción por lo tanto la sucesión $(m_n)_{n \geq 1}$ es acotada. Finalmente supongamos que m, m' son dos puntos de adherencia de dicha sucesión, entonces

$$e^{im\lambda} = e^{im'\lambda}, \quad \text{para toda } \lambda \in \mathbb{R},$$

lo que claramente implica que $m = m'$. Sea $\lim_{n \rightarrow \infty} m_n = m$, la transformada de Fourier de X satisface

$$\varphi_X(\lambda) = \exp \left\{ im\lambda - \frac{\sigma^2}{2} \lambda^2 \right\},$$

por lo tanto X es una variable aleatoria Gaussiana con media m y varianza σ^2 .

(ii) Para probar este inciso, recordemos que la transformada de Laplace de X_n satisface

$$\mathbb{E}[e^{\alpha X_n}] = \exp \left\{ m_n \alpha + \frac{\sigma_n^2}{2} \alpha^2 \right\} \quad \text{para toda } \alpha \in \mathbb{R}.$$

Como $e^{|x|} \leq e^x + e^{-x}$, es claro que

$$\sup_{n \geq 1} \mathbb{E}[e^{\alpha |X_n|}] < +\infty \quad \text{para toda } \alpha \in \mathbb{R}.$$

Luego, para $q \geq 0$ se tiene que $|x|^q \leq e^{\alpha |x|}$ para $|x|$ suficientemente grande de donde deducimos que

$$\sup_{n \geq 1} \mathbb{E}[|X_n|^q] < +\infty$$

y en consecuencia

$$\sup_{n \geq 1} \mathbb{E}[|X_n - X|^q] < +\infty.$$

Sea $p \geq 1$. La sucesión $(Y_n)_{n \in \mathbb{N}}$ con $Y_n = |X_n - X|^p$ converge a cero en probabilidad por hipótesis y es uniformemente integrable ya que está acotada en \mathcal{L}^2 tomando $q = 2p$. Por lo tanto,

$$\mathbb{E}[|X_n - X|^p] \longrightarrow 0.$$

qedhere □

Denotamos por I a un conjunto de índices subconjunto de \mathbb{R}^+ . Al considerar procesos definidos sobre un intervalo $[0, T]$ para $T \geq 0$, entenderemos que $I \subseteq [0, T]$. Sea $(\Omega, \mathcal{B}, \mathbb{P})$ un espacio de probabilidad abstracto.

Definición 47. (Trayectoria)

Para cada $\omega \in \Omega$, la trayectoria del proceso $X = (X_t)_{t \in I}$ correspondiente a un posible resultado ω es la función definida sobre I dada por $t \mapsto X_t(\omega)$.

Definición 48. (Vector Gaussiano)

Un vector aleatorio \mathbb{R}^n -valuado (X_1, \dots, X_n) es Gaussiano, si toda combinación lineal $\sum_{i=1}^n a_i X_i$ tiene distribución (unidimensional) Normal.

Definición 49. (Proceso Gaussiano)

Un proceso $(X_t)_{t \in I}$ es Gaussiano, si para toda n y todo vector $(t_1, \dots, t_n) \in I^n$, tenemos que $(X_{t_1}, \dots, X_{t_n})$ es un vector Gaussiano con valores en \mathbb{R}^n .

Definición 50. (Movimiento Browniano estándar)

Decimos que $B = (B_t)_{t \geq 0}$ es un movimiento Browniano estándar, si B es un proceso Gaussiano real centrado con trayectorias continuas casi seguramente, tal que $B_0 = 0$ y

$$\mathbb{E}[B_s B_t] = s \wedge t \quad \text{con } s, t \geq 0.$$

De manera análoga, es posible definir un movimiento Browniano $B = (B_t)_{t \in [0, T]}$ en un intervalo $[0, T]$, como un proceso Gaussiano centrado con trayectorias continuas casi seguramente y tal que

$$\mathbb{E}[B_s B_t] = s \wedge t \quad \text{con } s, t \in [0, T].$$

La siguiente proposición nos permite determinar cuando un proceso $X = (X_t)_{t \in I}$ es un movimiento Browniano.

Proposición 51. *Un proceso X es un movimiento Browniano si y sólo X tiene trayectorias continuas casi seguramente y satisface:*

- i) $X_0 = 0$ casi seguramente.*
- ii) Para $n \geq 2$ y para toda elección de puntos $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, los incrementos*

$$X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$$

son independientes.

- iii) Para todo $0 \leq s \leq t$, se tiene que $X_t - X_s \stackrel{(d)}{=} N(0, t - s)$.*

OBSERVACIÓN. Es común referirse a la propiedad *ii)* y *iii)* diciendo que el movimiento Browniano tiene incrementos independientes y estacionarios respectivamente.

Demostración. \Leftarrow] Consideramos $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$ y sea

$$(X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}),$$

un vector Gaussiano, entonces $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ también lo es y en consecuencia tendremos que X es un proceso Gaussiano ya que $X_t \stackrel{(d)}{=} N(0, t)$ para toda $t \geq 0$. Sean $s, t \geq 0$ y supongamos sin pérdida de generalidad que $s < t$, entonces X_s y $X_t - X_s$ son variables aleatorias gaussianas independientes, lo cual implica que

$$\mathbb{E}[X_t X_s] = \mathbb{E}[(X_t + X_s - X_s) X_s] = \mathbb{E}[X_s^2] = s.$$

En general $\mathbb{E}[X_t X_s] = s \wedge t$ para toda $s, t \geq 0$, por lo que X es un movimiento Browniano.

\Rightarrow] Supongamos ahora que X es un movimiento Browniano. Como $\mathbb{E}[X_0^2] = 0$, entonces $X_0 = 0$ c.s. Para $0 \leq s \leq t$, $X_t - X_s$ es una variable aleatoria Gaussiana centrada. Por otro lado observemos que

$$\mathbb{E}[(X_t - X_s)^2] = \mathbb{E}[X_t^2] + \mathbb{E}[X_s^2] - 2\mathbb{E}[X_s X_t] = t + s - 2s = t - s,$$

lo cual implica que $X_t - X_s \stackrel{(d)}{=} N(0, t - s)$. Finalmente probemos (ii).

Sean $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, como $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ es un vector Gaussiano entonces

$$(X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}})$$

también lo es y además para toda $i < j$, se tiene que

$$\begin{aligned} \text{Cov}(X_{t_{i+1}} - X_{t_i}, X_{t_{j+1}} - X_{t_j}) &= \mathbb{E}[X_{t_{i+1}} X_{t_{j+1}}] + \mathbb{E}[X_{t_i} X_{t_j}] \\ &\quad - \mathbb{E}[X_{t_{i+1}} X_{t_j}] - \mathbb{E}[X_{t_i} X_{t_{j+1}}] \\ &= t_{i+1} + t_i - t_i - t_{i+1} = 0. \end{aligned}$$

Por lo tanto $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ son independientes. \square

El siguiente teorema debido a Wiener prueba la existencia del movimiento Browniano.

Teorema 52. (Wiener - 1923) *El movimiento Browniano existe.*

Demostración. (Levy - 1948) Comenzamos por la construcción del movimiento Browniano en el intervalo $[0, 1]$. Para esto consideramos una familia de variables aleatorias Gaussianas centradas con varianza unitaria $(\xi_{k,n} : 0 \leq k \leq 2^n \quad n \geq 0)$ y definimos un nuevo proceso $(X_n(t) : t \in [0, 1] \quad n \geq 0)$ de manera recursiva:

1. En el primer paso, hacemos

$$X_0(0) \equiv 0, \quad X_0(1) \equiv \xi_{0,0}$$

y X_0 lineal en el intervalo $[0, 1]$.

2. En el segundo paso, hacemos

$$X_1(0) \equiv 0, \quad X_1(1/2) \equiv X_0(1/2) + \frac{\xi_{1,1}}{2}, \quad X_1(1) \equiv X_0(1)$$

y X_1 lineal en el intervalo $[0, 1/2]$ y $(1/2, 1]$,

y así sucesivamente.

En general tendremos que para toda $n \geq 0$, el mapeo $t \mapsto X_n(t)$ es lineal en cada uno de los intervalos $\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]$ y además

$$X_n\left(\frac{2j}{2^n}\right) \equiv X_{n-1}\left(\frac{2j}{2^{n-1}}\right), \quad X_n\left(\frac{2j+1}{2^n}\right) \equiv X_{n-1}\left(\frac{2j+1}{2^{n-1}}\right) + \frac{\xi_{2j+1,n}}{2^{\frac{n+1}{2}}}.$$

Afirmamos y probamos por inducción lo siguiente. Para toda $n \geq 0$, el vector $(X_n(k/2^n) : 0 \leq k \leq 2^n)$ es Gaussiano, centrado y de covarianza

$$\mathbb{E}[X_n(k/2^n)X_n(l/2^n)] = \frac{k}{2^n} \wedge \frac{l}{2^n}. \quad (3.5)$$

Notemos que el enunciado se cumple para el caso $n = 0$ y supongamos que es válido para los casos $1, \dots, n-1$. Primero observemos que $(X_n(k/2^n) : 0 \leq k \leq 2^n)$ es combinación lineal del vector Gaussiano $(X_{n-1}(k/2^{n-1}) : 0 \leq k \leq 2^{n-1})$ y de la familia de variables aleatorias Gaussianas $(\xi_{k,n} : 0 \leq k \leq 2^n)$ las cuales son independientes de $(X_{n-1}(k/2^{n-1}) : 0 \leq k \leq 2^{n-1})$; en consecuencia es un vector Gaussiano. Para probar (3.5), basta ver que

$$\mathbb{E}[X_n^2(k/2^n)] = \frac{k}{2^n},$$

los cual es claro para el caso $k = 2j$, por la hipótesis de inducción. Veamos que también se cumple para el caso $k = 2j + 1$,

$$\begin{aligned} X_n\left(\frac{2j+1}{2^n}\right) &= X_{n-1}\left(\frac{2j+1}{2^{n-1}}\right) + \frac{\xi_{2j+1,n}}{2^{\frac{n+1}{2}}} \\ &= \frac{1}{2}\left(X_{n-1}\left(\frac{j}{2^{n-1}}\right) + X_{n-1}\left(\frac{j+1}{2^{n-1}}\right)\right) + \frac{\xi_{2j+1,n}}{2^{\frac{n+1}{2}}}, \end{aligned}$$

lo cual implica que

$$\begin{aligned}\mathbb{E}\left[X_n\left(\frac{2j+1}{2^n}\right)\right] &= \frac{1}{4}\left(\mathbb{E}\left[X_{n-1}^2\left(\frac{j}{2^{n-1}}\right)\right] + \mathbb{E}\left[X_{n-1}^2\left(\frac{j+1}{2^{n-1}}\right)\right]\right) \\ &\quad + \frac{1}{2}\mathbb{E}\left[X_{n-1}\left(\frac{j}{2^{n-1}}\right)X_{n-1}\left(\frac{j+1}{2^{n-1}}\right)\right] + \frac{1}{2^{n+1}} \\ &= \frac{2j+1}{2^n}.\end{aligned}$$

Por lo que para toda $n \geq 0$, $(X_n(t))_{t \in [0,1]}$ es un proceso Gaussiano centrado.

Finalmente para construir al movimiento Browniano en el intervalo $[0, 1]$ definimos los siguientes conjuntos,

$$A_n = \left\{ \sup_{t \in [0,1]} |X_n(t) - X_{n-1}(t)| > \frac{1}{2^{\frac{n}{4}}} \right\}.$$

Notemos que

$$\begin{aligned}\mathbb{P}(A_n) &= \mathbb{P}\left(\bigcup_{j=0}^{2^n-1} \left\{ \sup_{t \in [j/2^n, (j+1)/2^n]} |X_n(t) - X_{n-1}(t)| > \frac{1}{2^{\frac{n}{4}}} \right\}\right) \\ &= \mathbb{P}\left(\bigcup_{j=0}^{2^n-1} \left\{ \sup_{t \in [j/2^n, (j+1)/2^n]} \left| \frac{\xi_{2j+1,n}}{2^{(n+1)/4}} \right| > \frac{1}{2^{\frac{n}{4}}} \right\}\right) \\ &\leq 2^{n-1} \mathbb{P}(|N(0, 1)| > 2^{(n+2)/4}) \\ &= 2^n \mathbb{P}(N(0, 1) > 2^{(n+2)/4}) \quad (\text{Principio de Reflexión}) \\ &\leq \frac{1}{\sqrt{\pi}} 2^{(3n-4)/4} e^{-2^{n/2}},\end{aligned}$$

lo cual implica que $\sum_{n \geq 0} \mathbb{P}(A_n) < \infty$. Gracias al Lema de Borel-Cantelli tendremos que existe un conjunto $\tilde{\Omega} \in \mathcal{F}$ con $\mathbb{P}(\tilde{\Omega}) = 1$ de manera que para cada $\omega \in \tilde{\Omega}$, existe un entero $N_0 = N_0(\omega)$ tal que para toda $n \geq N_0$,

$$\sup_{t \in [0,1]} |X_n(t) - X_{n-1}(t)| > \frac{1}{2^{n/4}},$$

esto es, casi seguramente el proceso $(X_n(\cdot))$ converge uniformemente en el intervalo $[0, 1]$. Como esto ocurre en el espacio $C([0, 1], \mathbb{R})$ dotado con la topología uniforme dicho límite es una función continua. Sea

$$X(t) = \lim_{n \rightarrow \infty} X_n(t), \quad t \in [0, 1],$$

por construcción el proceso $(X(t) : t \in [0, 1])$ es un proceso Gaussiano que satisface

$$\mathbb{E}[X(t)X(s)] = \lim_{n \rightarrow \infty} \mathbb{E}[X_n(s_n)X_n(t_n)] = \lim_{n \rightarrow \infty} s_n \wedge t_n = s \wedge t,$$

donde s_n y t_n son diádicos de la forma $j2^{-n}$ y tal que $s_n \rightarrow s$ y $t_n \rightarrow t$.

Para finalizar la prueba consideramos B^0, B^1, \dots movimientos Brownianos independientes definidos en el intervalo $[0, 1]$ y definimos al proceso,

$$B_t = \begin{cases} B_t^0 & \text{si } t \in [0, 1) \\ B_{t-n}^n + \sum_{m=0}^{n-1} B_1^m & \text{si } t \in [n, n+1) \end{cases}$$

el cual claramente es un movimiento browniano definido en \mathbb{R}_+ . \square

Consideramos ahora al espacio $C := C(\mathbb{R}_+, \mathbb{R})$, el espacio de funciones continuas definidas en \mathbb{R}_+ , con σ -álgebra de Borel \mathcal{C} generada por la topología uniforme.

Definición 53. (Proceso de Coordenadas)

Definimos al proceso de coordenadas $X = (X_t)_{t \in [0, 1]}$ en C mediante la aplicación

$$\begin{aligned} X_t : C &\longrightarrow \mathbb{R} \\ \omega &\longmapsto X_t(\omega) := \omega(t). \end{aligned}$$

Proposición 54. Las σ -álgebras \mathcal{C} y $\sigma(X_t, 0 \leq t \leq 1)$ coinciden.

Demostración. Notemos en primera instancia que para toda $t \geq 0$, X_t es continua en $C(\mathbb{R}_+, \mathbb{R})$, por lo tanto son medibles respecto a $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$, es decir tendremos que

$$\sigma(X_t : t \geq 0) \subseteq \mathcal{C}(\mathbb{R}_+, \mathbb{R})$$

Ahora bien, para toda $\omega_0 \in C(\mathbb{R}_+, \mathbb{R})$, vemos que

$$d_n(\omega, \omega_0) = \sup_{t \in [0, n] \cap \mathbb{Q}} |\omega(t) - \omega_0(t)|.$$

es $\sigma(X_t : t \geq 0)$ -medible. Sea F un conjunto cerrado en $C(\mathbb{R}_+, \mathbb{R})$, y $(\omega_n)_n$ una sucesión densa en F . Notemos que

$$\begin{aligned} F &= \{\omega \in C(\mathbb{R}_+, \mathbb{R}) : d(\omega, F) = 0\} \\ &= \{\omega \in C(\mathbb{R}_+, \mathbb{R}) : \inf_n d(\omega, \omega_n) = 0\} \in \sigma(X_t : t \geq 0) \end{aligned}$$

Como F fue un cerrado arbitrario de $C(\mathbb{R}_+, \mathbb{R})$ concluimos que

$$\mathcal{C}(\mathbb{R}_+, \mathbb{R}) \subseteq \sigma(X_t : t \geq 0).$$

Con esto finalizamos la prueba. □

Sea $Z = (Z_t)_{t \in [0,1]}$ un proceso con espacio de estados $(\mathbb{R}, \mathcal{R})$ y trayectorias continuas casi seguramente -recordemos que \mathcal{R}^n es la σ -álgebra de Borel de \mathbb{R}^n - y definido por la aplicación

$$\begin{aligned} \phi : \Omega &\longrightarrow C \\ \omega &\longmapsto \phi(\omega) := Z(\omega) = (Z_t(\omega), t \in [0, 1]). \end{aligned}$$

La función ϕ es \mathcal{B} - \mathcal{C} -medible ya que $X_t \circ \phi = Z_t$ es medible para cada $t \in [0, 1]$. Denotemos por \mathbb{P}_Z a la medida imagen de \mathbb{P} por la aplicación ϕ . Llanamente, para todo vector n -dimensional $(t_1, \dots, t_n) \in \mathbb{R}^n$ y todo Boreliano $A \in \mathcal{R}^n$

$$\mathbb{P}((Z_{t_1}, \dots, Z_{t_n}) \in A) = \mathbb{P}_Z((X_{t_1}, \dots, X_{t_n}) \in A),$$

es decir, los procesos X y Z tienen las mismas distribuciones finito dimensionales. La ley \mathbb{P}_Z queda entonces determinada por la leyes de

$$\{(Z_{t_1}, \dots, Z_{t_n}) : (t_1, \dots, t_n) \in [0, 1]^n, \quad n \geq 1\},$$

puesto que la clase de los conjuntos finito dimensionales

$$\{(X_{t_1}, \dots, X_{t_n}) \in A : (t_1, \dots, t_n) \in [0, 1]^n, \quad A \in \mathcal{R}^n, \quad n \geq 1\},$$

es una Clase Determinante, esto es, dos medidas sobre $C(\mathbb{R}_+, \mathbb{R})$ que atribuyen la misma masa a los conjuntos de este tipo son idénticas. Luego, el proceso de coordenadas X y el proceso Z son equivalentes. La medida de probabilidad \mathbb{P}_Z es llamada la *ley del proceso aleatorio Z* .

Como ejemplo de nuestro interés tomemos a Z como el movimiento Browniano. La ley de Z es llamada la *Medida de Wiener* sobre $C(\mathbb{R}_+, \mathbb{R})$, la cual denotamos por \mathbb{W} . La medida de Wiener es una medida probabilidad sobre C caracterizada por las siguientes propiedades:

$$(i) \quad \mathbb{W}\{x \in C : x(0) = 0\} = 1$$

(ii) Para todo entero n y $\alpha_i \geq 0$ con $i = 1, 2, \dots, n$,

$$\begin{aligned} & \mathbb{W} \left(\bigcap_{i=1}^n \{x \in C : x(t_i) - x(t_{i-1}) \leq \alpha_i\} \right) \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \int_{-\infty}^{\alpha_i} \exp \left\{ -\frac{u^2}{2(t_i - t_{i-1})} \right\} du. \end{aligned}$$

Llamamos a $(X_t)_{t \geq 0}$ en el espacio $(C, \mathcal{C}, \mathbb{W})$ el proceso canónico del movimiento Browniano.

Teorema 55. *La medida de Wiener existe y es la única medida de probabilidad en C , bajo la cual el proceso de coordenadas es un movimiento Browniano.*

Sea $X = (X_t)_{t \geq 0}$ el proceso canónico del movimiento Browniano. Gracias a la continuidad de las trayectorias tenemos que para toda $t > 0$,

$$\sup_{s \in [s, t]} X_s = \sup_{s \in [s, t] \cap \mathbb{Q}} X_s,$$

es una variable aleatoria. De la misma manera vemos que

$$\int_0^t X_s^2 ds,$$

también es una variable aleatoria. En general si B es un movimiento Browniano definido en un espacio arbitrario, entonces para toda $t > 0$

$$\sup_{s \in [s, t]} B_s, \quad \inf_{s \in [s, t]} B_s, \quad \int_0^t B_s^2 ds, \quad \int_0^t e^{B_s} ds, \quad \dots$$

son variables aleatorias.

Sea $x \in \mathbb{R}$, la medida \mathbb{W}^x es la medida imagen en C de \mathbb{W} por la aplicación,

$$\begin{aligned} C(\mathbb{R}_+, \mathbb{R}) &\longrightarrow C(\mathbb{R}_+, \mathbb{R}) \\ \omega &\longmapsto \omega + x. \end{aligned}$$

El proceso $(X_t)_{t \geq 0}$ bajo \mathbb{W}^x es llamado movimiento Browniano comenzado en x . Este proceso está caracterizado por :

(i) $X_0 = x$ c.s.

(ii) Es un proceso con incrementos independientes.

(iii) Para toda $s \leq t$, $X_t - X_s \stackrel{(d)}{=} N(0, t - s)$.

3.4. El Teorema de Donsker

Dada una sucesión de variables aleatorias $(\xi_i)_{i=1}^{\infty}$ definidas sobre un espacio de probabilidad $(\Omega, \mathcal{B}, \mathbb{P})$, consideramos las sumas parciales

$$S_n = \sum_{i=1}^n \xi_i$$

y definimos un elemento aleatorio X_n de C por:

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt](\omega)} + (nt - [nt]) \frac{1}{\sigma\sqrt{n}} \xi_{[nt]+1}(\omega). \quad (3.6)$$

Probaremos la convergencia en distribución de las X_n , pero antes presentamos un lema que nos auxiliará al momento de demostrar que dicha colección es tensa.

Lema 56. Sean $\xi_1, \xi_2, \dots, \xi_n$ variables aleatorias independientes con media 0 y varianzas finitas positivas σ_i^2 , $i = 1, 2, \dots, n$. Haciendo $S_i = \xi_1 + \dots + \xi_i$ y $\alpha_i^2 = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_i^2$ tenemos que

$$\mathbb{P}\left\{ \max_{i \leq n} |S_i| \geq \lambda \alpha_n \right\} \leq 2\mathbb{P}\left\{ |S_n| \geq (\lambda - \sqrt{2})\alpha_n \right\}.$$

Demostración. Notemos en primera instancia que si $\lambda \leq \sqrt{2}$, el resultado es trivial. Consideremos los eventos

$$E_k = \left\{ \max_{i \leq k} |S_i| < \lambda \alpha_n \leq |S_k| \right\}.$$

Si $\lambda > \sqrt{2}$ es fácil ver que

$$\begin{aligned}
\left\{ \max_{i \leq n} |S_i| \geq \lambda \alpha_n \right\} &= \bigcup_{i=1}^n \{ |S_i| \geq \lambda \alpha_n \} \\
&\subseteq \{ |S_n| \geq (\lambda - \sqrt{2}) \alpha_n \} \cup \left[\bigcup_{i=1}^{n-1} \{ |S_i| \geq \lambda \alpha_n \} \right] \\
&= \{ |S_n| \geq (\lambda - \sqrt{2}) \alpha_n \} \cup \left[\bigcup_{i=1}^{n-1} \{ |S_i| \geq \lambda \alpha_n \} \cap \{ |S_n| < (\lambda - \sqrt{2}) \alpha_n \} \right] \\
&= \{ |S_n| \geq (\lambda - \sqrt{2}) \alpha_n \} \\
&\quad \cup \left[\bigcup_{i=1}^{n-1} \left\{ \max_{k < i} |S_k| < \lambda \alpha_n \right\} \cap \{ |S_i| \geq \lambda \alpha_n \} \cap \{ |S_n| < (\lambda - \sqrt{2}) \alpha_n \} \right] \\
&= \{ |S_n| \geq (\lambda - \sqrt{2}) \alpha_n \} \cup \left[\bigcup_{i=1}^{n-1} E_i \cap \{ |S_n| < (\lambda - \sqrt{2}) \alpha_n \} \right]
\end{aligned}$$

donde en la segunda y tercera igualdad se han ajenizado las uniones, obteniendo que

$$\begin{aligned}
&\mathbb{P} \left\{ \max_{i \leq n} |S_i| \geq \lambda \alpha_n \right\} \\
&\leq \mathbb{P} \{ |S_n| \geq (\lambda - \sqrt{2}) \alpha_n \} + \sum_{i=1}^{n-1} \mathbb{P} (E_i \cap \{ |S_n| < (\lambda - \sqrt{2}) \alpha_n \}). \quad (3.7)
\end{aligned}$$

Como $|S_i| \geq \lambda \alpha_n$ y $|S_n| < (\lambda - \sqrt{2}) \alpha_n$ juntos implican $|S_n - S_i| \geq \alpha_n \sqrt{2}$, se sigue de la desigualdad de Chebyshev y la hipótesis de independencia de las ξ_i que la sumatoria en (3.7) es a lo mas

$$\begin{aligned}
&\sum_{i=1}^{n-1} \mathbb{P} (E_i) \mathbb{P} \{ |S_n - S_i| < \sqrt{2} \alpha_n \} \\
&\leq \sum_{i=1}^{n-1} \left[\mathbb{P} (E_i) \frac{1}{2 \alpha_n^2} \sum_{k=i+1}^n \sigma_k^2 \right] \\
&\leq \frac{1}{2} \sum_{i=1}^{n-1} \mathbb{P} (E_i) \leq \frac{1}{2} \mathbb{P} \left\{ \max_{i \leq n} |S_i| \geq \lambda \alpha_n \right\}. \quad (3.8)
\end{aligned}$$

Combinando (3.7) y (3.8), llegamos a la desigualdad deseada. \square

Teorema 57. (Donsker)

Supongamos que las variables aleatorias ξ_n son independientes e idénticamente distribuidas con media 0 y varianza $\sigma^2 > 0$:

$$\mathbb{E}[\xi_n] = 0, \quad \mathbb{E}[\xi_n^2] = \sigma^2.$$

Entonces las funciones aleatorias X_n definidas por (3.6), satisfacen

$$X_n \xrightarrow[n \rightsquigarrow \infty]{\mathfrak{D}} B,$$

donde B es un movimiento Browniano estándar.

Demostración. Primero mostraremos que las distribuciones finito dimensionales de las variables aleatorias X_n convergen a las de B . Para $k = 1$, consideramos un tiempo s ; queremos ver que

$$\pi_s(X_n) = X_n(s) \xrightarrow{\mathfrak{D}} B_s. \quad (3.9)$$

Como $|\lceil \alpha \rceil - \alpha| \leq 1$ para todo $\alpha \in \mathbb{R}$, obtenemos de (3.6) que

$$\left| X_n(s) - \frac{1}{\sigma\sqrt{n}} S_{[ns]} \right| \leq \frac{1}{\sigma\sqrt{n}} \xi_{[ns]+1} \xrightarrow{\mathbb{P}} 0 \quad (3.10)$$

donde la convergencia en probabilidad se sigue de la desigualdad de Markov. Habiendo mostrado esto, (3.9) se seguirá del Teorema 23 si probamos que

$$\frac{1}{\sigma\sqrt{n}} S_{[ns]} \xrightarrow{\mathfrak{D}} B_s.$$

Pero esto último es una clara consecuencia del Teorema del Límite Central y el hecho de que $[ns]/n \rightarrow s$ cuando $n \rightarrow \infty$.

Para $k = 2$, consideramos ahora dos tiempos t y s con $s < t$. Queremos probar que

$$\pi_{s,t}(X_n) = (X_n(s), X_n(t)) \xrightarrow{\mathfrak{D}} (B_s, B_t),$$

lo cual será consecuencia del Corolario 26 del Teorema 25 si probamos que

$$(X_n(s), X_n(t) - X_n(s)) \xrightarrow{\mathfrak{D}} (B_s, B_t - B_s),$$

ya que $(X_n, Y_n) \xrightarrow{\mathfrak{D}} (X, Y)$ implica $X_n + Y_n \xrightarrow{\mathfrak{D}} X + Y$.

Por (3.10) y ésta misma relación pero con t en vez de s , es suficiente probar que

$$\left(\frac{1}{\sigma\sqrt{n}} S_{[ns]}, \frac{1}{\sigma\sqrt{n}} (S_{[nt]} - S_{[ns]}) \right) \xrightarrow{\mathfrak{D}} (B_s, B_t - B_s).$$

En esta última relación las entradas del vector izquierdo son independientes, por la independencia de las ξ_n , por lo que la convergencia en distribución se sigue del Teorema del Límite Central y el Teorema 19. De esta manera podemos ir tomando

conjuntos finitos de tiempos e ir probando la convergencia en distribución de las distribuciones finito dimensionales para $k = 3, 4, \dots$, verificando inductivamente que las distribuciones finito dimensionales convergen propiamente.

Por último veremos que la colección de variables $\{X_n\}_{n \in \mathbb{N}}$ es tensa. Por el Lema 56 tenemos que para $\lambda > 2\sqrt{2}$,

$$\mathbb{P}\left\{\max_{k \leq n} |S_k| \geq \lambda \sigma \sqrt{n}\right\} \leq 2 \mathbb{P}\left\{|S_n| \geq \lambda \frac{1}{2} \sigma \sqrt{n}\right\}.$$

Por el Teorema del Límite Central,

$$\mathbb{P}\left\{|S_n| \geq \lambda \frac{1}{2} \sigma \sqrt{n}\right\} \xrightarrow{n \rightarrow \infty} \mathbb{P}\left\{|N_{(0,1)}| \geq \lambda \frac{1}{2}\right\} < \frac{2^3}{\lambda^3} \mathbb{E}\left[|N_{(0,1)}^3|\right].$$

Luego si $\epsilon > 0$, tenemos que para λ suficientemente grande

$$\limsup_{n \rightarrow \infty} \mathbb{P}\left\{\max_{k \leq n} |S_k| \geq \lambda \sigma \sqrt{n}\right\} < \frac{\epsilon}{\lambda^2}.$$

Por la definición de límite superior, tendremos que existe un entero no negativo N , tal que

$$\sup_{n \geq N} \mathbb{P}\left\{\max_{k \leq n} |S_k| \geq \lambda \sigma \sqrt{k}\right\} < \frac{\epsilon}{\lambda^2},$$

por lo que finalmente, para toda $n \geq N$,

$$\mathbb{P}\left\{\max_{k \leq n} |S_i| \geq \lambda \sigma \sqrt{k}\right\} < \frac{\epsilon}{\lambda^2}.$$

De esto último, concluimos que la colección $\{X_n\}_{n \in \mathbb{N}}$ es tensa por el Teorema 45. Finalmente, de la convergencia de las distribuciones finito dimensionales, se obtiene gracias a la Proposición 43 el resultado deseado. \square

Capítulo 4

Árboles Aleatorios

En un principio se explica como los árboles aleatorios discretos pueden ser codificados por trayectorias discretas como la función altura y la función contorno del árbol. Se da una prueba de que la función altura reescalada asociada de un bosque de árboles de Galton-Watson independientes converge en distribución hacia el Movimiento Browniano Reflejado sobre la recta real positiva. Se verán posteriormente ciertas consecuencias interesantes de esto último. Recuperamos un famoso teorema de Aldous mostrando que para una función contorno reescalada adecuada de un árbol de Galton-Watson condicionado a tener n -vértices converge en distribución hacia la excursión Browniana normalizada cuando n crece hacia infinito.

4.1. Árboles Discretos

En esta sección nos interesamos en árboles aleatorios (finitos) ordenados con raíz. Damos a continuación el conjunto de etiquetas

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

donde $\mathbb{N} = \{1, 2, 3, \dots\}$ y por convención $\mathbb{N}^0 = \{\emptyset\}$. Un elemento u en \mathcal{U} como lo sugiere la definición es un elemento $u = (u^1, u^2, \dots, u^m) \in \mathbb{N}^m$ para algún m natural y escribimos $|u| = m$ para denotar esta pertenencia; es común referirse a $|u|$ como la “*generación*” del elemento u .

Si $u = (u^1, u^2, \dots, u^m)$ y $v = (v^1, v^2, \dots, v^n)$ son elementos de \mathcal{U} , escribimos

$$uv = (u^1, u^2, \dots, u^m, v^1, v^2, \dots, v^n)$$

para la concatenación de u y v . En particular tendremos que

$$u\emptyset = \emptyset u = u \quad \text{y} \quad u = (u^1, u^2, \dots, u^m) = u^1 u^2 \dots u^m.$$

Finalmente, escribimos $u \preceq v$ si u es un “ancestro” de v , queriendo decir con esto que existe $w \in \mathcal{U}$ tal que $uw = v$. El conjunto (\mathcal{U}, \preceq) es Parcialmente Ordenado. Sean $u, v \in \mathcal{U}$, denotamos por $u \wedge v$ al “ancestro común” más grande de u y v , o bien al ancestro común más inmediato a ambos.

Al conjunto \mathcal{U} lo consideraremos con el orden lexicográfico (\leq); donde si $u, v \in \mathcal{U}$, tenemos que $u \leq v$ si y sólo si $u \preceq v$ o bien $u \wedge v$ es un ancestro estrictamente distinto de u y v de manera que $u^{|u \wedge v|+1} < v^{|u \wedge v|+1}$. El conjunto (\mathcal{U}, \leq) es Totalmente Ordenado.

Sea $\varphi : \mathcal{U} \setminus \{\emptyset\} \rightarrow \mathcal{U}$ una aplicación definida por

$$\varphi(u^1 u^2 \dots u^m) = u^1 u^2 \dots u^{m-1},$$

la cuál considerando el orden lexicográfico es creciente ($\varphi(u)$ es el “padre” de u).

Definición 58. (Árbol Finito Ordenado con Raíz)

Un árbol finito ordenado con raíz, es un subconjunto finito $\mathbf{t} \subset \mathcal{U}$ tal que:

- (i) $\emptyset \in \mathbf{t}$
- (ii) Si $u \in \mathbf{t} \setminus \{\emptyset\}$ entonces también $\varphi(u) \in \mathbf{t}$
- (iii) Para todo $u \in \mathbf{t}$ existe un entero $k_u(\mathbf{t}) \geq 0$, tal que $\forall i \in \mathbb{N}$:

$$ui \in \mathbf{t} \quad \text{si y solo si} \quad 1 \leq i \leq k_u(\mathbf{t}).$$

Notación. $\mathcal{A} = \{ \text{Árboles finitos ordenados con raíz} \}$.

Los elementos de \mathbf{t} son llamados *vértices*. Interpretamos al número real $k_u(\mathbf{t})$ como el *número de hijos* de u en \mathbf{t} ; debido a esta última interpretación, es posible denotar

$$k_u(\mathbf{t}) = \sup \{ i \geq 1 : ui \in \mathbf{t} \}.$$

La noción de *hermano*, *ancestro* y *descendiente* son naturalmente inducidas por la de *hijo* tal como en la vida cotidiana. La generación $|u|$ es igualmente conocida como la *altura* de u . Los vértices que no tienen hijos son llamados *hojas*.

En lo subsecuente, se considerará a los elementos de un árbol $\mathbf{t} \in \mathcal{A}$, como individuos de una población para la cuál \mathbf{t} es el “*árbol genealógico*”. La cardinalidad de \mathbf{t} , es el número de vértices en el árbol y lo denotamos por $\#(\mathbf{t})$.

Un árbol finito ordenado con raíz tiene la representación de una gráfica plana - de donde proviene su nombre alternativo de árbol plano - donde los vértices están conectados por un segmento de recta con cada uno de sus hijos y se encuentran ordenados de izquierda a derecha en orden lexicográfico.

Una observación importante es que los árboles $\mathbf{t} \in \mathcal{A}$ al ser subconjuntos finitos de (\mathcal{U}, \leq) y considerarlos con el orden lexicográfico (\mathbf{t}, \leq) , poseen un Buen Orden, con lo que es posible tomar elementos mínimos con respecto a \leq .

Notación. $\mathcal{A}_n = \{\mathbf{t} \in \mathcal{A} : \#(\mathbf{t}) = n\}$ de donde

$$\mathcal{A} = \bigcup_{n \geq 1} \mathcal{A}_n.$$

Así las cosas, si $\mathbf{t} \in \mathcal{A}$ entonces existe un entero $n = \#(\mathbf{t})$ tal que $\mathbf{t} \in \mathcal{A}_n$. Luego, si $\emptyset = u_1, u_2, \dots, u_{\#(\mathbf{t})}$ son los elementos de \mathbf{t} en orden lexicográfico es posible asociar a \mathbf{t} con el vector

$$(k_{u_1}(\mathbf{t}), \dots, k_{u_{\#(\mathbf{t})}}(\mathbf{t})),$$

cuya i -ésima entrada es respectivamente el número de hijos del i -ésimo elemento más grande de \mathbf{t} respecto al orden lexicográfico.

Formalmente con lo anterior es posible definir una función inyectiva

$$\begin{aligned} \mathcal{A} &\longrightarrow \bigcup_{n \geq 1} \mathbb{N}^n \\ \mathbf{t} &\longmapsto (k_{u_1}(\mathbf{t}), \dots, k_{u_{\#(\mathbf{t})}}(\mathbf{t})). \end{aligned}$$

La siguiente proposición nos indica que efectivamente esta aplicación es inyectiva.

Proposición 59. (Caracterización de árboles en \mathcal{A})

Sean $\mathbf{t}_1, \mathbf{t}_2 \in \mathcal{A}$ tales que $\#(\mathbf{t}_1) = \#(\mathbf{t}_2) = n$. Sean u_1, \dots, u_n y v_1, \dots, v_n respectivamente los elementos de \mathbf{t}_1 y \mathbf{t}_2 en orden lexicográfico. Si para cada $\ell \in \mathbb{N}$ se tiene que $k_{u_\ell}(\mathbf{t}_1) = k_{v_\ell}(\mathbf{t}_2)$, entonces $\mathbf{t}_1 = \mathbf{t}_2$. En particular $u_\ell = v_\ell$ para toda $\ell \leq n$.

Demostración. Sean $\mathbf{t}_1, \mathbf{t}_2 \in \mathcal{A}$ y supongamos que $\mathbf{t}_1 \neq \mathbf{t}_2$. Entonces:

Caso 1 Si $k_\emptyset(\mathbf{t}_1) \neq k_\emptyset(\mathbf{t}_2)$ ya acabamos.

Caso 2 Si $k_\emptyset(\mathbf{t}_1) = k_\emptyset(\mathbf{t}_2)$, sea

$$z = \text{mín} \{u \in (\mathcal{U}, \leq) : u \in (\mathbf{t}_1 \setminus \mathbf{t}_2) \cup (\mathbf{t}_2 \setminus \mathbf{t}_1)\},$$

esto es, z es el primer elemento en \mathcal{U} tal que $z \notin \mathbf{t}_1 \cap \mathbf{t}_2$; sin pérdida de generalidad supongamos que $z \in \mathbf{t}_1 \setminus \mathbf{t}_2$. Por minimalidad de z tendremos que si es el ℓ -ésimo elemento más grande en \mathbf{t}_1 ($z = u_\ell$) entonces $(u_1, \dots, u_{\ell-1}) = (v_1, \dots, v_{\ell-1})$, lo cual es equivalente a decir que $\mathbf{t}_1, \mathbf{t}_2$ comparten sus primeros $\ell - 1$ elementos. Veamos ahora porque

$$(k_{u_1}(\mathbf{t}_1), \dots, k_{u_{\ell-1}}(\mathbf{t}_1)) \neq (k_{v_1}(\mathbf{t}_2), \dots, k_{v_{\ell-1}}(\mathbf{t}_2)).$$

Notemos que $\varphi(z)$ el padre de z debe pertenecer a ambos árboles $\mathbf{t}_1, \mathbf{t}_2$; más aún, ocurre que debe tener la misma etiqueta en ambos (digamos $\varphi(z) = u_r = v_r$ con $r \leq \ell - 1$) puesto que antes de z los elementos en ambos árboles son los mismos y deben mantener el mismo orden lexicográfico. Sin embargo, claramente tenemos que $k_{\varphi(z)}(\mathbf{t}_1) > k_{\varphi(z)}(\mathbf{t}_2)$. Efectivamente, sea $z = \varphi(z) \cdot \alpha$ donde $\alpha \in \mathbb{N}$, entonces

$$z \in \mathbf{t}_1 \Rightarrow z = u_r \cdot \alpha \text{ donde } \alpha \leq k_{u_r}(\mathbf{t}_1),$$

por otro lado

$$z \notin \mathbf{t}_2 \Rightarrow z = v_r \cdot \alpha \text{ donde } \alpha > k_{v_r}(\mathbf{t}_2).$$

Por lo tanto

$$(k_{u_1}(\mathbf{t}_1), \dots, k_{u_r}(\mathbf{t}_1), \dots, k_{u_{\ell-1}}(\mathbf{t}_1)) \neq (k_{v_1}(\mathbf{t}_2), \dots, k_{v_r}(\mathbf{t}_2), \dots, k_{v_{\ell-1}}(\mathbf{t}_2)).$$

□

Ahora comenzaremos a entender como es que los árboles pueden ser codificados por funciones discretas. Introducimos la siguiente manera de codificar un árbol. Para cada $n \in \mathbb{N}$ consideramos a las sucesiones finitas de enteros no-negativos (m_1, \dots, m_n) tales que:

$$(1) \ m_1 + m_2 + \dots + m_i \geq i \quad \forall i \in \{1, 2, \dots, n-1\},$$

$$(2) \ m_1 + m_2 + \dots + m_n = n - 1.$$

Notación. Para cada $n \geq 1$ denotamos por

$$\mathcal{S}_n = \{(m_1, \dots, m_n) \in \mathbb{N}^n : \text{La sucesión finita cumple (1) y (2)}\}, \text{ y}$$

$$\mathcal{S} = \bigcup_{n \geq 1} \mathcal{S}_n.$$

Recordemos $u_1 = \emptyset, u_2, \dots, u_{\#(\mathbf{t})}$ son los elementos de \mathbf{t} en orden lexicográfico.

Proposición 60. *La función*

$$\Phi : \mathbf{t} \longmapsto (k_{u_1}(\mathbf{t}), k_{u_2}(\mathbf{t}), \dots, k_{u_{\#(\mathbf{t})}}(\mathbf{t}))$$

define una biyección entre \mathcal{A} y \mathcal{S} .

Demostración. Sea $\#(\mathbf{t}) = N$, la suma $k_{u_1}(\mathbf{t}) + k_{u_2}(\mathbf{t}) + \dots + k_{u_{\#(\mathbf{t})}}(\mathbf{t})$ es el total de hijos de todos los individuos del árbol y en consecuencia es igual a $N - 1$ (Obsérvese que \emptyset no es hijo de nadie por lo que no es contado entre estos). Más aún, si $i \in \{1, 2, \dots, N - 2\}$, $k_{u_1}(\mathbf{t}) + k_{u_2}(\mathbf{t}) + \dots + k_{u_i}(\mathbf{t})$ es el número de hijos de u_1, u_2, \dots, u_i , por lo tanto es mayor igual que i , porque u_2, u_3, \dots, u_i son contados entre estos hijos (en el orden lexicográfico se pasa por un individuo antes que por su hijo). Incluso es una desigualdad estricta ya que el padre de u_{i+1} está en $\{u_1, u_2, \dots, u_i\}$.

De esta información concluimos fácilmente que Φ es sobreyectiva, ya que para una sucesión $(m_1, \dots, m_\ell) \in \mathcal{S}$ bastará asociarle el elemento $\mathbf{t} \in \mathcal{A}$, tal que $k_{u_i}(\mathbf{t}) = m_i$ para toda $i \in \{1, \dots, \ell\}$ y $\#(\mathbf{t}) = \ell$.

Sean ahora $\mathbf{t}_1, \mathbf{t}_2 \in \mathcal{A}$ tal que $\Phi(\mathbf{t}_1) = \Phi(\mathbf{t}_2)$, esto ocurre si y solo si

$$\left(k_{u_1}(\mathbf{t}_1), k_{u_2}(\mathbf{t}_1), \dots, k_{u_{\#(\mathbf{t}_1)}}(\mathbf{t}_1) \right) = \left(k_{v_1}(\mathbf{t}_2), k_{v_2}(\mathbf{t}_2), \dots, k_{v_{\#(\mathbf{t}_2)}}(\mathbf{t}_2) \right),$$

lo que a su vez implica que

$$\#(\mathbf{t}_1) = \#(\mathbf{t}_2) = n \text{ y } k_{u_i}(\mathbf{t}_1) = k_{v_i}(\mathbf{t}_2) \text{ para toda } i \in \{1, \dots, n\}.$$

Se sigue de la Proposición 59 que $\mathbf{t}_1 = \mathbf{t}_2$. Por lo tanto Φ es inyectiva. \square

Esta última proposición en realidad nos da más información, ya que nos dice que para cada $n \in \mathbb{N}$, la restricción de Φ a \mathcal{A}_n es una biyección sobre \mathcal{S}_n .

Sea $\mathbf{t} \in \mathcal{A}_n$. Más que hacer referencia a $\Phi(\mathbf{t}) = (m_1, \dots, m_n)$, consideraremos frecuentemente la sucesión finita de enteros (x_1, \dots, x_n) definida por

$$x_k = \sum_{i=1}^k (m_i - 1) \text{ con } 0 \leq k \leq n,$$

que satisface las siguientes propiedades:

- (a) $x_0 = 0$ y $x_n = -1$,
- (b) $x_i \geq 0$ para todo $0 \leq i \leq n - 1$,
- (c) $(x_i - x_{i-1}) = (m_i - 1) \geq -1$ para todo $1 \leq i \leq n$.

Tal sucesión es llamada la *trayectoria de Lukasiewicz* del árbol \mathbf{t} .

Notación. Para cada $n \geq 1$ denotamos por

$$\mathcal{L}_n = \{(x_1, \dots, x_n) \in \mathbb{N}^n : \text{La sucesión finita cumple (a), (b) y (c)}\},$$

$$\mathcal{L}_{\mathcal{A}} = \bigcup_{n \geq 1} \mathcal{L}_n.$$

De manera muy natural podemos notar que la función Φ de la proposición anterior induce una biyección entre los árboles en \mathcal{A} y sus trayectorias de Lukasiewicz. Más aún, resulta ser que para cada $n \geq 1$, la biyección prevalece para \mathcal{A}_n , \mathcal{S}_n y \mathcal{L}_n .

Definición 61. (Función de altura $h_{\mathbf{t}}$)

Sea $\mathbf{t} \in \mathcal{A}$, denotemos por $u_o = \emptyset, u_1, u_2, \dots, u_{\#(\mathbf{t}-1)}$ a los elementos de \mathbf{t} en orden lexicográfico. La función de altura asociada al árbol \mathbf{t} , es la función

$$h_{\mathbf{t}} : \{n \in \mathbb{N} : 0 \leq n \leq \#(\mathbf{t}) - 1\} \longrightarrow \{|u_j| : 0 \leq j \leq \#(\mathbf{t}) - 1\}$$

definida por $h_{\mathbf{t}}(n) = |u_n|$.

La función altura puede entenderse como la sucesión ordenada de las generaciones de los individuos de \mathbf{t} , cuando estos individuos están enlistados en orden lexicográfico. A todo árbol en \mathcal{A} lo caracteriza su función altura, ya que nos da información acerca de la cardinalidad del árbol y el número de hijos de cada uno de sus vértices. Es posible verificar esto último inductivamente con la definición dada en términos de un supremo para el número de hijos de un vértice u en un árbol \mathbf{t} .

Existe otra función que nos permite caracterizar un árbol, la cuál es más fácil de visualizar sobre el plano.

Supongamos que tenemos un árbol \mathbf{t} en \mathcal{A} sobrepuesto sobre el plano superior de manera que cada borde tenga longitud 1. Imaginemos el movimiento de una partícula

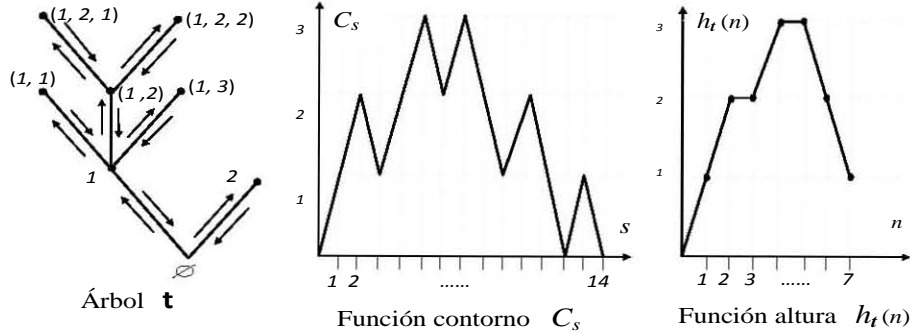


Figura 4.2: Funciones que codifican al árbol \mathbf{t} con $\#(\mathbf{t}) = 8$ y $\zeta(\mathbf{t}) = 14$ (Le Gall ,[11]).

que comienza al tiempo $t = 0$ en la raíz del árbol y lo explora de izquierda a derecha, moviéndose continuamente a lo largo de los segmentos de recta que unen a los vértices a una velocidad unitaria hasta que todos los segmentos han sido explorados y la partícula ha regresado a la raíz. Ya que es claro que cada segmento de recta será recorrido dos veces en esta evolución, tendremos que el tiempo total necesario para la exploración del árbol es de

$$\zeta(\mathbf{t}) = 2(\#(\mathbf{t}) - 1).$$

Definición 62. (Función contorno C_t)

Sea $\mathbf{t} \in \mathcal{A}$, la función contorno asociada al árbol \mathbf{t} , se define para $s \in [0, \zeta(\mathbf{t})]$ como

$$C_{\mathbf{t}}(s) = \begin{array}{l} \text{Distancia sobre el árbol, entre la posición} \\ \text{de la partícula al tiempo } s \text{ y la raíz.} \end{array}$$

Por convención, $C_{\mathbf{t}}(s) = 0$ para $s \geq \zeta(\mathbf{t})$.

Observemos ahora la relación existente entre la función de altura y la trayectoria de Lukasiewicz de un árbol en \mathcal{A} . Sea $\mathbf{t} \in \mathcal{A}$ y $u_0, u_1, \dots, u_{\#(\mathbf{t})-1}$ sus elementos en orden lexicográfico, de nuestras definiciones es fácil ver que la relación $u \prec v$ ocurre si y sólo si v es estrictamente un descendiente de u .

Proposición 63. Sea $h_{\mathbf{t}}$ la función altura de un árbol $\mathbf{t} \in \mathcal{A}$, entonces es posible expresar $h_{\mathbf{t}}$ en términos de la trayectoria de Lukasiewicz del respectivo árbol de la siguiente manera:

$$h_{\mathbf{t}}(n) = \# \left\{ i \in \{0, 1, \dots, n-1\} : x_i = \inf_{i \leq k \leq n} x_k \right\},$$

para todo $n \in \{0, 1, \dots, \#(\mathbf{t}) - 1\}$.

Demostración. Claramente

$$h_{\mathbf{t}}(n) = |u_n| = \#\{i \in \{0, 1, \dots, n-1\} : u_i \prec u_n\},$$

por lo que bastaría probar que para $i \in \{0, 1, \dots, n-1\}$, $u_i \prec u_n$ ocurre si y sólo si $x_i = \inf_{i \leq k \leq n} x_k$. Para este fin, es suficiente verificar que

$$\inf_k \{k \geq i : x_k < x_i\} = \begin{cases} \#(\mathbf{t}) & \text{si } u_i \prec u_k, \text{ para toda } k \in \{i+1, \dots, \#(\mathbf{t})\} \\ l^* & \text{si existe } k > i \text{ tal que } u_i \not\prec u_k \end{cases}$$

donde l^* es el primer índice $k > i$ tal que $u_i \not\prec u_k$.

Sin embargo, escribiendo

$$x_k - x_i = \sum_{r=i+1}^k (m_r - 1)$$

es posible ver que para todo entero $k > i$, tal que $u_i \prec u_k$ se tiene que

$$\sum_{r=i+1}^k m_r \geq (k - i) \quad \text{por lo que } x_k - x_i \geq 0.$$

Por otro lado, $x_{l^*} - x_i = -1$ si l^* es el primer índice $k > i$ tal que u_k no es descendiente de u_i ($k = \#(\mathbf{t})$ si tal l^* no existe). Esto completa la prueba. \square

4.2. Árboles Galton-Watson

La idea de utilizar procesos de ramificación para el estudio de árboles aleatorios se basa en la imagen intuitiva de un árbol como la realización del proceso de ramificación. El proceso de ramificación de Galton-Watson es quizá el más conveniente para la formalización de esta idea.

La familia de árboles Galton-Watson se define de la siguiente manera: Dada una medida de probabilidad μ con distribución $(\mu(k))_{k=0}^{\infty}$ sobre \mathbb{Z}_+ , o bien, equivalentemente, dada una variable aleatoria ξ con distribución $(\mu(k))_{k=0}^{\infty}$, construimos recursivamente un árbol θ , comenzando con la raíz y dotando a cada vértice de cierto número de hijos dado por una copia independiente de ξ . En otras palabras, el número de hijos de cada vértice será una variable aleatoria independiente de las demás con distribución $(\mu(k))_{k=0}^{\infty}$. Es común referirse a la medida de probabilidad μ con distribución $(\mu(k))_{k=0}^{\infty}$ como la distribución de hijos o progenie.

Sea μ una distribución de hijos crítica o subcrítica. Esto quiere decir que μ es una medida de probabilidad sobre \mathbb{Z}_+ tal que

$$\sum_{k=0}^{\infty} k\mu(k) \leq 1$$

Se excluirá el caso trivial cuando $\mu(1) = 1$.

Haremos uso de la siguiente construcción explícita de árboles de Galton-Watson:

Sea $(K_u, u \in \mathcal{U})$ una colección de variables aleatorias independientes con distribución μ , indexadas por el conjunto de etiquetas \mathcal{U} . Denotemos por θ al subconjunto aleatorio de \mathcal{U} definido por

$$\theta = \{u = u^1 \cdots u^n \in \mathcal{U} : u^i \leq K_{u^1 \dots u^{i-1}} \text{ para toda } 1 \leq i \leq n\} \cup \{\emptyset\}$$

Proposición 64.

θ es un árbol c.s. Más aún, si

$$Z_n = \#\{u \in \theta : |u| = n\}$$

entonces $(Z_n, n \geq 0)$ es un proceso de Galton-Watson con distribución de hijos μ y valor inicial $Z_0 = 1$.

OBSERVACIÓN. Claramente $k_u(\theta) = K_u \quad \forall u \in \theta$.

Demostración. Primero demostremos que θ es un elemento de \mathcal{A} casi seguramente. Basta probar que si $\theta \neq \emptyset$, entonces θ cumple las propiedades de la Definición (58).

Sea $\theta \neq \emptyset$. Como θ es no vacío, sabemos que cuenta con algún elemento u tal que $u = u^1 \cdots u^n$ para alguna $n \geq 1$. De la definición de θ , tendremos que

$$u^i \leq K_{u^1 \dots u^{i-1}} \text{ para toda } 1 \leq i \leq n.$$

Como $\varphi(u) = u^1 \cdots u^{n-1}$, por lo anterior tenemos que en particular

$$u^i \leq K_{u^1 \dots u^{i-1}} \text{ para toda } 1 \leq i \leq n-1.$$

Concluimos que $\varphi(u) \in \theta$ para toda $u \in \theta \setminus \{\emptyset\}$. De esta última propiedad, obtenemos que para $u_1 \in \theta \cap \mathbb{N}$, $\varphi(u_1) = \emptyset$. Considerando ahora la v.a. K_u , tendremos nuevamente por la definición de θ que $u \cdot u^{n+1} \in \theta$ si y sólo si $u^{n+1} \leq K_u$. Por lo tanto θ es un árbol.

La finitud del árbol θ se sigue al identificar a Z como un proceso de Galton-Watson con distribución de hijos μ (crítica o subcrítica), por lo que se extingue casi seguramente, de manera que $Z_n = 0$ para n suficientemente grande.

Hacemos para $n \geq 0$, $A_n = \{u \in \theta : |u| = n\}$. Como el número de individuos en la $(n + 1)$ -ésima es exactamente el total de hijos de los individuos de la n -ésima generación, es posible escribir

$$\#A_n = \sum_{v \in A_{n-1}} \sup \{i \geq 1 : v \cdot i \in \theta\} = \sum_{v \in A_{n-1}} K_v.$$

Entonces es posible representar al proceso Z de la siguiente manera,

$$\begin{aligned} Z_0 &= \#A_0 = 1. \\ Z_1 &= \#A_1 = K_\emptyset. \\ \text{Para } n \geq 2, \quad Z_n &= \#A_n = \sum_{i=1}^{Z_{n-1}} K_{v_{i,n-1}}, \quad \text{donde } v_{i,n-1} \in A_{n-1}. \end{aligned}$$

Por lo tanto Z_n denota el número de individuos de la n -ésima generación. Por este mismo argumento, se tiene que $Z_n = 0$ implica que $Z_{n+1} = 0$. Con esto hemos probado que Z es un proceso de Galton-Watson. \square

Al árbol θ y cualquier árbol aleatorio con la misma distribución, se le conoce como un árbol de Galton-Watson con distribución de hijos μ o de manera más corta un árbol μ -Galton-Watson. Escribimos Π_μ para la distribución de θ sobre el espacio \mathcal{A} .

Definición 65. (Desplazamiento de un árbol en \mathcal{A})

Sea $\mathbf{t} \in \mathcal{A}$. Para $1 \leq i \leq k_\emptyset(\mathbf{t})$, el árbol \mathbf{t} desplazado en i es el conjunto:

$$T_i \mathbf{t} = \{u \in \mathcal{U} : iu \in \mathbf{t}\}.$$

Nótese que $T_i \mathbf{t}$ es un árbol.

Entonces la distribución Π_μ puede ser caracterizada por las siguientes dos propiedades

- (i) $\Pi_\mu(k_\emptyset = j) = \mu(j) \quad j \in \mathbb{Z}_+$
- (ii) Para toda $j \geq 1$ con $\mu(j) \geq 0$, los árboles desplazados $T_1 \mathbf{t}, \dots, T_j \mathbf{t}$ son independientes bajo la probabilidad condicional $\Pi_\mu(d\mathbf{t} \mid k_\emptyset = j)$ y su distribución de probabilidad es Π_μ .

La propiedad (ii) es conocida como la “Propiedad de ramificación de los árboles de Galton-Watson”.

Se dará ahora una fórmula explícita para Π_μ

Proposición 66. Para todo $\mathbf{t} \in \mathcal{A}$,

$$\Pi_\mu(\mathbf{t}) = \prod_{u \in \mathbf{t}} \mu(k_u(\mathbf{t}))$$

Demostración. Es fácil convencerse de la igualdad (ver Proposición 59)

$$\{\theta = \mathbf{t}\} = \bigcap_{u \in \mathbf{t}} \{K_u = k_u(\mathbf{t})\},$$

de donde es inmediato que

$$\Pi_\mu(\mathbf{t}) = \mathbb{P}(\theta = \mathbf{t}) = \prod_{u \in \mathbf{t}} \mathbb{P}(K_u = k_u(\mathbf{t})) = \prod_{u \in \mathbf{t}} \mu(k_u(\mathbf{t})). \quad \square$$

Recordemos de la Proposición 60, la definición de la biyección Φ .

Proposición 67. Sea θ un árbol μ -Galton-Watson. Entonces

$$\Phi(\theta) \stackrel{(d)}{=} (M_1, M_2, \dots, M_T)$$

donde las variables aleatorias M_1, M_2, \dots son independientes con distribución μ , y

$$T = \inf\{n \geq 1 : M_1 + M_2 + \dots + M_n < n\}$$

OBSERVACIÓN El hecho de que $T < \infty$ c.s. es consecuencia de nuestro acercamiento aunque es posible probarlo directamente por la Ley Fuerte de los Grandes números una vez demostrado que son v.a.i.i.d. con distribución crítica o subcrítica μ .

Demostración. Asumamos que θ está dado por la construcción explícita anteriormente descrita. Sean entonces $U_0 = \emptyset, U_1, \dots, U_{\#(\theta)-1}$ los elementos de θ escritos en orden lexicográfico, de manera que

$$\Phi(\theta) = (K_{U_0}, K_{U_1}, \dots, K_{U_{\#(\theta)-1}}).$$

Sabemos entonces que

$$K_{U_0} + K_{U_1} + \cdots + K_{U_n} \geq n + 1, \quad \forall n \in \{0, 1, \dots, \#(\theta) - 1\}$$

$$K_{U_0} + K_{U_1} + \cdots + K_{U_{\#(\theta)-1}} = \#(\theta) - 1.$$

Definimos para $p \geq \#(\theta)$, a U_p como la concatenación

$$U_p = U_{\#(\theta)-1} \cdot (1, 1, 1, \dots, 1) = U_{\#(\theta)-1} \cdot 1 \cdot 1 \cdot 1 \cdots 1$$

donde en la parte derecha se han agregado $p - \#(\theta) + 1$ etiquetas 1; entonces la prueba se reduce a mostrar que, para todo $p \geq 0$, $K_{U_0}, K_{U_1}, \dots, K_{U_p}$ son independientes con distribución μ .

Debemos notar que las etiquetas U_i son aleatorias ya que dependen de la colección $(K_u, u \in \mathcal{U})$ y por tanto no es posible únicamente argumentar que las variables K_u son i.i.d. con distribución μ . Se procederá por inducción sobre p . Para $p = 0$ o $p = 1$, el resultado es obvio ya que $U_0 = \emptyset$ y $U_1 = 1$ son deterministas.

Sea $p \geq 2$ y supongamos que el resultado es válido para $k \leq p - 1$. Utilizamos la notación $u \leq v$ para el orden lexicográfico. Observemos que para toda $u \in \mathcal{U}$ fija, el conjunto aleatorio

$$\theta \cap \{v \in \mathcal{U} : v \leq u\}$$

es medible respecto a la σ -álgebra $\sigma(K_v, v < u)$. Esto último se sigue de la construcción de θ . Como consecuencia, el evento

$$\{U_p = u\} \cap \{\#(\theta) > p\}$$

es medible respecto a $\sigma(K_v, v \leq u)$. Fácilmente se puede observar que esta propiedad de medibilidad se mantiene para el evento

$$\{U_p = u\} \cap \{\#(\theta) \leq p\}$$

Concluimos entonces que $\{U_p = u\}$ es medible respecto a $\sigma(K_v, v < u)$.

Finalmente, si g_0, g_1, \dots, g_p son funciones no-negativas sobre $\{0, 1, \dots\}$,

$$\begin{aligned} & \mathbb{E}[g_0(K_{U_0}) \cdot g_1(K_{U_1}) \cdots g_p(K_{U_p})] \\ = & \sum_{u_0 < u_1 < \cdots < u_p} \mathbb{E}[\mathbf{1}_{\{U_0=u_0, \dots, U_p=u_p\}} \cdot g_0(K_{u_0}) \cdots g_p(K_{u_p})] \\ = & \sum_{u_0 < u_1 < \cdots < u_p} \mathbb{E}[\mathbf{1}_{\{U_0=u_0, \dots, U_p=u_p\}} \cdot g_0(K_{u_0}) \cdots g_{p-1}(K_{u_{p-1}})] \mathbb{E}[g_p(K_{u_p})] \end{aligned}$$

porque K_{u_p} es independiente de $\sigma(K_v, v < u_p)$, y se ocupo la propiedad de medibilidad antes mencionada. Entonces $\mathbb{E}[g_p(K_{u_p})] = \mu_{(g_p)}$ no depende de u_p , y tomando $g_p \equiv Id$ en la fórmula anterior observamos que

$$\mathbb{E}[g_1(K_{U_0}) \cdot g_2(K_{U_2}) \cdots g_p(K_{U_p})] = \mathbb{E}[g_1(K_{U_0}) \cdot g_2(K_{U_2}) \cdots g_{p-1}(K_{U_{p-1}})] \times \mu_{(Id)}.$$

Asumiendo la prueba por inducción de ésta última fórmula, la prueba está completa. \square

Corolario 68. *Sea $(S_n, n \geq 0)$ una caminata aleatoria sobre \mathbb{Z} con valor inicial S_0 y distribución de salto $\nu(k) = \mu(k+1) \forall k \geq -1$. Definamos*

$$T = \inf\{n \geq 1 : S_n = -1\}.$$

Entonces las trayectorias de Lukasiewicz de un árbol θ , μ -Galton-Watson tiene la misma distribución que (S_0, S_1, \dots, S_T) . En particular, $\#(\theta)$ y T tienen la misma distribución.

Recordemos que si ν es una distribución de probabilidad sobre los enteros no negativos, entonces un proceso discreto $(S_n, n \geq 0)$ es una caminata aleatoria con distribución de salto ν si es posible escribir

$$S_n = Y_1 + \cdots + Y_n,$$

donde Y_1, Y_2, \dots, Y_n son v.a.i.i.d. con distribución ν .

Demostración. Si (X_0, \dots, X_T) denota a las trayectorias de Lukasiewicz del árbol θ donde $X_k = (K_{U_1} - 1) + \cdots + (K_{U_k} - 1)$, fácilmente observamos que si $\ell \leq T$

$$\mathbb{P}(X_\ell - X_{\ell-1} = k) = \mathbb{P}(K_\ell = k+1) = \mu(k+1) = \mathbb{P}(S_\ell = k).$$

Finalmente tendremos que

$$\begin{aligned} T &= \inf\{k \geq 1 : M_1 + \cdots + M_k < k\} \\ &\stackrel{(d)}{=} \inf\{k \geq 1 : (K_{U_1} - 1) + \cdots + (K_{U_k} - 1) < 0\} \\ &= \inf\{k \geq 1 : X_k < 0\} = \#(\theta). \end{aligned}$$

Con esto finalizamos la prueba del Corolario. \square

4.3. Convergencia al Movimiento Browniano

Nuestra meta ahora será mostrar que las funciones de altura (o funciones contorno) de árboles de Galton-Watson (respectivamente de bosques de Galton-Watson) converge en distribución, módulo un reescalamiento apropiado, hacia la excursión Browniana (respectivamente movimiento Browniano reflejado).

Fijamos entonces una distribución crítica de hijos μ con varianza finita $\sigma^2 > 0$. Nótese que ahora tenemos que

$$\sum_{k=0}^{\infty} k\mu(k) = 1$$

Sean $\theta_1, \theta_2, \dots$ una sucesión de árboles μ -Galton-Watson independientes. A cada θ_i lo asociamos con su función altura $(h_{\theta_i}(n), 0 \leq n \leq \#(\theta_i) - 1)$.

Definición 69. (Proceso de Altura)

El proceso de altura $(H_n, n \geq 0)$ del bosque $\theta_1, \theta_2, \dots$, se define como la concatenación de las funciones $h_{\theta_1}, h_{\theta_2}, \dots$, esto es:

$$H_n = h_{\theta_i}(n - [\#(\theta_1) + \dots + \#(\theta_{i-1})])$$

si $\#(\theta_1) + \dots + \#(\theta_{i-1}) \leq n < \#(\theta_1) + \dots + \#(\theta_i)$.

Claramente la función $(H_n, n \geq 0)$ determina la sucesión de árboles.

Los valores de H entre su k -ésimo cero y el siguiente, a lo cual nos referimos como la “ k -ésima excursión” de H desde el 0, es simplemente la función altura del k -ésimo árbol en la sucesión.

Combinando el Corolario 68 y la Proposición 63, obtenemos el siguiente resultado

Proposición 70. *Para todo $n \geq 0$*

$$H_n = \#\{k \in \{0, 1, \dots, n-1\} : S_k = \inf_{k \leq i \leq n} S_i\}. \quad (4.1)$$

donde $(S_n, n \geq 0)$ es una caminata aleatoria con la distribución descrita en el Corolario 68.

Este último resultado será el ingrediente principal en la demostración del siguiente teorema.

Por definición, un movimiento Browniano reflejado (comenzado en el origen) es el valor absoluto del movimiento Browniano estándar. Denotamos por $[x]$ a la parte entera de x .

Teorema 71. *Sea $\theta_1, \theta_2, \dots$ una sucesión de árboles μ -Galton-Watson críticos e independientes, y sea $(H_n, n \geq 0)$ su proceso de altura asociado. Entonces*

$$\left(\frac{1}{\sqrt{n}} H_{[nt]}, t \geq 0 \right) \xrightarrow[n \rightarrow \infty]{\mathfrak{D}} \left(\frac{2}{\sigma} \beta_t, t \geq 0 \right),$$

donde β es un movimiento Browniano reflejado. La convergencia se satisface débilmente sobre el espacio de Skorokhod $D(\mathbb{R}_+, \mathbb{R}_+)$.

Demostración. Primero establezcamos la convergencia débil de las distribuciones finito-dimensionales.

Sea $S = (S_n, n \geq 0)$ como en la Proposición 70, es decir una caminata aleatoria con valor inicial S_0 sobre \mathbb{Z} con distribución de salto $\nu(k) = \mu(k+1)$ para $k \geq -1$, donde μ es la distribución (crítica) de un árbol aleatorio μ -Galton-Watson. Nótese entonces que la distribución de salto ν tiene media 0 y varianza finita σ^2 , ya que respectivamente tendremos que

$$\begin{aligned} \sum_{k=-1}^{\infty} \nu(k)k &= \sum_{k=0}^{\infty} \mu(k)(k-1) = \sum_{k=0}^{\infty} \mu(k)k - \sum_{k=0}^{\infty} \mu(k) = 1 - 1 = 0, \\ \sum_{k=-1}^{\infty} \nu(k)k^2 &= \sum_{k=0}^{\infty} \mu(k)(k-1)^2 \leq \sum_{k=0}^{\infty} \mu(k)k^2 < \infty. \end{aligned}$$

De esto último y por la Ley Débil de los Grandes Números tendremos que $S_n/n \xrightarrow{\mathbb{P}} 0$, cuando $n \rightarrow \infty$ lo que implica que la caminata aleatoria S es recurrente (Véase [5]; Teorema de Chung-Fuchs). También introducimos la notación

$$M_n = \sup_{0 \leq k \leq n} S_k, \quad I_n = \inf_{0 \leq k \leq n} S_k.$$

Por el Teorema de Invarianza de Donsker se tiene que

$$\left(\frac{1}{\sqrt{n}} S_{[nt]}, t \geq 0 \right) \xrightarrow[n \rightarrow \infty]{\mathfrak{D}} \left(\sigma B_t, t \geq 0 \right) \quad (4.2)$$

donde B es un movimiento Browniano estándar.

Para todo $n \geq 0$, incorporamos la caminata aleatoria retornada en el tiempo \hat{S}^n definida por

$$\hat{S}_k^n = S_n - S_{(n-k)^+}$$

y notamos que por la homogeneidad espacial de la caminata aleatoria S , el proceso $\hat{S}^n = (\hat{S}_k^n, 0 \leq k \leq n)$ tiene la misma distribución que $S^n = (S_k, 0 \leq k \leq n)$. Ahora bien, establecemos para cualquier trayectoria discreta $\omega = (\omega_1, \omega_2, \dots)$ que

$$\Psi_n(\omega) = \#\{k \in \{1, 2, \dots, n\} : \omega_k = \sup_{0 \leq j \leq k} \omega_j\}.$$

De esto último y la fórmula (4.1), observemos que

$$\begin{aligned} \Psi_n(\hat{S}^n) &= \#\{k \in \{1, \dots, n\} : \hat{S}_k^n = \sup_{0 \leq i \leq k} \hat{S}_i^n\} \\ &= \#\{k \in \{1, \dots, n\} : S_n - S_{(n-k)^+} = \sup_{0 \leq i \leq k} S_n - S_{(n-i)^+}\} \\ &= \#\{k \in \{1, \dots, n\} : S_{(n-k)^+} = \inf_{-k \leq -i \leq 0} S_{(n-i)^+}\} \\ &= \#\{n-k \in \{0, \dots, n-1\} : S_{(n-k)^+} = \inf_{n-k \leq n-i \leq n} S_{(n-i)^+}\} = H_n, \end{aligned}$$

en donde la última igualdad ocurre considerando $r = n - k$ y $j = n - i$. También establecemos

$$E_n = \Psi_n(S) = \#\{k \in \{1, \dots, n\} : S_k = M_k\}.$$

El único objetivo del siguiente lema es permitirnos probar un segundo lema; necesario para continuar con la prueba del teorema.

Lema 72. *Definamos una sucesión de tiempos de paro $(T_j)_{j \in \mathbb{N}}$ inductivamente haciendo*

$$T_0 = 0 \quad \text{y} \quad T_j = \inf\{n > T_{j-1} : S_n = M_n\} \quad \text{para toda } j \geq 1.$$

Entonces las variables aleatorias $S_{T_j} - S_{T_{j-1}}$, $j = 1, 2, \dots$ son independientes e idénticamente distribuidas, con distribución

$$\mathbb{P}[S_{T_1} = k] = \nu([k, \infty)), \quad k \geq 0.$$

Demostración. El que los incrementos $S_{T_j} - S_{T_{j-1}}$, para $j \in \mathbb{N}$ sean independientes e idénticamente distribuidos es consecuencia directa de la propiedad fuerte de Markov. Nos falta únicamente obtener la distribución de S_{T_1} .

La medida invariante de la caminata aleatoria recurrente S , es la medida de conteo sobre \mathbb{Z} . Es un resultado estándar, que si $R_0 = \inf\{n \geq 1 : S_n = 0\}$, entonces para toda $i \in \mathbb{Z}$,

$$\mathbb{E} \left[\sum_{n=0}^{R_0-1} \mathbf{1}_{\{S_n=i\}} \right] = 1.$$

Nótese que $T_1 \leq R_0$ y que la caminata aleatoria toma valores positivos en (T_1, R_0) . Se sigue de esto último, que para todo $i \leq 0$

$$\mathbb{E} \left[\sum_{n=0}^{T_1-1} \mathbf{1}_{\{S_n=i\}} \right] = 1.$$

Por lo cual, para cualquier función $g : \mathbb{Z} \rightarrow \mathbb{Z}_+$

$$\begin{aligned} \mathbb{E} \left[\sum_{n=0}^{T_1-1} g(S_n) \right] &= \sum_{i=0}^{\infty} \mathbb{E} \left[\sum_{n=0}^{T_1-1} g(i) \mathbf{1}_{\{S_n=i\}} \right] \\ &= \sum_{i=0}^{\infty} g(i) \mathbb{E} \left[\sum_{n=0}^{T_1-1} \mathbf{1}_{\{S_n=i\}} \right] = \sum_{i=0}^{\infty} g(i). \end{aligned} \quad (4.3)$$

Luego, para cualquier función $f : \mathbb{Z} \rightarrow \mathbb{Z}_+$

$$\begin{aligned} \mathbb{E}[f(S_{T_1})] &= \mathbb{E} \left[\sum_{k=0}^{\infty} \mathbf{1}_{\{k < T_1\}} f(S_{k+1}) \mathbf{1}_{\{S_{k+1} \geq 0\}} \right] = \sum_{k=0}^{\infty} \mathbb{E} \left[\mathbf{1}_{\{k < T_1\}} f(S_{k+1}) \mathbf{1}_{\{S_{k+1} \geq 0\}} \right] \\ &= \sum_{k=0}^{\infty} \mathbb{E} \left[\mathbf{1}_{\{k < T_1\}} \mathbb{E} [f(S_{k+1}) \mathbf{1}_{\{S_{k+1} \geq 0\}} | S_{k+1} - S_k = j] \right] \\ &= \sum_{k=0}^{\infty} \mathbb{E} \left[\mathbf{1}_{\{k < T_1\}} \sum_{j=0}^{\infty} \nu(j) f(S_k + j) \mathbf{1}_{\{S_k + j \geq 0\}} \right] \\ &= \sum_{j=0}^{\infty} \nu(j) \mathbb{E} \left[\sum_{k=0}^{\infty} \mathbf{1}_{\{k < T_1\}} f(S_k + j) \mathbf{1}_{\{S_k + j \geq 0\}} \right] \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \nu(j) f(i + j) \mathbf{1}_{\{i+j \geq 0\}} \\ &= \sum_{m=0}^{\infty} f(m) \sum_{j=m}^{\infty} \nu(j) \\ &= \sum_{m=0}^{\infty} f(m) \nu([m, \infty)), \end{aligned}$$

lo cual nos da la distribución de S_{T_1} en el caso en que $f(x) := \mathbf{1}_{\{k\}}(x)$, para cada $k \in \mathbb{N}$. En la tercera igualdad se utilizó la propiedad de Markov al tiempo k y en la quinta igualdad se aplicó (4.3). \square

Notemos ahora que la distribución de S_{T_1} tiene primer momento finito:

$$\mathbb{E}[S_{T_1}] = \sum_{k=0}^{\infty} k\nu([k, \infty)) = \sum_{k=0}^{\infty} \frac{k(k+1)}{2} \nu(k) = \frac{\sigma^2}{2}.$$

Recordemos que

$$E_n = \Psi_n(S) = \#\{k \in \{1, \dots, n\} : S_k = M_k\} \quad \text{y} \quad \Psi_n(\hat{S}^n) = H_n.$$

El siguiente lema es central para la primera parte de la prueba.

Lema 73. *Tenemos la siguiente convergencia en probabilidad*

$$\frac{H_n}{S_n - I_n} \xrightarrow[n \rightsquigarrow \infty]{\mathbb{P}} \frac{2}{\sigma^2}$$

Demostración. En primera instancia, consideramos los índices j para los cuales los tiempos de paro T_j definidos anteriormente, cumplan que $0 \leq T_j \leq n$ y observemos que la cantidad de índices debe ser finita, ya que dichos tiempos de paro toman únicamente valores enteros, por lo que $E_n = \#\{j \in \{0, 1, \dots, n\} : S_j = M_j\} < +\infty$. Afirmamos que se tienen las siguientes igualdades,

$$M_n = \sup_{0 \leq k \leq n} S_k = \sum_{T_j \leq n} (S_{T_j} - S_{T_{j-1}}) = \sum_{j=1}^{E_n} (S_{T_j} - S_{T_{j-1}}).$$

Notemos que en la segunda igualdad tenemos una suma telescópica, por lo que basta ver que efectivamente que $M_n = S_{T_{E_n}}$. Supongamos que existe un índice $T_{E_n} < j' < n$ tal que $S_{j'} = M_n$, entonces $S_{j'} = M_{j'}$, por lo que se tendría que $j' = T_{E_n+1} < n$ contradiciendo la definición de E_n .

Ahora bien, del Lema 72 sabemos que los incrementos $S_{T_j} - S_{T_{j-1}}$, son independientes e idénticamente distribuidos con media $\mathbb{E}[S_{T_1}]$; utilizando la ley fuerte de los grandes números, tendremos que (Note $E_n \rightarrow \infty$ cuando $n \rightarrow \infty$)

$$\frac{M_n}{E_n} \xrightarrow[n \rightsquigarrow \infty]{c.s.} \mathbb{E}[S_{T_1}] = \frac{\sigma^2}{2}.$$

Reemplazando S con la caminata aleatoria retornada en el tiempo \hat{S}^n y recordando la definición de E_n , vemos que para toda n ,

$$(M_n, E_n) \stackrel{(d)}{=} \left(\sup_{0 \leq k \leq n} S_n - S_{(n-k)^+}, \Psi_n(\hat{S}^n) \right) \stackrel{(d)}{=} (S_n - I_n, H_n)$$

En consecuencia la convergencia previa implica

$$\frac{S_n - I_n}{H_n} \xrightarrow[n \rightsquigarrow \infty]{\mathbb{P}} \frac{\sigma^2}{2}$$

de donde claramente por el Corolario 27 se sigue el resultado deseado. \square

De (4.2) y por la continuidad de la funcional ínfimo sobre el espacio de Skorohod, tenemos que para cualquier elección de puntos $0 \leq t_1 \leq t_2 \leq \dots \leq t_m$,

$$\frac{1}{\sqrt{p}} (S_{[pt_1]} - I_{[pt_1]}, \dots, S_{[pt_m]} - I_{[pt_m]}) \xrightarrow[p \rightsquigarrow \infty]{\mathfrak{D}} \sigma (B_{t_1} - \inf_{0 \leq s \leq t_1} B_s, \dots, B_{t_m} - \inf_{0 \leq s \leq t_m} B_s).$$

Se sigue por el Lema 73 que

$$\frac{1}{\sqrt{p}} (H_{[pt_1]}, \dots, H_{[pt_m]}) \xrightarrow[p \rightsquigarrow \infty]{\mathfrak{D}} \frac{2}{\sigma} (B_{t_1} - \inf_{0 \leq s \leq t_1} B_s, \dots, B_{t_m} - \inf_{0 \leq s \leq t_m} B_s).$$

Sin embargo, por el teorema de caracterización de Levy, tenemos que el proceso

$$\beta_t \equiv B_t - \inf_{0 \leq s \leq t} B_s$$

es un movimiento browniano reflejado. Con esto concluimos la prueba de convergencia débil sobre el espacio de Skorohod de las marginales finito dimensionales del Teorema 71.

Finalmente para probar la convergencia débil, tenemos que probar que la colección de variables aleatorias definidas por

$$\frac{1}{\sqrt{n}} H_{[nt]}, \quad \forall n \in \mathbb{N}$$

son tensas. Para este fin, daremos detalles del argumento que se utilizará en el caso en que la distribución μ tiene momentos exponenciales pequeños, es decir que existe un $\lambda > 0$ tal que

$$\sum_{k=0}^{\infty} e^{\lambda k} \mu(k) < \infty.$$

Enunciamos un lema que será probado al completar la prueba del presente teorema.

Lema 74. Sea $\epsilon \in (0, \frac{1}{4})$. Es posible encontrar una $\hat{\epsilon} > 0$ y un entero $N \geq 1$ tal que, para todo $n \geq N$ y $\ell \in \{0, 1, 2, \dots, n\}$

$$\mathbb{P}\left(\left|M_\ell - \frac{\sigma^2}{2}E_\ell\right| > n^{\frac{1}{4}+\epsilon}\right) < \exp\{-n^{\hat{\epsilon}}\}.$$

Aplicamos este lema con $\epsilon = 1/8$. Ya que para toda n , se tiene que (M_n, E_n) se distribuye de la misma manera que $(S_n - I_n, H_n)$, se sigue que para toda n suficientemente grande y $\ell \in \{0, 1, \dots, n\}$,

$$\mathbb{P}\left(\left|S_\ell - I_\ell - \frac{\sigma^2}{2}H_\ell\right| > n^{\frac{3}{8}}\right) < \exp(-n^{\hat{\epsilon}})$$

por lo que

$$\begin{aligned} \mathbb{P}\left(\sup_{0 \leq \ell \leq n} \left|S_\ell - I_\ell - \frac{\sigma^2}{2}H_\ell\right| > n^{\frac{3}{8}}\right) &= \mathbb{P}\left(\bigcup_{\ell=1}^n \left|S_\ell - I_\ell - \frac{\sigma^2}{2}H_\ell\right| > n^{\frac{3}{8}}\right) \\ &\leq \sum_{\ell=1}^n \mathbb{P}\left(\left|S_\ell - I_\ell - \frac{\sigma^2}{2}H_\ell\right| > n^{\frac{3}{8}}\right) < \sum_{\ell=1}^n \exp(-n^{\hat{\epsilon}}). \end{aligned}$$

y en consecuencia

$$\mathbb{P}\left(\sup_{0 \leq \ell \leq n} \left|S_\ell - I_\ell - \frac{\sigma^2}{2}H_\ell\right| > n^{\frac{3}{8}}\right) < n \exp(-n^{\hat{\epsilon}}).$$

Sea $A \geq 1$ un entero fijo. Deducimos de la cota anterior que, para toda p suficientemente grande,

$$\mathbb{P}\left(\sup_{0 \leq t \leq A} \left|S_{[pt]} - I_{[pt]} - \frac{\sigma^2}{2}H_{[pt]}\right| > (pA)^{\frac{3}{8}}\right) < pA \exp(-(Ap)^{\hat{\epsilon}}). \quad (4.4)$$

Observemos que

$$\lim_{p \rightarrow \infty} \frac{(p+1)e^{-(A(p+1))^{\hat{\epsilon}}}}{pe^{-(Ap)^{\hat{\epsilon}}}} = \lim_{p \rightarrow \infty} \left(1 + \frac{1}{p}\right) e^{-A^{\hat{\epsilon}}[(p+1)^{\hat{\epsilon}} - p^{\hat{\epsilon}}]} = 0 < 1,$$

por lo que la serie

$$\sum_{p=1}^{\infty} Ap \exp(-(Ap)^{\hat{\epsilon}}) < +\infty.$$

De lo anterior y gracias al lema de Borel-Cantelli, obtenemos que

$$\sup_{0 \leq t \leq A} \left| \frac{S_{[pt]} - I_{[pt]}}{\sqrt{p}} - \frac{\sigma^2 H_{[pt]}}{2\sqrt{p}} \right| \xrightarrow[p \rightarrow \infty]{c.s.} 0.$$

Finalmente observemos que por el Teorema 23, el teorema se sigue de la convergencia

$$\left(\frac{1}{\sqrt{p}} (S_{[pt]} - I_{[pt]}), t \geq 0 \right) \xrightarrow[p \rightarrow \infty]{\mathfrak{D}} \left(\sigma (B_t - \inf_{0 \leq s \leq t} B_s), t \geq 0 \right).$$

el cual es consecuencia inmediata de (4.2) y el Corolario 27. \square

Aún debemos probar el Lema 74. Antes, enunciamos un sencillo lema de “desviaciones moderadas” para sumas de variables aleatorias independientes.

Lema 75. *Sea $(Y_n)_{n \in \mathbb{N}}$ una sucesión de variables aleatorias reales independientes e idénticamente distribuidas. Supongamos que existe $\lambda > 0$ tal que*

$$\mathbb{E}[\exp(\lambda |Y_1|)] < \infty \quad y \quad \mathbb{E}[Y_1] = 0.$$

Entonces, para todo $\alpha > 0$, es posible escoger N suficientemente grande de manera que para toda $n \geq N$ y $\ell \in \{1, 2, \dots, n\}$,

$$\mathbb{P}(|Y_1 + Y_2 + \dots + Y_\ell| > n^{\frac{1}{2} + \alpha}) \leq \exp(-n^{\alpha/2}).$$

Demostración. Nuestras suposiciones en el lema implican que

$$\mathbb{E}[e^{\lambda Y_1}] = \mathbb{E} \left[\sum_{k=0}^{\infty} \frac{(\lambda Y_1)^k}{k!} \right] = 1 + c\lambda^2 + O[\lambda^2], \quad \text{conforme } \lambda \rightarrow 0,$$

donde $c = \frac{1}{2} \text{Var}(Y_1)$. Por lo tanto, es posible encontrar una constante C tal que para toda $\lambda > 0$ suficientemente pequeña,

$$\mathbb{E}[e^{\lambda Y_1}] \leq e^{C\lambda^2}.$$

Se sigue, que para toda $\lambda > 0$ suficientemente pequeña,

$$\mathbb{P}(Y_1 + Y_2 + \dots + Y_\ell > n^{\frac{1}{2} + \alpha}) \leq e^{-\lambda n^{\frac{1}{2} + \alpha}} \mathbb{E}[e^{\lambda(Y_1 + \dots + Y_\ell)}] \leq e^{-\lambda n^{\frac{1}{2} + \alpha}} e^{Cn\lambda^2}.$$

Si n es suficientemente grande podemos tomar $\lambda = n^{-1/2}$ y el resultado deseado se sigue (después de reemplazar Y_i con $-Y_i$). \square

Demostración del Lema 74.

Escogemos $\alpha \in (0, \frac{\epsilon}{2})$ y para simplificar la notación hacemos $m_n = \lceil n^{\frac{1}{2}+\alpha} \rceil$, entonces para todo $\ell \in \{0, 1, 2, \dots, n\}$,

$$\mathbb{P}\left(\left|M_\ell - \frac{\sigma^2}{2}E_\ell\right| > n^{\frac{1}{4}+\epsilon}\right) \leq \mathbb{P}(E_\ell > m_n) + \mathbb{P}\left(\left|M_\ell - \frac{\sigma^2}{2}E_\ell\right| > n^{\frac{1}{4}+\epsilon}; E_\ell \leq m_n\right). \quad (4.5)$$

Recordando que $M_n = \sum_{i=0}^{E_n} (S_{T_i} - S_{T_{i-1}})$, tenemos que

$$\begin{aligned} & \mathbb{P}\left(\left|M_\ell - \frac{\sigma^2}{2}E_\ell\right| > n^{\frac{1}{4}+\epsilon}; E_\ell \leq m_n\right) \\ & \leq \mathbb{P}\left(\sup_{0 \leq k \leq m_n} \left|\sum_{i=0}^k (S_{T_i} - S_{T_{i-1}}) - \frac{\sigma^2}{2}k\right| > n^{\frac{1}{4}+\epsilon}\right) \\ & \leq \mathbb{P}\left(\sup_{0 \leq k \leq m_n} \left|\sum_{i=0}^k (S_{T_i} - S_{T_{i-1}}) - \frac{\sigma^2}{2}k\right| > m_n^{\frac{1}{2}+\epsilon}\right) \leq m_n \exp(-m_n^{\epsilon/2}), \end{aligned}$$

donde la última cota existe para n suficientemente grande por el Lema 75. Recordemos que consideramos que μ tiene momentos exponenciales pequeños; y lo mismo es válido para la ley de S_{T_1} por el Lema 72, el cual nos permite aplicar el Lema 75.

Nos falta acotar el término $\mathbb{P}(E_\ell > m_n)$ de la suma en la desigualdad (4.5). Llanamente,

$$\mathbb{P}(E_\ell > m_n) \leq \mathbb{P}(E_n > m_n) \leq \mathbb{P}(S_{T_{m_n}} \leq M_n),$$

donde la primera desigualdad se debe a que $\ell \leq n$ y la segunda a que

$$S_{T_{m_n}} = M_{m_n} \quad \text{y} \quad m_n < E_n < n \quad \text{para cada } n \in \mathbb{N},$$

de donde concluimos que $S_{T_{m_n}} \leq M_n$ ya que $M_{m_n} \leq M_{E_n} \leq M_n$; entonces

$$\mathbb{P}(E_\ell > m_n) \leq \mathbb{P}(S_{T_{m_n}} \leq n^{\frac{1}{2}+\frac{\alpha}{2}}) + \mathbb{P}(M_n > n^{\frac{1}{2}+\frac{\alpha}{2}}).$$

Observemos que para n grande tendremos que

$$\mathbb{P}(M_n > n^{\frac{1}{2}+\frac{\alpha}{2}}) \leq n \sup_{1 \leq \ell \leq n} \mathbb{P}(S_\ell > n^{\frac{1}{2}+\frac{\alpha}{2}}) \leq n \exp(n^{-\frac{\alpha}{4}}),$$

donde la primera desigualdad se debe a que podemos escribir

$$\begin{aligned} \mathbb{P}(M_n > n^{\frac{1}{2} + \frac{\alpha}{2}}) &= \mathbb{P}\left(\sup_{0 \leq \ell \leq n} S_\ell > n^{\frac{1}{2} + \frac{\alpha}{2}}\right) \\ &= \mathbb{P}\left(\bigcup_{\ell=1}^n \{S_\ell > n^{\frac{1}{2} + \frac{\alpha}{2}}\}\right) \\ &\leq \sum_{\ell=1}^n \mathbb{P}(S_\ell > n^{\frac{1}{2} + \frac{\alpha}{2}}) \leq n \sup_{0 \leq \ell \leq n} \mathbb{P}(S_\ell > n^{\frac{1}{2} + \frac{\alpha}{2}}), \end{aligned}$$

y la segunda es debido al Lema 75. Finalmente,

$$\mathbb{P}(S_{T_{m_n}} \leq n^{\frac{1}{2} + \frac{\alpha}{2}}) = \mathbb{P}(S_{T_{m_n}} - \frac{\sigma^2}{2}m_n \leq n^{\frac{1}{2} + \frac{\alpha}{2}} - \frac{\sigma^2}{2}m_n)$$

y ya que $S_{T_{m_n}} - \frac{\sigma^2}{2}m_n$ es la suma de m_n variables aleatorias i.i.d. centradas con momentos exponenciales pequeños, es posible aplicar una vez más el Lema 75 para obtener la cota buscada. Esto completa la prueba del Lema 74. \square

4.4. Algunas Aplicaciones

Ahora estamos interesados en utilizar el resultado de convergencia obtenido en la sección anterior para una aplicación a los procesos que definiremos en breve como excursión browniana y tiempo local del movimiento browniano.

4.4.1. Tiempo Local y Excursión Browniana

Sea $B = (B_t)_{t \geq 0}$ un movimiento browniano real estándar (comenzado en el origen). El proceso definido por $\beta_t = |B_t|$ es llamado un movimiento browniano reflejado.

Definición 76. (Tiempo Local del Movimiento Browniano)

Existe un proceso continuo y creciente que denotamos por $(L_t^0)_{t \geq 0} = (L_t^0(B))_{t \geq 0}$, llamado el “Tiempo Local del proceso B en 0”, el cual es posible definir por la siguiente aproximación:

$$L_t^0 = \lim_{\epsilon \searrow 0} \frac{1}{2\epsilon} \int_0^t \mathbf{1}_{[-\epsilon, \epsilon]}(B_s) ds = \lim_{\epsilon \searrow 0} \frac{1}{2\epsilon} \int_0^t \mathbf{1}_{[0, \epsilon]}(\beta_s) ds, \quad \forall t \geq 0 \quad c.s.$$

La continuidad del proceso que se incluye en la definición es posible obtenerla gracias al criterio de continuidad de Kolmogorov.

Una definición alternativa del Tiempo Local, cuya utilidad reside en ser un poco más interpretativa, considera una variable que escribiremos $N_\epsilon(t)$ y que denota el número de excursiones positivas de B lejos del 0 con altura mayor que ϵ y completado antes del tiempo t . Entonces tendremos que

$$L_t^0 = \lim_{\epsilon \searrow 0} 2\epsilon N_\epsilon(t) \quad \forall t \geq 0 \quad c.s.$$

Entonces el proceso $(L_t^0)_{t \geq 0}$ es continuo y creciente; más aún, el conjunto de puntos donde crece la función

$$t \mapsto L_t^0,$$

que denotamos $\{t \geq 0 : L_t^0 - L_{t-}^0 > 0\}$, coincide c.s. con el conjunto

$$\mathcal{Z} = \{t \geq 0 : \beta_t = 0\}$$

de ceros del proceso β y el cual es fácil observar que al estar definido como el valor absoluto B , igualmente \mathcal{Z} coincidirá con el conjunto de ceros del movimiento browniano; esto es, el soporte topológico de la medida $d_t L_t^0$ coincide c.s. con el conjunto $\{t \geq 0 : B_t = 0\}$; en consecuencia, $\mathcal{Z} = B^{-1}(\{0\})$ es un conjunto cerrado c.s.

Introducimos la “*Inversa Continua por la Derecha*” del proceso de Tiempo Local,

$$\sigma_\ell := \inf\{t \geq 0 : L_t^0 > \ell\}, \quad \text{para toda } \ell \geq 0.$$

Tendremos entonces que

$$\mathcal{Z} = \{\sigma_\ell : \ell \geq 0\} \cup \{\sigma_{\ell-} : \ell \in D_\sigma\},$$

donde D_σ es el conjunto (numerable) de discontinuidades de la aplicación $\ell \mapsto \sigma_\ell$.

Definición 77. (Intervalo de Excursión Browniana)

A cada componente del conjunto abierto $\mathbb{R}_+ \setminus \mathcal{Z}$, se le conoce como Intervalo de Excursión del movimiento browniano lejos del 0. Estos intervalos coinciden c.s. con los intervalos de excursión de β lejos del 0; más aún, estos intervalos son exactamente

$$(\sigma_{\ell-}, \sigma_\ell), \quad \text{para } \ell \in D_\sigma.$$

A partir de esto último podemos definir lo siguiente.

Definición 78. (Excursión Browniana)

Para cada $\ell \in D_\sigma$, definimos la excursión $e_\ell = (e_\ell(t))_{t \geq 0}$ asociada al intervalo $(\sigma_{\ell-}, \sigma_\ell)$ por:

$$e_\ell(t) = \begin{cases} \beta_{\sigma_{\ell-}+t} & \text{si } 0 \leq t \leq \sigma_\ell - \sigma_{\ell-}, \\ 0 & \text{si } t > \sigma_\ell - \sigma_{\ell-}. \end{cases}$$

Se puede ver a e_ℓ como un elemento del espacio de excursiones \mathbf{E} , definido por

$$\mathbf{E} = \{e \in C(\mathbb{R}_+, \mathbb{R}_+) : e(0) = 0 \text{ y } \sigma(e) = \sup \{s \geq 0 : e_s > 0\} < (0, +\infty)\},$$

y $\sup \{\emptyset\} = 0$ por convención. Notemos que excluimos la función constante cero al pedir $\sigma(e) > 0$. El espacio E debe considerarse equipada con la métrica d dada por:

$$d(e, e') = \sup_{t \geq 0} |e(t) - e'(t)| + |\sigma(e) - \sigma(e')|$$

y la σ -álgebra de Borel asociada. Observemos que $\sigma(e_\ell) = \sigma_\ell - \sigma_{\ell-}$ para toda $\ell \in D$.

El siguiente teorema debido a Levy será de gran utilidad.

Teorema 79. (Teorema de Lévy)

Sea B un movimiento browniano lineal estándar. Sea $(L_t^0, t \geq 0)$ su tiempo local al tiempo cero, entonces

$$(L_t^0, t \geq 0) \stackrel{(d)}{=} \left(\sup_{0 \leq s \leq t} B_s, t \geq 0 \right).$$

Más aún, este resultado puede ser extendido de forma que

$$\left(B_t - \inf_{0 \leq s \leq t} B_s, \sup_{0 \leq s \leq t} B_s \right) \stackrel{(d)}{=} (|B_t|, L_t^0) \quad \text{para toda } t \geq 0.$$

Para una prueba consúltese [17] Teorema VI (2.3).

4.4.2. Convergencia a la Excursión Browniana

Manteniendo la notación de la sección anterior, para una sucesión de árboles μ -Galton-Watson $(\theta_i)_{i \in \mathbb{N}}$ y considerando al proceso altura asociado H , establecemos para $n \geq 0$,

$$\Lambda_n = k \quad \text{sii} \quad \#(\theta_1) + \cdots + \#(\theta_{k-1}) \leq n < \#(\theta_1) + \cdots + \#(\theta_k)$$

de manera que k tiene que ser el índice del árbol para el cual el n -ésimo vértice visitado por proceso de altura pertenece a la gráfica de la función de altura h_{θ_k} .

Por el Teorema de Levy, la convergencia del Teorema 71 puede ser extendida de la siguiente manera:

$$\left(\frac{1}{\sqrt{n}} H_{[nt]}, \frac{1}{\sqrt{n}} \Lambda_{[nt]}; t \geq 0 \right) \xrightarrow[n \rightarrow \infty]{\mathfrak{D}} \left(\frac{2}{\sigma} |B_t|, \sigma L_t^0; t \geq 0 \right). \quad (4.6)$$

Esto es consecuencia de los argumentos que mencionamos a continuación: No es difícil ver que

$$\Lambda_n = 1 - \inf_{0 \leq j \leq n} S_j = 1 - I_n.$$

Por otro lado, se probó, que para toda $A > 0$,

$$\sup_{0 \leq t \leq A} \left| \frac{S_{[nt]} - I_{[nt]}}{\sqrt{n}} - \frac{\sigma^2 H_{[nt]}}{2 \sqrt{n}} \right| \xrightarrow[n \rightarrow \infty]{c.s.} 0.$$

Esto nos permite afirmar, combinando el resultado de la sección anterior, el Teorema de Donsker y el Teorema 23 que

$$\left(\frac{1}{\sqrt{n}} I_{[nt]}, t \geq 0 \right) \xrightarrow[n \rightarrow \infty]{\mathfrak{D}} \left(\sigma B_t - \sigma (B_t - \inf_{0 \leq s \leq t} B_s), t \geq 0 \right) \stackrel{(d)}{=} \left(\sigma \sup_{0 \leq s \leq t} B_s, t \geq 0 \right),$$

de donde obtenemos que

$$\left(\frac{1}{\sqrt{n}} H_{[nt]}, \frac{1}{\sqrt{n}} \Lambda_{[nt]}; t \geq 0 \right) \xrightarrow[n \rightarrow \infty]{\mathfrak{D}} \left(\frac{2}{\sigma} (B_t - \inf_{0 \leq s \leq t} B_s), \sigma \sup_{0 \leq s \leq t} B_s; t \geq 0 \right).$$

Aplicando el Teorema de Levy mencionado en la subsección anterior da como resultado (4.6).

Ahora ocuparemos (4.6) para el estudio de las asintóticas de un árbol μ -Galton-Watson condicionado a ser grande. Sea $\theta \in \mathcal{A}$ escribimos

$$h(\theta) = \sup\{|v| : v \in \theta\},$$

para denotar la altura del árbol θ . Fijamos $x > 0$ y para todo $n \geq 1$ denotamos por $\theta^{\{x\sqrt{n}\}}$ al árbol con distribución

$$\Pi_\mu(da | h(a) \geq x\sqrt{n}),$$

donde recordemos que Π_μ es la ley de un árbol Galton-Watson con distribución de hijos μ . Así mismo, denotamos por $H^{\{x\sqrt{n}\}}$ la función altura de $\theta^{\{x\sqrt{n}\}}$. Por convención,

$$H_m^{\{x\sqrt{n}\}} = 0, \quad \text{si } m \geq \#(\theta^{\{x\sqrt{n}\}}).$$

El siguiente resultado es un corolario del Teorema 71 extendido con ayuda del Teorema de Levy como ya mencionamos.

Corolario 80. *Tenemos que*

$$\left(\frac{1}{\sqrt{p}} H_{[pt]}^{\{x\sqrt{p}\}}; t \geq 0 \right) \xrightarrow[p \rightsquigarrow \infty]{\mathfrak{D}} \left(\frac{2}{\sigma} e_t^{\sigma x/2}; t \geq 0 \right)$$

donde $e_t^{\sigma x/2}$ es una excursión Browniana condicionada a tener altura mayor que $\sigma x/2$.

OBSERVACIÓN. La excursión Browniana $e_t^{\sigma x/2}$ puede ser construida explícitamente de la siguiente manera. Hacemos

$$\begin{aligned} T &= \inf \{t \geq 0 : |B_t| \geq \sigma x/2\} \\ G &= \sup \{t \leq T : B_t = 0\} \\ D &= \inf \{t \geq T : B_t = 0\}, \end{aligned}$$

donde B es un browniano estándar y entonces podemos tomar

$$e_t^{\sigma x/2} = |B_{(G+t) \wedge D}| = \beta_{(G+t) \wedge D}.$$

A continuación probamos el corolario,

Demostración. La prueba recae en el resultado (4.6). Del teorema de representación de Skorohod, podemos encontrar para todo $p \geq 1$ un proceso que denotaremos $(H_t^{(p)}, \Lambda_t^{(p)})$ tal que

$$(H_t^{(p)}, \Lambda_t^{(p)})_{t \geq 0} \stackrel{(d)}{=} \left(\frac{1}{\sqrt{p}} H_{[pt]}, \frac{1}{\sqrt{p}} \Lambda_{[pt]} \right)_{t \geq 0}$$

y

$$(H_t^{(p)}, \Lambda_t^{(p)})_{t \geq 0} \xrightarrow[p \rightsquigarrow \infty]{c.s.} \left(\frac{2}{\sigma} \beta_t, \sigma L_t^0 \right)_{t \geq 0}, \quad (4.7)$$

e uniformemente sobre todo compacto.

Hacemos

$$\begin{aligned} T^{(p)} &= \inf \{t \geq 0 : |H_t^{(p)}| \geq x\} \\ G^{(p)} &= \sup \{t \leq T^{(p)} : H_t^{(p)} = 0\} - \frac{1}{p} \\ D^{(p)} &= \inf \{t \geq T^{(p)} : H_t^{(p)} = 0\}. \end{aligned}$$

El término $-p^{-1}$ en la fórmula de $G^{(p)}$ se debe a la parte entera en $\frac{1}{\sqrt{p}}H_{[pt]}$: Queremos que el proceso $H_{G^{(p)}+t}^{(p)}$ permanezca en el 0 durante el intervalo $[0, p^{-1})$.

Por construcción $(H_{G^{(p)}+t}^{(p)}, t \geq 0)$ tiene la misma distribución que el proceso altura (reescalado) del primer árbol en la sucesión $\theta_1, \theta_2, \dots$ con altura mayor que $x\sqrt{p}$, el cual se distribuye como $\theta^{\{x\sqrt{p}\}}$. Por tanto,

$$(H_{(G^{(p)}+t) \wedge D^{(p)}}^{(p)}, t \geq 0) \stackrel{(d)}{=} \left(\frac{1}{\sqrt{p}} H_{[pt]}^{\{x\sqrt{p}\}}, t \geq 0 \right).$$

El corolario se seguirá entonces de la convergencia en (4.7) si podemos probar que

$$G^{(p)} \xrightarrow{c.s.} G \quad D^{(p)} \xrightarrow{c.s.} D.$$

Utilizando el hecho de que inmediatamente después del tiempo T el proceso $|\beta|$ toca niveles estrictamente mayores que $\sigma x/2$, obtenemos fácilmente de (4.7) que

$$T^{(p)} \xrightarrow[p \rightarrow \infty]{c.s.} T.$$

De esto último y (4.7) nuevamente se sigue que

$$\begin{aligned} \liminf D^{(p)} &\geq D \quad c.s. \\ \limsup G^{(p)} &\geq G \quad c.s. \end{aligned}$$

Probemos ahora que también tendremos que $\limsup D^{(p)} \leq D \quad c.s.$ (el mismo argumento sirve para $\liminf G^{(p)}$). Fijemos $t > 0$. De la propiedad del soporte del tiempo local, tenemos que

$$L_t^0 > L_D^0 \quad c.s. \text{ en } \{D < t\}.$$

Gracias a (4.7), obtenemos que

$$\Lambda_t^{(p)} > \sigma L_D^0 \quad \text{para } p \text{ grande, } \quad c.s. \text{ en } \{D < t\}.$$

Notamos que $L_D^0 = L_T^0$. Ya que $T^{(p)} \xrightarrow{c.s.} T$, (4.7) también nos muestra que

$$\Lambda_t^{(p)} > \Lambda_{T^{(p)}}^{(p)} \quad \text{para } p \text{ grande, } c.s. \text{ en } \{D < t\}.$$

Observando que $\Lambda^{(p)}$ se mantiene constante sobre el intervalo $[T^{(p)}, D^{(p)})$, concluimos que

$$t \geq D^{(p)} \quad \text{para } p \text{ grande, } c.s. \text{ en } \{D < t\}.$$

Con esto basta para concluir que $\limsup D^{(p)} \leq D$, finalizando así la prueba. \square

Reemplazando p por p^2 y tomando $x = 1$, deducimos del corolario (de hecho más que nada de su prueba) que

$$\frac{1}{p^2} \#(\theta^{\{p\}}) \xrightarrow[p \rightsquigarrow \infty]{\mathfrak{D}} \zeta_{\sigma/2},$$

donde $\zeta_{\sigma/2} = D - G$ es la longitud de la excursión $e^{\sigma/2}$. Efectivamente, esto se sigue de la convergencia

$$D^{(p)} - G^{(p)} \xrightarrow[p \rightsquigarrow \infty]{c.s.} D - G,$$

y el hecho de que por construcción

$$\#(\theta^{\{p\}}) \stackrel{(d)}{=} p^2(D^{(p^2)} - G^{(p^2)}) + 1$$

con notación de la prueba precedente.

Recordemos que $h(\theta) = \sup\{|v| : v \in \theta\}$ y que $H^{\{p\}}$ denota la función altura de $\theta^{\{p\}}$.

Corolario 81. *Se tiene que*

$$\frac{1}{p} h(\theta^{\{p\}}) \xrightarrow[p \rightsquigarrow \infty]{\mathfrak{D}} \frac{2}{\sigma} \sup_{t \in [0, \zeta_{\sigma/2}]} e_t^{\sigma/2}.$$

Demostración.

$$\begin{aligned} \frac{1}{p} h(\theta^{\{p\}}) &= \frac{1}{p} \sup \left\{ H_{[p^2 t]}^{\{p\}} ; 0 \leq [p^2 t] \leq \#(\theta^{\{p\}}) - 1 \right\} \\ &= \frac{1}{p} \sup \left\{ H_{[p^2 t]}^{\{p\}} ; 0 \leq [p^2 t] \leq p^2(D^{(p^2)} - G^{(p^2)}) \right\} \\ &= \sup \left\{ \frac{1}{p} H_{[p^2 t]}^{\{p\}} ; 0 \leq \frac{[p^2 t]}{p^2} \leq D^{(p^2)} - G^{(p^2)} \right\} \end{aligned}$$

donde

$$\frac{[p^2 t]}{p^2} \xrightarrow{p \rightarrow \infty} t \quad \text{y} \quad D^{(p)} - G^{(p)} \xrightarrow[p \rightarrow \infty]{c.s.} D - G = \zeta_{\sigma/2}.$$

Entonces gracias a que $h(x) = \sup_{t \geq 0} x(t)$ es una funcional continua sobre el espacio de Skorohod y a la convergencia en distribución del proceso $\left(\frac{1}{p} H_{[p^2 t]}^{(p)}, t \geq 0\right)$, por el corolario anterior obtenemos el resultado. \square

Ahora discutiremos “medidas de ocupación”. Más que considerar un solo árbol como anteriormente, estaremos interesados en un bosque finito cuyo tamaño tienda a infinito con p . Precisamente, fijamos $b > 0$, y establecemos

$$H_n^p = \begin{cases} H_n & \text{si } \Lambda_n \leq bp, \\ 0 & \text{si } \Lambda_n > bp, \end{cases}$$

de manera que H^p es el proceso de altura para una colección de $[bp]$ árboles de Galton-Watson independientes. Entonces se sigue fácilmente de (4.6) que

$$\left(\frac{1}{p} H_{[p^2 t]}^p, t \geq 0\right) \xrightarrow[p \rightarrow \infty]{\mathfrak{D}} \left(\frac{2}{\sigma} \beta_{t \wedge \tau_{b/\sigma}}, t \geq 0\right) \quad (4.8)$$

donde, recordemos que β un movimiento browniano reflejado y para toda $r > 0$,

$$\tau_r = \inf\{t \geq 0 : L_t^0 > r\} \stackrel{(d)}{=} \inf\{t \geq 0 : \beta_t > r\}.$$

Efectivamente, podremos escribir

$$\frac{1}{p} H_{[p^2 t]}^p = \frac{1}{p} H_{[p^2(t \wedge \tau_b^{(p)})]}$$

donde

$$\tau_b^{(p)} = \frac{1}{p^2} \inf\{n \geq 0 : \Lambda_n > bp\} = \inf\{t \geq 0 : \frac{1}{p} \Lambda_{[p^2 t]} > b\}.$$

Entonces podemos observar de (4.6) que

$$\left(\left(\frac{1}{p} H_{[p^2 t]}, t \geq 0\right), \tau_b^{(p)}\right) \xrightarrow[p \rightarrow \infty]{\mathfrak{D}} \left(\left(\frac{2}{\sigma} \beta_t, t \geq 0\right), \tau_{b/\sigma}\right)$$

de donde se sigue claramente (4.8).

Tomando $b = 1$, es posible deducir de (4.8) que, para toda $x > 0$,

$$\mathbb{P} \left[\sup_{1 \leq i \leq p} h(\theta_i) > px \right] \xrightarrow{p \rightarrow \infty} \mathbb{P} \left[\sup_{t \leq \tau_{1/\sigma}} \frac{2}{\sigma} \beta_t > x \right] = 1 - \exp \left(- \frac{2}{\sigma^2 x} \right).$$

La última igualdad es una consecuencia de la teoría de excursión para el movimiento Browniano real (Capítulo XII [17]). Tendremos entonces que

$$\mathbb{P} \left[\sup_{1 \leq i \leq p} h(\theta_i) > px \right] = 1 - (1 - \mathbb{P} [h(\theta) > px])^p$$

obteniendo con esto un hecho clásico de la teoría de procesos de ramificación:

$$\mathbb{P} [h(\theta) \geq n] \sim \frac{2}{n\sigma^2}, \quad \text{cuando } n \rightarrow \infty.$$

4.4.3. Convergencia a un Proceso de Difusión de Feller

Hacemos ahora $Z_0^p = p$, y para toda $n \geq 1$

$$Z_n^p = \sum_{i=1}^p \# \{u \in \theta_i : |u| = n\} = \# \{k \geq 0 : H_k^p = n\}.$$

Ya que cada θ_i es un árbol μ -Galton-Watson, de la Proposición 64 sabemos que $(Z_n^p, n \geq 0)$ es un proceso de ramificación Galton-Watson con distribución de descendencia μ . Es posible entonces aplicar la siguiente aproximación clásica por difusión a este proceso.

Teorema 82. *Se tiene que*

$$\left(\frac{1}{p} Z_{[pt]}^p, t \geq 0 \right) \xrightarrow{p \rightarrow \infty} (X_t, t \geq 0),$$

donde el proceso límite X es un proceso de difusión con generador infinitesimal

$$\frac{1}{2} \sigma^2 x \frac{d^2}{dx^2},$$

el cual puede obtenerse como la única solución a la ecuación diferencial estocástica

$$dX_t = \sigma \sqrt{X_t} dB_t, \quad X_0 = 1.$$

Una prueba a este teorema puede verse en [7] Teorema 9.1.3.

Afirmamos que

- Z_n^p es una martingala,
- $(Z_n^p)^2 - \sigma^2 \sum_{k=0}^{n-1} Z_k^p$ es una martingala,

lo cual fuertemente nos sugiere que el proceso limite X es de la forma establecida en el teorema anterior. El proceso X es llamada *Difusión ramificada de Feller*.

Fijemos ahora una distribución μ con varianza $\sigma = 2$. Sean f_1, \dots, f_r , r funciones continuas con soporte compacto de \mathbb{R}_+ en \mathbb{R}_+ . Recordemos que H^p es el proceso de altura para cierta colección de árboles Galton-Watson independientes, entonces como consecuencia de la convergencia en (4.8) obtenemos que

$$\left(\int_0^{\tau_1^{(p)}} f_i \left(\frac{1}{p} H_{[p^2 t]^p} \right) dt, 1 \leq i \leq r \right) \xrightarrow[p \rightarrow \infty]{\mathfrak{D}} \left(\int_0^{\tau_{1/2}} f_i(\beta_t) dt, 1 \leq i \leq r \right).$$

Por otro lado,

$$\int_0^{\tau_1^{(p)}} f_i \left(\frac{1}{p} H_{[p^2 t]^p} \right) dt = \frac{1}{p^2} \sum_{n=0}^{\infty} Z_n^p f_i \left(\frac{1}{p} \right) = \int_0^{\infty} da f_i \left(\frac{[pa]}{p} \right) \frac{1}{p} Z_{[pa]}^p.$$

Utilizando el teorema anterior, observamos finalmente que

$$\left(\int_0^{\tau_{1/2}} f_i(\beta_t) dt, 1 \leq i \leq r \right) \stackrel{(d)}{=} \left(\int_0^{\infty} da f_i(a) X_a, 1 \leq i \leq r \right).$$

En otras palabras, la medida de ocupación de β sobre el intervalo de tiempo $[0, \tau_{1/2}]$, es decir la medida

$$f \longrightarrow \int_0^{\tau_{1/2}} f(\beta_t) dt$$

tiene la misma distribución que la medida $X_a da$.

4.5. Árboles Galton-Watson con Progenie Fija

Del Teorema 71 es posible obtener un famoso resultado de Aldous relacionado con árboles Galton-Watson condicionados a tener una gran cantidad (fija) de vértices.

Aldous trató con la función de contorno más que con la de altura, lo cual realmente es más o menos equivalente como se verá más adelante. Asumiremos como en la parte final de la prueba del Teorema 71 que la distribución μ tiene momentos exponenciales pequeño. El resultado en verdad se sigue sin esta asunción, pero nos permite simplificar la prueba.

Para toda $p \geq 1$ denotamos por $\theta^{(p)}$ a un árbol μ -Galton-Watson condicionado con $\#(\theta) = p$, lo cual puede interpretarse como que dicho árbol tiene p vértices contando la raíz. Para que esto tenga sentido necesitamos que $\mathbb{P}(\#(\theta) = p) > 0$ para toda $p \geq 1$, lo cual ocurre si $\mu(1) > 0$.

Recordemos que la distribución de $\theta^{(p)}$ es $\Pi_\mu(da | \#(a) = p)$ y que denotamos por $(H_k^{(p)}, 0 \leq k \leq p)$ al proceso de altura de $\theta^{(p)}$, con la convención $H_p^{(p)} = 0$.

Introducimos la excursión Browniana normalizada $\mathbf{e} = (\mathbf{e}_t, 0 \leq t \leq 1)$. Esta es simplemente una excursión Browniana condicionada a tener duración 1. Por lo tanto, debemos fijarnos en la primera excursión positiva β (partiendo del 0) con duración mayor a 1, escribimos $[G, D]$ para el correspondiente intervalo de tiempo en el que ocurre, y hacemos

$$\mathbf{e}_t = \frac{1}{\sqrt{D-G}} \beta_{(G+(D-G)t) \wedge D}, \quad 0 \leq t \leq 1.$$

Por convención $\mathbf{e}_t = 0$ si $t > 1$.

Teorema 83. *Tenemos que*

$$\left(\frac{1}{\sqrt{p}} H_{[pt]}^{(p)}, 0 \leq t \leq 1 \right) \xrightarrow[p \rightarrow \infty]{\mathfrak{D}} \left(\frac{2}{\sigma} \mathbf{e}_t, 0 \leq t \leq 1 \right)$$

Este resultado claramente es similar al obtenido anteriormente en el Teorema 71 y Corolario 80. Sin embargo, debido a cierta “degeneración” en el límite (no hay excursión Browniana con duración exactamente 1), no es posible utilizar la misma estrategia de prueba.

Demostración. Sea $H = (H_n, n \geq 0)$ el proceso de altura como en el Teorema 71 asociado a una sucesión independiente de árboles μ -Galton-Watson. Ya que es posible, asumiremos que H está dado en términos de la caminata aleatoria que la define.

Denotemos por V_1 el número de vértices del primer árbol en la sucesión, es decir,

$$V_1 = \inf \{n \geq 1 : H_n = 0\} = \inf \{n \geq 1 : S_n = -1\}.$$

Un argumento de combinatoria –considerar todas las posibles permutaciones circulares de p incrementos de la caminata aleatoria S sobre el intervalo $[0, p]$ – muestra que para toda $p \geq 1$,

$$\mathbb{P}(V_1 = p) = \frac{1}{p} \mathbb{P}(S_p = -1). \quad (4.9)$$

Por otro lado, resultados clásicos para caminatas aleatorias (Véase [18] Capítulo II) nos indican que

$$\lim_{p \rightarrow \infty} \sqrt{p} \mathbb{P}(S_p = -1) = \frac{1}{\sigma \sqrt{2\pi}},$$

de donde se sigue que

$$\mathbb{P}(V_1 = p) \sim \frac{1}{\sigma \sqrt{2\pi p^3}}, \quad \text{para } p \text{ grande.} \quad (4.10)$$

De la última parte de la prueba al Teorema 71, recordemos que es posible encontrar $\epsilon > 0$ tal que, para p suficientemente grande

$$\mathbb{P} \left(\sup_{0 \leq t \leq 1} \left| \frac{H_{pt}}{\sqrt{p}} - \frac{2}{\sigma^2} \frac{S_{[pt]} - I_{[pt]}}{\sqrt{p}} \right| > p^{-1/8} \right) < \exp(-p^\epsilon).$$

Observemos que utilizando (4.10), también tendremos que para p grande

$$\mathbb{P} \left(\sup_{0 \leq t \leq 1} \left| \frac{H_{pt}}{\sqrt{p}} - \frac{2}{\sigma^2} \frac{S_{[pt]} - I_{[pt]}}{\sqrt{p}} \right| > p^{-1/8} \mid V_1 = p \right) < \exp(-p^{\epsilon'}), \quad (4.11)$$

para toda $\epsilon' < \epsilon$. Ya que $I_n = 0$ para $0 \leq n < V_1$, también tenemos que

$$\mathbb{P} \left(\sup_{0 \leq t \leq 1} \left| \frac{H_{pt}}{\sqrt{p}} - \frac{2}{\sigma^2} \frac{S_{[pt]}}{\sqrt{p}} \right| > p^{-1/8} \mid V_1 = p \right) < \exp(-p^{\epsilon'}),$$

Lo resultados (4.9), (4.10) y (4.11) pueden consultarse en [15] Lema 3.8.6. , Lema 3.8.7. y Teorema 4.4.5. respectivamente.

Ahora bien, $(H_k^{(p)}, 0 \leq k \leq p)$ se distribuye igual que $(H_k, 0 \leq k \leq p)$ bajo $\mathbb{P}(\cdot | V_1 = p)$. Tendremos entonces que el presente teorema será resultado de la cota anterior y el lema presentado a continuación que relaciona a la excursión Browniana normalizada con la excursión de una caminata aleatoria de larga duración fija. \square

Lema 84. *La distribución de el proceso*

$$\left(\frac{1}{\sigma \sqrt{p}} S_{[pt]}, 0 \leq t \leq 1 \right),$$

bajo la probabilidad condicional $\mathbb{P}(\cdot | V_1 = 0)$ converge cuando $p \rightarrow \infty$ a la ley de la excursión Browniana normalizada.

La prueba se omite, sin embargo este resultado puede ser visto como una versión condicional del Teorema de Donsker (Véase [10]).

Inmediatamente se deduce del teorema anterior que

$$\lim_{p \rightarrow \infty} \mathbb{P}(h(\theta) > x\sqrt{p} \mid \#(\theta) = p) = \mathbb{P}\left(\sup_{0 \leq t \leq 1} \mathbf{e}_t > \frac{\sigma x}{2}\right).$$

CONSECUENCIAS EN COMBINATORIA. Para varias elecciones particulares de μ , la medida $\Pi_\mu(da \mid \#(a) = p)$ coincide con la medida de probabilidad uniforme sobre cierta clase de “árboles combinatorios” con p vértices. El teorema anterior nos da información acerca de la proporción de árboles en dicha clase que satisfacen ciertas propiedades.

Siendo más explícitos, consideremos el caso en que μ es una distribución geométrica de parámetro $1/2$,

$$\mu(k) = \frac{1}{2^{k+1}}.$$

Entonces $\Pi_\mu(da \mid \#(a) = p)$ es la distribución uniforme sobre el conjunto de árboles ordenados con raíz y p -vértices \mathcal{A}_p (esto es consecuencia de la Proposición 66). El resultado anterior nos muestra que la altura de un árbol escogido aleatoriamente de \mathcal{A}_p es de orden \sqrt{p} , de manera más precisa nos da la proporción asintótica de aquellos árboles en \mathcal{A}_p con altura mayor a $x\sqrt{p}$. Argumentos similares son enunciados al utilizar otras funcionales distintas a la función altura.

Similarmente, si μ es una distribución Poisson tal que

$$\mu(k) = e^{-1} \frac{1}{k!},$$

entonces $\Pi_\mu(da \mid \#(a) = p)$ resulta ser la distribución uniforme sobre cierta clase de árboles con raíz y p -vértices, conocidos como árboles de Cayley. Los árboles de Cayley son árboles sin orden en sus vértices a los que simplemente se les asignan etiquetas enumeradas $1, 2, \dots, p$. La raíz puede ser entonces cualquiera de los p -vértices. Es un resultado conocido que habrá exactamente p^{p-2} árboles de Cayley con p -vértices, en consecuencia habrá p^{p-1} árboles de Cayley con raíz y p -vértices. De esta manera, si comenzamos a partir de un árbol ordenado con raíz con distribución de acuerdo a $\Pi_\mu(da \mid \#(a) = p)$ y posteriormente asignamos etiquetas $1, 2, \dots, p$ de manera equiprobable a los vértices, y finalmente nos olvidamos del orden con el cual comenzamos, lo que obtenemos es un árbol aleatorio que se distribuye uniformemente sobre el conjunto de árboles de Cayley con raíz y p -vértices. Por lo tanto,

los resultados de la presente sección nos estarían dado de igual manera información sobre las propiedades de árboles de Cayley condicionados a ser grandes.

Puede que sea inesperado el hecho de que de distintas clases de árboles combinatorios obtengamos el mismo tipo de comportamiento en el límite y en particular, que la ley obtenida en el límite sea la misma en cada caso. Sin embargo, obsérvese que la constante σ si puede variar: $\sigma^2 = 1$ en el caso de árboles binarios o de Cayley, mientras que $\sigma^2 = 2$ para árboles ordenados con raíz.

4.6. Convergencia de Funciones Contorno

Explicaremos ahora brevemente como es que algunos resultados precedentes también pueden ser enunciados en términos del proceso contorno de un bosque en vez del proceso de altura. Recordemos la definición de la función de contorno de un árbol enunciada en la Definición 62. Notemos que en contraste con el proceso de altura es conveniente indexar al proceso de contorno por un parámetro real.

Daremos el resultado homólogo correspondiente al Teorema 71.

Sean $\theta_1, \theta_2, \dots$ una sucesión de árboles μ -Galton-Watson independientes. A cada θ_i lo asociamos con su función contorno $C_t(\theta_i)$ para $t \in [0, \zeta(\theta_i)]$, donde $\zeta(\theta_i) = 2(\#(\theta_i) - 1)$.

Este último intervalo de definición nos da como consecuencia, que la función de contorno de un árbol que consista únicamente de la raíz es trivial. Por esta razón y ya que deseamos que al igual que el proceso de altura el proceso de contorno defina la sucesión de árboles, haremos la convención de que la función de contorno $C_t(\theta_i)$ estará definida para $t \in [0, \xi(\theta_i)]$, donde $\xi(\theta_i) = 2\#(\theta_i) - 1$; tomando $C_t(\theta_i) = 0$ para $t \in [\zeta(\theta_i), \xi(\theta_i)]$.

Definición 85. (Proceso de Contorno)

El proceso de contorno $(C_t, t \geq 0)$ del bosque $\theta_1, \theta_2, \dots$ se define como la concatenación de las funciones $C_t(\theta_1), C_t(\theta_2), \dots$, esto es:

$$C_t = C_{t - [\xi(\theta_1) + \dots + \xi(\theta_{i-1})]}(\theta_i)$$

si $\xi(\theta_1) + \dots + \xi(\theta_{i-1}) \leq t < \xi(\theta_1) + \dots + \xi(\theta_i)$.

Para toda $n \geq 0$, definimos

$$J_n = 2n - H_n + I_n.$$

Nótemos que la sucesión J_n es estrictamente creciente y que $J_n > n$.

Recordemos que el valor al tiempo n del proceso altura, corresponde a la generación del individuo visitado en el paso n , donde los individuos son visitados en orden lexicográfico un árbol tras otro. Es fácil verificar inductivamente sobre n que $[J_n, J_{n+1}]$ es exactamente el intervalo durante el cual el proceso de contorno va del individuo n al individuo $n + 1$. De esta observación, obtenemos que

$$\sup_{t \in [J_n, J_{n+1}]} |C_t - H_n| \leq |H_{n+1} - H_n| + 1.$$

Un argumento más explícito para justificar esta última cota se sigue de la fórmula para C_t en términos del proceso altura: Para $t \in [J_n, J_{n+1}]$,

$$C_t = \begin{cases} H_n - (t - J_n) & \text{si } t \in [J_n, J_{n+1} - 1] \\ (H_{n+1} - (J_{n+1} - t))^+ & \text{si } t \in [J_{n+1} - 1, J_{n+1}]. \end{cases}$$

No es difícil verificar estas fórmulas por inducción sobre n .

Definimos una función aleatoria $\varphi : \mathbb{R}_+ \rightarrow \{0, 1, \dots\}$ dada por

$$\varphi(t) = n \quad \text{si y solo si } t \in [J_n, J_{n+1}).$$

De la cota previa obtendremos que para todo entero $m \geq 1$,

$$\sup_{t \in [0, m]} |C_t - H_{\varphi(t)}| \leq \sup_{t \in [0, J_m]} |C_t - H_{\varphi(t)}| \leq 1 + \sup_{n \leq m} |H_{n+1} - H_n|. \quad (4.12)$$

De manera similar, se sigue de la definición de J_n que

$$\sup_{t \in [0, m]} |\varphi(t) - \frac{t}{2}| \leq \sup_{t \in [0, J_m]} |\varphi(t) - \frac{t}{2}| \leq \frac{1}{2} \sup_{n \leq m} H_n + \frac{1}{2} |J_m| + 1. \quad (4.13)$$

Teorema 86. *Tenemos que*

$$\left(\frac{1}{\sqrt{p}} C_{2pt}, t \geq 0 \right) \xrightarrow[p \rightarrow \infty]{\mathfrak{D}} \left(\frac{2}{\sigma} \beta_t, t \geq 0 \right)$$

donde β es un movimiento Browniano reflejado.

Demostración. Para toda $p \geq 1$, establecemos

$$\varphi_p(t) = p^{-1} \varphi(pt).$$

Por la desigualdad (4.12), tenemos que para todo $m \geq 1$,

$$\sup_{t \leq m} \left| \frac{1}{\sqrt{p}} C_{2pt} - \frac{1}{\sqrt{p}} H_{p\varphi_p(2t)} \right| \leq \frac{1}{\sqrt{p}} + \frac{1}{\sqrt{p}} \sup_{t \leq 2m} |H_{[pt]+1} - H_{[pt]}| \xrightarrow[p \rightsquigarrow \infty]{\mathbb{P}} 0, \quad (4.14)$$

donde la convergencia en probabilidad se da gracias a la convergencia en distribución del Teorema 71 y que el proceso límite es la constante 0. Por otro lado, tenemos la convergencia dada en la prueba del Teorema 71,

$$\frac{1}{\sqrt{p}} S_{[pt]} \xrightarrow[p \rightsquigarrow \infty]{\mathfrak{D}} \sigma B_t$$

lo cual, por la continuidad de la funcional ínfimo (supremo) sobre $C(\mathbb{R}_+, \mathbb{R}_+)$, implica que para todo $m \geq 1$,

$$\frac{1}{\sqrt{p}} I_{mp} \xrightarrow[p \rightsquigarrow \infty]{\mathfrak{D}} \sigma \inf_{t \leq m} B_t, \quad (4.15)$$

donde $\inf_{[pt] \leq mp} S_{[pt]} = I_{mp}$. Luego, de la desigualdad (4.13) se tiene que

$$\sup_{t \leq m} |\varphi_p(2t) - t| \leq \frac{1}{p} \sup_{k \leq 2mp} H_k + \frac{1}{p} |I_{2mp}| + \frac{2}{p} \xrightarrow[p \rightsquigarrow \infty]{\mathbb{P}} 0, \quad (4.16)$$

donde la convergencia en probabilidad se obtiene nuevamente por el Teorema 71 y (4.15). Combinando (4.14), (4.16) y el Teorema 71 se sigue el enunciado del presente teorema. \square

Ejemplo. Existe un caso especial en que el Teorema 86 no es difícil de obtener y que además no es necesario hacer referencia al Teorema 71. Este caso se da cuando tenemos que la distribución de hijos μ es una distribución geométrica:

$$\mu(k) = \frac{1}{2^{k+1}},$$

la cual satisface ser una distribución crítica con varianza positiva $\sigma^2 = 2$.

No es difícil ver que lejos del origen, el proceso de contorno $(C_n, n \geq 0)$ se comporta como una caminata aleatoria simple. Efectivamente, por las propiedades de la distribución geométrica, la probabilidad de que un individuo tenga al menos $n+1$ hijos dado que se sabe que al menos tiene n , es $1/2$ independientemente de la n . Esto es, si X tiene distribución μ ,

$$\mathbb{P}(X \geq n+1 | X \geq n) = \frac{\mathbb{P}(X \geq n+1, X \geq n)}{\mathbb{P}(X \geq n)} = \frac{\mathbb{P}(X \geq n+1)}{\mathbb{P}(X \geq n)} = \frac{\frac{1}{2^n}}{\frac{1}{2^{n-1}}} = \frac{1}{2}.$$

Además es claro que la caminata deberá estar condicionada a ser siempre positiva, por lo que el resultado del Teorema 86 en este caso se sigue del Teorema de invarianza de Donsker ($\sigma = \sqrt{2}$).

Claramente tendremos que podremos enunciar el resultado del Corolario 81 en terminos del proceso contorno de los respectivos árboles. Simplemente reemplazamos $H_{[pt]}^{\{x\sqrt{p}\}}$ por $C_{[2pt]}^{\{x\sqrt{p}\}}$, es decir, intercambiamos la función altura por la función de contorno del árbol $\theta^{\{x\sqrt{p}\}}$.

Corolario 87. *Tenemos que*

$$\left(\frac{1}{\sqrt{p}} C_{[2pt]}^{\{x\sqrt{p}\}}; t \geq 0 \right) \xrightarrow[p \rightsquigarrow \infty]{\mathfrak{D}} \left(\frac{2}{\sigma} e_t^{\sigma x/2}; t \geq 0 \right)$$

donde $e_t^{\sigma x/2}$ es una excursión Browniana condicionada a tener altura mayor que $\sigma x/2$.

Los distintos resultados de este capítulo, nos muestran que los procesos de altura con su respectivo reescalamiento (o bien procesos de contorno) de árboles Galton-Watson “grandes” convergen en distribución hacia excursiones Brownianas. Aun así, todavía no podemos afirmar que los árboles por si solos convergen. De hecho, para poder expresar matemáticamente esto último es necesario definir formalmente a que nos referimos cuando hablamos de un límite de árboles aleatorios y su convergencia.

El siguiente paso es dar de manera precisa la definición de árboles continuos y discutir una topología sobre el espacio de los árboles continuos. Esto nos permitirá reinterpretar los resultados hasta ahora dados como teoremas de convergencia de árboles aleatorios.

Capítulo 5

Árboles Reales y su Codificación por Excursiones Brownianas.

En este capítulo describimos en primera instancia el formalismo de los árboles reales, el cual será utilizado para dar un significado matemático preciso de la convergencia de árboles discretos con un reescalamiento adecuado hacia objetos continuos. Posteriormente mostramos como es que un árbol real puede ser codificado por funciones continuas de manera similar a la codificación de árboles discretos por su funciones contorno.

5.1. Árboles Reales

Definición 88. (Árbol Real). *Un espacio métrico compacto (\mathcal{T}, d) es un árbol real si cumple que para cualesquiera dos puntos $a, b \in \mathcal{T}$:*

1. *Existe una única isometría $f_{a,b} : [0, d(a, b)] \rightarrow \mathcal{T}$ tal que*

$$f_{a,b}(0) = a \text{ y } f_{a,b}(d(a, b)) = b.$$

2. *Si q es una función continua e inyectiva de $[0, 1]$ en \mathcal{T} , tal que $q(0) = a$ y $q(1) = b$, tendremos que*

$$q([0, 1]) = f_{a,b}([0, d(a, b)]).$$

Un árbol real con raíz es un árbol real (\mathcal{T}, d) con un vértice distinguido $\rho = \rho(\mathcal{T})$ al cual se le llama raíz. En lo subsecuente, los árboles reales en consideración tendrán siempre raíz, aunque no se mencione explícitamente.

Consideremos un árbol real (\mathcal{T}, d) , cualesquiera dos elementos $a, b \in \mathcal{T}$ y ρ su raíz.

- El rango de la isometría $f_{a,b}$ la denotamos por $[[a, b]]$, siendo este último el segmento de recta entre a y b en el árbol. En particular, $[[\rho, a]]$ es la trayectoria que va de la raíz hacia a , a la cual nos referimos como la línea ancestral del vértice a .

Más precisamente, definimos un orden parcial sobre el árbol de la siguiente manera:

- $a \preceq b$ (a es un ancestro de b) si y solo si $a \in [[\rho, b]]$.

Además, tendremos que:

- Para $a, b \in \mathcal{T}$, existe un único $c \in \mathcal{T}$ tal que $[[\rho, a]] \cap [[\rho, b]] = [[\rho, c]]$. En dicho caso, escribimos $c = a \wedge b$ y llamamos a c el ancestro común más reciente de a y b .
- Por definición, la multiplicidad de un vértice $a \in \mathcal{T}$ es el número de componente conexas (conexas maximales) de $\mathcal{T} \setminus \{a\}$. A los vértices en $\mathcal{T} \setminus \{\rho\}$ con multiplicidad 1, los llamamos hojas.

Notemos que un árbol real \mathcal{T} puede verse como la unión de todas las líneas ancestrales, es decir,

$$\mathcal{T} = \bigcup_{a \in \mathcal{T}} [[\rho, a]].$$

Una isometría es abierta, continua y cuando definida entre espacios métricos es inyectiva. Por otro lado, un arco es cualquier espacio homeomorfo al intervalo $[0, 1]$; sea h el homeomorfismo entre $[0, 1]$ y $[0, d(\rho, a)]$ con $a \in \mathcal{T}$ fijo, entonces $f_{\rho,a} \circ h$ es un homeomorfismo entre el intervalo $[0, 1]$ y $[[\rho, a]]$.

Por lo tanto, todo árbol real es unión de arcos y sus hojas serán los puntos extremos de los arcos maximales $[[\rho, a]]$ contenidos en \mathcal{T} , en consecuencia \mathcal{T} es un espacio arcoconexo.

De hecho una caracterización bastante útil es que un espacio métrico (\mathcal{T}, d) es un árbol real si y sólo si es completo, conexo por trayectorias y cumple la “condición de los cuatro puntos”,

$$d(x_1, x_2) + d(x_3, x_4) \leq \max \{d(x_1, x_3) + d(x_3, x_2), d(x_1, x_4) + d(x_2, x_4)\}$$

para cualesquiera cuatro puntos $x_1, x_2, x_3, x_4 \in \mathcal{T}$.

Nuestro objetivo ahora, es estudiar la convergencia de árboles aleatorios reales. Claramente será necesario tener una noción de distancia entre dos árboles reales. Utilizaremos la distancia de *Gromov-Hausdorff* entre espacios métricos compactos la cual fué introducida por Gromov en vista de sus aplicaciones geométricas.

A continuación, se da una breve introducción a los hiperespacios, que principalmente son espacios auxiliares utilizados para examinar de mejor manera propiedades de espacios compactos, métricos y conexos a través del estudio de sucesiones de subconjuntos del espacio en cuestión.

Definición 89. (Hiperespacios) *Para un espacio topológico E , sea:*

- (i.) $2^E = \{A : A \neq \emptyset, A \subseteq E \text{ y } A \text{ es cerrado}\}$
- (ii.) $C_n(E) = \{A \in 2^E : A \text{ tiene a lo más } n \text{ componentes.}\}$

Supongamos ahora, que E es un espacio métrico con métrica d . Con el propósito de dotar de una métrica apropiada a 2^E , definimos para cada $\epsilon > 0$ y para cada $A \in 2^E$, la nube de radio $\epsilon > 0$ del conjunto A como

$$N_\epsilon^d(A) = \{x \in E : d(x, a) < \epsilon \text{ para alguna } a \in A\}.$$

Finalmente, para $A, B \in 2^E$, definimos a la métrica de Hausdorff inducida por d como:

$$d_H(A, B) = \inf \{\epsilon > 0 : A \subseteq N_\epsilon^d(B), B \subseteq N_\epsilon^d(A)\}$$

Métricas equivalentes para E inducen la misma métrica de Hausdorff; los espacios 2^E y $C_n(E)$ con la métrica d_H son llamados hiperespacios de E .

Intuitivamente dos espacios A, B están suficientemente cerca con respecto a d_H , si cada punto de A está cercano a un punto de B y A y B son aproximadamente del mismo «tamaño geométrico».

Además, si E es un espacio compacto, métrico y conexo (un continuo) de igual manera lo serán sus hiperespacios, más aún $2^E, C_n(E)$ serán arcoconexos.

Definición 90. (Distancia de Gromov-Hausdorff)

Sean \mathcal{T} y \mathcal{T}' dos espacios métricos compactos con raíces ρ y ρ' respectivamente, definimos su distancia de Gromov-Hausdorff como

$$d_{GH}(\mathcal{T}, \mathcal{T}') = \inf\{d_H(\varphi(\mathcal{T}), \varphi'(\mathcal{T}')) \vee d(\varphi(\rho), \varphi'(\rho'))\},$$

donde el ínfimo es tomado sobre todas las posibles elecciones de un espacio métrico (E, d) y todos los encajes isométricos, $\varphi : \mathcal{T} \rightarrow E$ y $\varphi' : \mathcal{T}' \rightarrow E$ en (E, d) .

Definimos una relación de equivalencia (\sim) entre la colección de los espacios métricos compactos con raíz, dada por:

$$\mathcal{T}_1 \sim \mathcal{T}_2 \text{ si existe una isometría de } \mathcal{T}_1 \text{ sobre } \mathcal{T}_2 \text{ que preserve la raíz.}$$

En dado caso, diremos que los espacios son equivalentes. Es fácil ver que $d_{GH}(\mathcal{T}, \mathcal{T}')$ únicamente depende de las clases de equivalencia de \mathcal{T} y \mathcal{T}' . Entonces d_{GH} define una métrica sobre la colección de todas las clases de equivalencia de los espacios métricos compactos con raíz.

Denotamos por \mathbb{T} al conjunto de todas las clases de equivalencia de árboles reales con raíz.

Teorema 91. *El espacio métrico (\mathbb{T}, d_{GH}) es completo y separable.*

La prueba detallada a este teorema puede consultarse en [8], Teorema 1. Sin embargo, la idea de la demostración de completitud se basa en probar que el espacio de árboles reales compactos \mathbb{T} es un subconjunto cerrado del espacio de espacios métricos compactos considerado con la métrica de Gromov-Hausdorff; para esto se demuestra que el límite de cualquier sucesión de elementos en \mathbb{T} verifica ser conexo por trayectorias y satisface la condición de los cuatro puntos; esto último, gracias a un teorema que nos permite afirmar que dicha propiedad es preservada bajo límites de la métrica de Gromov-Hausdorff. Para la separabilidad, se demuestra que el conjunto de árboles finitos con longitud racional entre los vértices es numerable y denso en (\mathbb{T}, d_{GH}) .

Utilizaremos la siguiente definición alternativa de d_{GH} . Sean (\mathcal{T}_1, d_1) y (\mathcal{T}_2, d_2) dos espacios métricos compactos. Una correspondencia entre \mathcal{T}_1 y \mathcal{T}_2 es un subconjunto \mathcal{R} de $\mathcal{T}_1 \times \mathcal{T}_2$ tal que para todo $x_1 \in \mathcal{T}_1$ existe al menos un $x_2 \in \mathcal{T}_2$ de manera que $(x_1, x_2) \in \mathcal{R}$ e inversamente, para todo $y_2 \in \mathcal{T}_2$ existe al menos un $y_1 \in \mathcal{T}_1$ de manera que $(y_1, y_2) \in \mathcal{R}$. La distorsión de la correspondencia \mathcal{R} se define como

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

Entonces, si \mathcal{T} y \mathcal{T}' son dos espacios compactos métricos con raíces ρ y ρ' respectivamente, tendremos que

$$d_{GH}(\mathcal{T}, \mathcal{T}') = \frac{1}{2} \inf \{ \text{dis}(\mathcal{R}) : \mathcal{R} \in \mathcal{C}(\mathcal{T}, \mathcal{T}'), (\rho, \rho') \in \mathcal{R} \}$$

donde $\mathcal{C}(\mathcal{T}, \mathcal{T}')$ denota el conjunto de todas las correspondencias entre $\mathcal{T}, \mathcal{T}'$.

5.2. Codificando Árboles Reales

Describamos a continuación un método para construir árboles reales, el cual, es particularmente adecuado para las aplicaciones que se darán posteriormente para árboles aleatorios.

Consideremos una función continua (determinista) $g : [0, \infty) \rightarrow [0, \infty)$ con soporte compacto, esto es, $\overline{\{x : g(x) \neq 0\}}$ es compacto y que además $g(0) = 0$. Para evitar resultados triviales, supondremos que g no es la función idénticamente cero. Para todo $s, t \geq 0$, establecemos

$$m_g(s, t) = \inf_{r \in [s \wedge t, s \vee t]} g(r),$$

y definimos a la función $d_g : [0, \infty) \times [0, \infty) \rightarrow [0, \infty)$ por:

$$d_g(s, t) = g(s) + g(t) - 2m_g(s, t).$$

Claramente $d_g(s, t) = d_g(t, s)$ y es fácil verificar que se cumple la desigualdad triangular

$$d_g(s, u) \leq d_g(s, t) + d_g(t, u)$$

para todo $s, t, u \geq 0$.

Introducimos la siguiente relación de equivalencia (\sim) la cual puede ser definida por:

- (1) $s \sim t$ si y solo si $d_g(s, t) = 0$,
- (2) $s \sim t$ si y solo si $g(s) = g(t) = m_g(s, t)$.

Claramente (2) \Rightarrow (1), para ver que (1) \Rightarrow (2) observemos que

$$\begin{aligned} d_g(s, t) = 0 &\Leftrightarrow g(s) + g(t) - 2m_g(s, t) = 0 \\ &\Leftrightarrow 0 \leq g(t) - m_g(s, t) = m_g(s, t) - g(s) \leq 0, \end{aligned}$$

de donde se concluye que $g(s) = g(t) = m_g(s, t)$.

Dada la relación anterior, es posible definir el espacio cociente

$$\mathcal{T}_g = [0, \infty) / \sim .$$

Obviamente la función d_g induce una métrica sobre \mathcal{T}_g , mantendremos la notación d_g para ésta métrica. A la función $p_g : [0, \infty) \rightarrow \mathcal{T}_g$ dada por $s \mapsto [s]_{\sim}$ donde

$$[s]_{\sim} = \{t \in [0, \infty) : t \sim s\} \in \mathcal{T}_g,$$

la llamamos la proyección canónica .

Denotamos $\rho = p_g(0) = \{s \in [0, \infty) : g(s) = 0\}$. Si $\zeta > 0$ es el supremo del soporte de g , tenemos que $p_g(t) = \rho$ para toda $t \geq \zeta$. En particular,

$$\mathcal{T}_g = p_g([0, \zeta]) = [0, \zeta] / \sim,$$

es compacto y arcoconexo. Notemos que la línea de ancestros de un vértice $p_g(s)$ que denotamos por $[[\rho, p_g(s)]]$, es isométrico al intervalo $[0, g(s)]$.

En este capítulo, estaremos interesados en el caso en que g sea una excursión browniana con duración $\zeta = 1$.

Proposición 92. *La función proyección p_g es una función cociente, es decir, el espacio métrico (\mathcal{T}_g, d_g) está dotado con la topología cociente (topología fuerte) inducida por p_g , es decir,*

$$U \subseteq \mathcal{T}_g \text{ es abierto si y solo si } p_g^{-1}(U) \subseteq [0, \zeta] \text{ es abierto.}$$

Antes de proceder la prueba de la proposición, presentamos dos resultados topológicos que nos facilitaran el esquema de la demostración:

Teorema 93. *Si A, B son espacios topológicos y $f : A \rightarrow B$ es continua sobreyectiva y cerrada, entonces la topología de B es exactamente la topología cociente inducida por f .*

Demostración. Sea f continua y cerrada. Por definición la topología cociente es la topología fuerte en el caso particular de una función, por lo que basta ver que si τ_B es la topología de B entonces $\tau_f \subseteq \tau_B$ ya que por definición $\tau_B \subseteq \tau_f$. Sea $U \in \tau_f$ entonces por definición de topología cociente $f^{-1}(U) \in \tau_A$ por lo que $A \setminus f^{-1}(U) = f^{-1}(B \setminus U)$ es cerrado en A . Ahora bien, como f es cerrada $f[f^{-1}(B \setminus U)] = B \setminus U$ es cerrado en (B, τ_B) y por tanto $U \in \tau_B$. \square

Lema 94. Si para una función continua $f : A \rightarrow B$ existe un conjunto $M \subset A$ tal que $f(M) = B$ y la restricción $f|_M$ es cociente, entonces f es un cociente.

La prueba es inmediata de las propiedades de una función cociente en composición con una función continua (Véase [6] Cor. 2.4.6., pág. 92).

Demostración de la Proposición 92.

Ya que $p_g([0, \zeta]) = \mathcal{T}_g$, por los resultados anteriores basta verificar que $p_g : [0, \zeta] \rightarrow \mathcal{T}_g$ es continua y cerrada. La continuidad se obtiene cuando se considera al espacio $[0, \zeta]$ equipado con la métrica euclidiana y al espacio métrico (\mathcal{T}_g, d_g) . Dada una sucesión $t_1, t_2, \dots \in [0, \zeta]$, tal que $t_n \rightarrow t$ cuando $n \rightarrow \infty$, veamos que $p_g(t_n) \rightarrow p_g(t)$. Sea $\epsilon > 0$, entonces

$$d_g(p_g(t), p_g(t_n)) = g(t) + g(t_n) - 2m_g(t_n, t),$$

por otro lado y gracias a la continuidad de la función g , existe $\delta > 0$ tal que $|s - t| < \delta$ implica

$$|g(s) - g(t)| < \frac{\epsilon}{3},$$

usando esta desigualdad elegimos $N \in \mathbb{N}$, de manera que para toda $n \geq N$ tendremos que $|t_n - t| < \delta$ y en consecuencia $|g(t_n) - g(t)| < \frac{\epsilon}{3}$ para toda $n \geq N$.

Es importante tener presente que $m_g(t_n, t) = g(s)$ para algún $s \in [t_n \wedge t, t_n \vee t]$, ya que entonces tendremos que $|s - t| < \delta$ para toda $n \geq N$. Luego, para toda $n \geq N$

$$g(t_n) - m_g(t, t_n) \leq |g(t_n) - g(t)| + g(t) - m_g(t, t_n) < \frac{\epsilon}{3} + \frac{\epsilon}{3} = \frac{2\epsilon}{3}$$

por lo que finalmente

$$d_g(p_g(t), p_g(t_n)) \leq [g(t_n) - m_g(t_n, t)] + [g(t) - m_g(t_n, t)] < \frac{2\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

con esto hemos probado que $p_g(t_n) \rightarrow p_g(t)$ cuando $n \rightarrow \infty$, por lo tanto p_g es continua. Claramente será cerrada, ya que toda función continua de un espacio compacto en un espacio de Hausdorff (es posible separar cualesquiera dos puntos por abiertos ajenos) es cerrada. \square

Teorema 95. El espacio métrico (\mathcal{T}_g, d_g) es un árbol real con raíz $\rho = p_g(0)$.

Es posible probar que el recíproco también es cierto, esto es, que todo árbol real con raíz es posible representarlo de la forma \mathcal{T}_g .

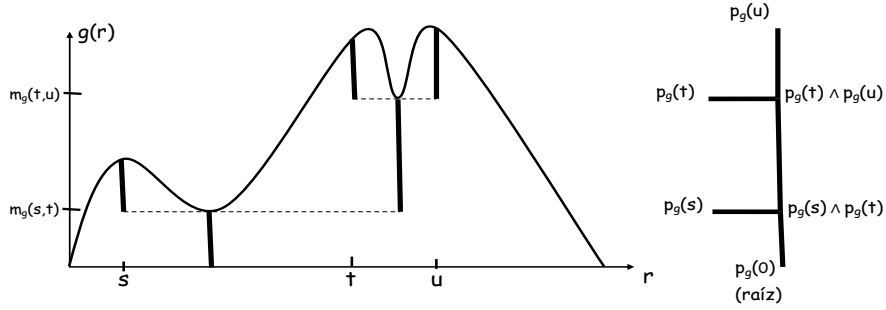


Figura 5.3: Sub-árbol de \mathcal{T}_g unión de las líneas ancestrales con vértices $p_g(s)$, $p_g(t)$, $p_g(u)$ (Le Gall [11]).

Ilustramos ahora, antes de probar el teorema anterior, lo que llamamos un sub-árbol de \mathcal{T}_g también denominado un árbol reducido, unión de las líneas ancestrales con vértices $p_g(s)$, $p_g(t)$, $p_g(u)$.

Las líneas ancestrales de $p_g(s)$, $p_g(t)$ y $p_g(u)$ son segmentos de recta de longitud $g(s)$, $g(t)$ y $g(u)$ respectivamente. Las líneas ancestrales de $p_g(s)$ y $p_g(t)$ comparten un segmento de recta isométrico al intervalo $[0, m_g(s, t)]$ (como se intenta ilustrar en la figura), y por supuesto una propiedad similar persiste para las líneas ancestrales de $p_g(s)$ y $p_g(u)$, o $p_g(u)$ y $p_g(t)$.

Daremos una prueba elemental del Teorema (95), la cual únicamente utiliza la definición de árbol real, a la vez será de gran ayuda para entender las nociones de “línea ancestral” y “ancestro común más reciente” en \mathcal{T}_g . Otro argumento que depende del Teorema (91) se presentará más adelante.

Antes de proceder con la prueba del Teorema (95), establecemos y probamos el siguiente lema de cambio de raíz, el cual es de interés independiente al recién mencionado teorema. Para su demostración recurrimos al siguiente resultado topológico.

Sean A, B dos conjuntos con relaciones de equivalencia r, R , respectivamente. Una función $f : A \rightarrow B$ se dice que *preserva-relación* si $ara' \Rightarrow f(a)Rf(a')$. Además, recordemos que las proyecciones son las funciones denotadas por p_A, p_B , donde

$$a \mapsto p_A(a) = \{a' \in A : a'ra\} \quad \text{y} \quad b \mapsto p_B(b) = \{b' \in A : b'Rb\}.$$

Teorema 96. Sean A, B dos conjuntos con relaciones de equivalencia r, R respectivamente y $f : A \rightarrow B$ una función que *preserva-relación*, entonces existe una y sólo

una función $f^* : A/r \rightarrow B/R$, tal que el diagrama

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ p_A \downarrow & & \downarrow p_B \\ A/r & \xrightarrow{f^*} & B/R \end{array}$$

conmuta ($f^* \circ p_A = p_B \circ f$) y donde la continuidad de f implica la de f^* . Además, f^* es llamada la función inducida por f en el “paso al cociente”. Inversamente si f, f^* son cualesquiera dos funciones tales que el diagrama anterior conmuta, entonces necesariamente f preserva-relación y f^* es la función inducida por f .

Véase [3] Cap. I, Teo. 7.7. y Cap. VI, Teo. 4.3., págs. 17, 127.

Lema 97. (Cambio de Raíz).

Sea $s_0 \in [0, \zeta]$. Para cualquier número real $r \geq 0$, denotamos por \bar{r} al único elemento de $[0, \zeta)$ tal que $r - \bar{r}$ es un entero múltiplo de ζ . Establecemos

$$\begin{aligned} g'(s) &= g(s_0) + g(\overline{s_0 + s}) - 2m_g(s_0, \overline{s_0 + s}), & \text{para todo } s \in [0, \zeta] \\ g'(s) &= 0, & \text{para } s > \zeta. \end{aligned}$$

Entonces, la función g' es continua con soporte compacto y satisface $g'(0) = 0$ de manera que es posible definir $\mathcal{T}_{g'}$. Más aún, para todo $s, t \in [0, \zeta]$, tenemos que

$$d_{g'}(s, t) = d_g(\overline{s_0 + s}, \overline{s_0 + t}) \quad (5.1)$$

y existe una única isometría \mathcal{R} de $\mathcal{T}_{g'}$ sobre \mathcal{T}_g tal que, para toda $s \in [0, \zeta]$,

$$\mathcal{R}(p_{g'}(s)) = p_g(\overline{s_0 + s}). \quad (5.2)$$

Asumiendo la veracidad del Teorema (95), es posible observar que $\mathcal{T}_{g'}$ coincidirá con el árbol real \mathcal{T}_g cambiando la raíz a $p_g(s_0)$. Por lo que el lema anterior nos estaría indicando, cual función codifica al árbol \mathcal{T}_g con cambio de raíz en un vértice arbitrario.

Demostración. Queremos probar que g' es continua en $[0, \zeta)$ con soporte compacto y cumple que $g'(0) = 0$, con lo que tendría sentido hablar de $\mathcal{T}_{g'}$. Sea $s_0 \in [0, \zeta)$ fijo. Si $s \in [0, \zeta]$ entonces $s_0 + s \in [0, 2\zeta)$ y por definición $\overline{s_0 + s} \in [0, \zeta)$, por lo que

$$s_0 + s - \overline{s_0 + s} = \begin{cases} 0 & \text{si } s_0 + s < \zeta \Leftrightarrow \overline{s_0 + s} = s_0 + s, \\ \zeta & \text{si } s_0 + s \geq \zeta \Leftrightarrow \overline{s_0 + s} = s_0 + s - \zeta. \end{cases}$$

Luego, tendremos que la función g' que obviamente dependerá de $s_0 \in [0, \zeta)$ fija, la podremos expresar de la siguiente manera:

$$g'(s) = \begin{cases} d_g(s_0, s_0 + s) & \text{si } s \in [0, \zeta - s_0), \\ d_g(s_0, s_0 + s - \zeta) & \text{si } s \in [\zeta - s_0, \zeta), \\ 0 & \text{si } s \geq \zeta \end{cases}$$

De esto último fácilmente nos percatamos que $g'(0) = 0$ y que g' es no-negativa.

Observemos que para toda $\epsilon > 0$, g es continua en los compactos $[0, \zeta - s_0 - \epsilon]$ y $[\zeta - s_0, \zeta - \epsilon]$ y alcanza su ínfimo en cada uno de ellos. De esto tendremos que g' como antes definida será continua sobre cada uno de estos intervalos y por lo tanto sobre su unión. Finalmente basta probar que en $\zeta - s_0$ es continua por la izquierda, lo cual obtenemos del hecho que $d_g(s_0, s_0 + (\zeta - s_0)) = d_g(s_0, s_0 + (\zeta - s_0) - \zeta)$ ya que $g(0) = g(\zeta) = 0 = m_g(0, s_0) = m_g(s_0, \zeta)$. Por lo tanto g' será continua en $[0, \infty)$.

Ahora si, tiene sentido considerar al espacio $\mathcal{T}_{g'}$. Por el Teorema 96 basta ver que dada $s_0 \in [0, \zeta)$, la función $f : [0, \zeta] \rightarrow [0, \zeta]$ definida por $f(t) = \overline{s_0 + t}$ preserva las relaciones de equivalencia que dan origen a los espacios cocientes \mathcal{T}_g y $\mathcal{T}_{g'}$; probado esto, podremos afirmar que existe una única función \mathcal{R} tal que el siguiente diagrama conmuta:

$$\begin{array}{ccc} [0, \zeta] & \xrightarrow{f} & [0, \zeta] \\ p_{g'} \downarrow & & \downarrow p_g \\ \mathcal{T}_{g'} & \xrightarrow{\mathcal{R}} & \mathcal{T}_g \end{array}$$

Así las cosas, la clave está en poder verificar la relación (5.1), ya que nos indica que efectivamente la función f antes mencionada preserva la relación. En los siguientes casos consideraremos sin pérdida de generalidad que $s, t \in [0, \zeta]$ con $s < t$.

Caso 1. $s, t \in [0, \zeta - s_0)$. Entonces, $\overline{r + s_0} = r + s_0$ para toda $r \in [s, t]$ y

$$m_{g'}(s, t) = \inf_{r \in [s, t]} g'(r) = \inf_{r \in [s, t]} [g(s_0) + g(s_0 + r) - 2m_g(s_0, s_0 + r)].$$

Existen dos posibilidades:

(i) $m_g(s_0 + s, s_0 + t) \geq m_g(s_0, s_0 + s)$

$$\Rightarrow m_g(s_0, s_0 + r) = m_g(s_0, s_0 + s) = m_g(s_0, s_0 + t) \quad \forall r \in [s, t],$$

$$\Rightarrow m_{g'}(s, t) = g(s_0) + m_g(s_0 + s, s_0 + t) - 2m_g(s_0, s_0 + s).$$

Se seguirá entonces que

$$\begin{aligned}
d_{g'}(s, t) &= g'(s) + g'(t) - 2m_{g'}(s, t) \\
&= g(s_0 + s) - 2m_g(s_0, s_0 + s) + g(s_0 + t) - 2m_g(s_0, s_0 + t) \\
&\quad - 2[m_g(s_0 + s, s_0 + t) - 2m_g(s_0, s_0 + s)] \\
&= g(s_0 + s) + g(s_0 + t) - 2m_g(s_0 + s, s_0 + t) \\
&= d_g(s_0 + s, s_0 + t) \\
&= d_g(\overline{s_0 + s}, \overline{s_0 + t}).
\end{aligned}$$

(ii) $m_g(s_0 + s, s_0 + t) < m_g(s_0, s_0 + s)$

$$\Rightarrow -m_g(s_0 + s, s_0 + t) \geq -m_g(s_0, s_0 + r) \geq -m_g(s_0, s_0 + s), \quad \forall r \in [s, t]$$

y tendremos que es posible calcular explícitamente

$$\inf_{r \in [s, t]} g'(r) = \inf_{r \in [s, t]} [g(s_0) + g(s_0 + r) - 2m_g(s_0, s_0 + r)],$$

puesto que es posible minimizar gracias a lo anterior, obteniendo que,

$$\inf_{r \in [s, t]} g'(r) = \inf_{r \in [s, t]} [g(s_0) + g(s_0 + r) - 2m_g(s_0, s_0 + s)],$$

por último como para toda $r \in [s, t]$ tenemos que

$$g(s_0 + r) - m_g(s_0 + s, s_0 + t) > g(s_0 + r) - m_g(s_0, s_0 + s) \geq 0,$$

minimizamos esta cantidad nuevamente, eligiendo al primer $r^* \in [s, t]$ tal que

$$g(s_0 + r^*) = m_g(s_0, s_0 + s),$$

obteniendo así que,

$$\inf_{r \in [s, t]} g'(r) = m_{g'}(s, t) = g(s_0) - m_g(s_0, s_0 + s).$$

Se sigue que

$$\begin{aligned}
d_{g'}(s, t) &= g'(s) + g'(t) - 2m_{g'}(s, t) \\
&= g(s_0) + g(s_0 + s) - 2m_g(s_0, s_0 + s) \\
&\quad + g(s_0) + g(s_0 + t) - 2m_g(s_0, s_0 + t) \\
&\quad - 2[g(s_0) - m_g(s_0, s_0 + s)] \\
&= g(s_0 + s) - 2m_g(s_0, s_0 + s) + g(s_0 + t) - 2m_g(s_0, s_0 + t) \\
&\quad + 2m_g(s_0, s_0 + s) \\
&= d_g(s_0 + s, s_0 + t).
\end{aligned}$$

Análogamente se verifican sin pérdida de generalidad para $s < t$, los siguientes casos:

Caso 2. $s, t \in (\zeta - s_0, \zeta]$,

Caso 3. $s \in (\zeta - s_0, \zeta]$ y $t \in [0, \zeta - s_0]$.

Por (5.1) si $s, t \in [0, \zeta]$ son tal que $d_{g'}(s, t) = 0$, tenemos que $d_g(\overline{s_0 + s}, \overline{s_0 + t}) = 0$, así que $p_g(\overline{s_0 + s}) = p_g(\overline{s_0 + t})$ y la función $f(t) = \overline{s_0 + t}$ efectivamente preserva la relación de equivalencia, de donde tendremos que el diagrama previo conmuta y de (5.1) que la función inducida en el paso al cociente \mathcal{R} es una isometría. Finalmente, $\mathcal{T}_{g'} = p_{g'}([0, \zeta])$ (el supremo de el soporte de g' es menor igual que ζ) y \mathcal{R} está determinada de manera única por la relación (5.2) y el Teorema (96). \square

Definición 98. (Punto de Corte).

Sea (S, τ) un espacio topológico conexo y $\mathbf{p} \in S$. Si $S \setminus \{\mathbf{p}\}$ es conexo, entonces \mathbf{p} es llamado un punto de no-corte de S . Si $S \setminus \{\mathbf{p}\}$ es desconexo, entonces \mathbf{p} es llamado un punto de corte de S .

Notación $\mathbf{S} = \mathbf{U|V}$. Si S es un espacio topológico, escribimos $S = U|V$ para abreviar el hecho de que $S = U \cup V$ donde $U \neq \emptyset \neq V$, $U \cap V = \emptyset$ y U, V son abiertos en S .

Por lo tanto, $S = U|V$ si y sólo si S es desconexo y decimos que los conjuntos U, V conforman una separación de S .

Damos algunos preliminares. Para $\sigma, \sigma' \in \mathcal{T}_g$, definimos la relación (\preceq) por:

$$\sigma \preceq \sigma' \text{ si y sólo si } d_g(\sigma, \sigma') = d_g(\rho, \sigma') - d_g(\rho, \sigma).$$

Si $\sigma = p_g(s)$ y $\sigma' = p_g(t)$, se sigue entonces de nuestras definiciones que

$$\sigma \preceq \sigma' \text{ si y sólo si } m_g(s, t) = g(s),$$

ya que σ comparte una línea ancestral de longitud $m_g(s, t)$ con σ' y para que σ sea ancestro de σ' basta que su línea ancestral de longitud $g(s)$ se encuentre completamente contenida en la de σ' con lo que necesariamente $m_g(s, t) = g(s)$. Es inmediato verificar que la relación (\preceq) induce un orden parcial no estricto en \mathcal{T}_g , esto quiere decir que es una relación reflexiva, transitiva y asimétrica.

Para cualesquiera $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathcal{T}_g$, notemos que por la transitividad antes mencionada

$$\begin{aligned} \mathbf{a} \preceq \mathbf{c} & \text{ si y sólo si } d_g(\mathbf{a}, \mathbf{c}) = d_g(\rho, \mathbf{c}) - d_g(\rho, \mathbf{a}), \\ \mathbf{c} \preceq \mathbf{b} & \text{ si y sólo si } d_g(\mathbf{c}, \mathbf{b}) = d_g(\rho, \mathbf{b}) - d_g(\rho, \mathbf{c}), \end{aligned}$$

lo cual implica que

$$\mathbf{a} \preceq \mathbf{c} \preceq \mathbf{b} \text{ si y sólo si } d_g(\mathbf{a}, \mathbf{b}) = d_g(\mathbf{a}, \mathbf{c}) + d_g(\mathbf{c}, \mathbf{b}).$$

Con esto último es natural establecer

$$[[\mathbf{a}, \mathbf{b}]] = \{\mathbf{c} \in \mathcal{T}_g : d_g(\mathbf{a}, \mathbf{b}) = d_g(\mathbf{a}, \mathbf{c}) + d_g(\mathbf{c}, \mathbf{b})\}.$$

Si $\sigma = p_g(s)$ y $\sigma' = p_g(t)$, entonces es fácil verificar que $[[\rho, \sigma]] \cap [[\rho, \sigma']] = [[\rho, \gamma]]$ donde $\gamma = p_g(r)$ si r es cualquier tiempo para el cual se alcanza el mínimo de g entre s y t . Entonces escribimos $\gamma = \sigma \wedge \sigma'$.

Hacemos,

$$\mathcal{T}_g[\sigma] := \{\sigma' \in \mathcal{T}_g : \sigma \preceq \sigma'\}.$$

Si $\mathcal{T}_g[\sigma] \neq \{\sigma\}$ y $\sigma \neq \rho$, entonces haciendo $U = \mathcal{T}_g \setminus \mathcal{T}_g[\sigma]$ y $V = \mathcal{T}_g[\sigma] \setminus \{\sigma\}$, afirmamos que $\mathcal{T}_g \setminus \{\sigma\} = U \vee V$. Claramente $U \cap V = \emptyset$. Para ver que U es abierto, sea s tal que $\sigma = p_g(s)$ y notemos que de nuestras definiciones

$$\mathcal{T}_g[\sigma] := p_g(\{u \in [0, \zeta] : m_g(s, u) = g(s)\}),$$

donde $\{u \in [0, \zeta] : m_g(s, u) = g(s)\}$ es un conjunto cerrado y recordemos p_g es una función cerrada. El conjunto V es abierto, ya que si $\sigma' \in \mathcal{T}_g[\sigma]$ y $\sigma' \neq \sigma$, se sigue fácilmente de nuestras definiciones que la bola abierta centrada en σ' de radio $d_g(\sigma, \sigma')$ se queda contenida en V . Por lo tanto, todo $\sigma \in \mathcal{T}_g \setminus \{\rho\}$ es punto de corte de \mathcal{T}_g .

Con la información anterior a nuestro alcance nos encontramos en posición de probar que (\mathcal{T}_g, d_g) es un árbol real con raíz $\rho = p_g(0)$.

Demostración del Teorema 95. Probamos la propiedad (1) de la definición de árbol real. Fijando $\sigma_1, \sigma_2 \in \mathcal{T}_g$ hay que probar la existencia e unicidad de la isometría f_{σ_1, σ_2} . Utilizando el Lema 97 y sea s_0 tal que $p_g(s_0) = \sigma_1$, es posible asumir que trabajamos sobre el mismo árbol recodificado a manera que $\sigma_1 = \rho$.

Sea entonces $\sigma \in \mathcal{T}_g$ fijo, debemos probar la existencia de una única isometría $f = f_{\rho, \sigma} : [0, d_g(\rho, \sigma)] \rightarrow \mathcal{T}_g$ tal que $f(0) = \rho$ y $f(d_g(\rho, \sigma)) = \sigma$. Si $s \in p_g^{-1}(\{\sigma\})$, se tiene que $g(s) = d_g(\rho, \sigma)$. Luego, para toda $a \in [0, d_g(\rho, \sigma)]$, si hacemos

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

tendremos que $g(v(a)) = a$ por continuidad de g en $[0, s]$. Definimos a f por

$$f(a) = p_g(v(a)).$$

Tenemos que $f(0) = \rho$ y $f(d_g(\rho, \sigma)) = \sigma$, esto último ya que

$$m_g(v(g(s)), s) = g(s) = g(v(g(s))),$$

de donde $p_g(v(g(s))) = p_g(s) = \sigma$. Es fácil verificar que f es una isometría: Si $a, b \in [0, d_g(\rho, \sigma)]$ con $a \leq b$, es inmediato que $m_g(v(a), v(b)) = a$, y entonces

$$d_g(f(a), f(b)) = g(v(a)) + g(v(b)) - 2a = b - a.$$

Para probar la unicidad, supongamos que existe f^* una isometría que satisface las mismas propiedades que f . Si $a \in [0, d_g(\rho, \sigma)]$ entonces $f^*(a) \preceq \sigma$, ya que

$$d_g(f^*(a), \sigma) = d_g(\rho, \sigma) - a = d_g(\rho, \sigma) - d_g(\rho, f^*(a)).$$

Recordemos que $\sigma = p_g(s)$ y sea r tal que $p_g(r) = f^*(a)$. Notemos que

$$g(r) = d_g(\rho, p_g(r)) = a$$

y ya que $f^*(a) \preceq \sigma$ también $g(r) = m_g(r, s)$. De esto último tendremos que

$$m_g(v(a), s) = a = g(r) = g(v(a)) = m_g(v(a), r),$$

y entonces $d_g(r, v(a)) = 0$, de manera que $f^*(a) = p_g(r) = p_g(v(a)) = f(a)$. Con esto, la prueba de (1.) está completa.

Para probar la parte (2.) de la definición de árbol real, dada una función inyectiva y continua $q : [0, 1] \rightarrow \mathcal{T}_g$, queremos mostrar que

$$q([0, 1]) = f_{q(0), q(1)}([0, d_g(q(0), q(1))]).$$

Por el Lema de cambio de raíz y sin pérdida de generalidad podemos asumir que $q(0) = \rho$ y hacer $q(1) = \sigma$. Así, $f_{\rho, \sigma}([0, d_g(\rho, \sigma)]) = [[\rho, \sigma]]$.

Es fácil ver de inicio que $q([0, 1])$ es un arco por ser homeomorfo al espacio $[0, 1]$. Probamos primero que $[[\rho, \sigma]] \subseteq q([0, 1])$. Supongamos $\eta \in [[\rho, \sigma]] \setminus q([0, 1])$, en particular $\rho \neq \eta \neq \sigma$. Entonces

$$q([0, 1]) \subseteq S = U \cup V \quad \text{donde } U = \mathcal{T}_g \setminus \mathcal{T}_g[\eta] \text{ y } V = \mathcal{T}_g[\eta] \setminus \{\eta\}.$$

Pero recordemos que $q(0) = \rho \in U$ y $q(1) = \sigma \in V$ y en consecuencia

$$q([0, 1]) \cap U \neq \emptyset \neq q([0, 1]) \cap V,$$

contradiciendo así la conexidad de $q([0, 1])$. Por lo tanto, $[[\rho, \sigma]] \subseteq q([0, 1])$.

Para la otra contención, supongamos que $\eta \in q([0, 1]) \setminus [[\rho, \sigma]]$, entonces $\eta = q(a)$ con $a \in (0, 1)$. Sea $\gamma = \eta \wedge \sigma$, entonces $\gamma \in [[\rho, \eta]] \cap [[\rho, \sigma]]$, de donde

$$d_g(\eta, \sigma) = d_g(\eta, \gamma) + d_g(\gamma, \sigma).$$

De esto último concluimos que $\gamma \in [[\eta, \sigma]]$. Ahora bien, de la primera parte de la prueba de (2.) podemos afirmar que $[[\rho, \gamma]] \subset q([0, a])$ y en particular $\gamma \in q([0, a])$; por un argumento de cambio de raíz y nuevamente por la primera parte de la prueba también tendremos que $\gamma \in [[\eta, \sigma]] \subseteq q([a, 1])$. Como q es inyectiva, lo anterior es posible sólo si $\gamma = q(a) = \eta$, lo cual contradice el hecho de que $\eta \notin [[\rho, \sigma]]$. \square

Ya que hemos probado que (\mathcal{T}_g, d_g) es un árbol real, es posible verificar rápidamente que las notaciones $\sigma \preceq \sigma'$, $[[\sigma, \sigma']]$, $\sigma \wedge \sigma'$ utilizadas durante la prueba son consistentes con las definiciones introducidas en la primera sección de este capítulo para un árbol real en general.

Discutamos brevemente las multiplicidades en el árbol \mathcal{T}_g . Si $\sigma \in \mathcal{T}_g$ no es una hoja entonces debe ocurrir que $\ell(\sigma) \leq r(\sigma)$, donde

$$\ell(\sigma) := \inf p_g^{-1}(\{\sigma\}), \quad r(\sigma) := \sup p_g^{-1}(\{\sigma\})$$

son respectivamente la preimagen más pequeña y más grande de σ ó bien el elemento más pequeño y el más grande respectivamente de la clase de equivalencia σ en $[0, \zeta]$. Notemos además que como $d_g(\ell(\sigma), r(\sigma)) = 0$ tendremos

$$m_g(\ell(\sigma), r(\sigma)) = g(\ell(\sigma)) = g(r(\sigma)) = d_g(\rho, \sigma).$$

Recordemos rápidamente que un espacio localmente conexo es todo aquel que en cada punto posee una base de abiertos conexos. Todo subconjunto abierto de un espacio localmente conexo vuelve a ser un espacio localmente conexo. Es una caracterización de éstos espacios que las componentes de todo subconjunto abierto son abiertas (y cerradas también), además de poseer una cantidad finita de componentes.

Dado un punto x , la componente de x es un conexo maximal ó bien la unión de todos los conexos que lo tienen. La cuasi-componente de un punto x es la intersección de todos los abiertos-cerrados que tienen al punto x . Debido a que un conjunto conexo no puede contener propiamente conjuntos abiertos-cerrados, se tiene que toda componente está contenida en una cuasi-componente. En espacios localmente conexos las componentes y cuasicomponentes coinciden.

Sea \mathcal{I} un conjunto finito de índices, por la mencionada caracterización es posible denotar por (a_i, b_i) con $i \in \mathcal{I}$ a las componentes conexas del conjunto abierto

$$A = (\ell(\sigma), r(\sigma)) \cap \{t \geq 0 : g(t) > d_g(\rho, \sigma)\}$$

Observemos que en caso de que σ sea una hoja se tiene que $\mathcal{I} = \emptyset$.

Proposición 99. *Las componentes conexas del abierto $\mathcal{T}_g \setminus \{\sigma\}$ son exactamente los conjuntos $p_g((a_i, b_i))$ con $i \in \mathcal{I}$ y en caso de que σ no sea la raíz $\mathcal{T}_g \setminus \mathcal{T}_g[\sigma]$ también es componente.*

Demostración. Se probó que el conjunto $\mathcal{T}_g \setminus \mathcal{T}_g[\sigma]$ es abierto, al igual que $\mathcal{T}_g[\sigma] \setminus \{\sigma\}$. Nótese que en caso de que σ sea la raíz $\mathcal{T}_g[\sigma] \setminus \{\sigma\}$ y $\mathcal{T}_g \setminus \{\sigma\}$ coinciden.

Como el conjunto A es localmente conexo al ser subconjunto abierto de $[0, \zeta]$ y p_g es continua, sobreyectiva y cerrada, tendremos que $p_g(A) = \mathcal{T}_g[\sigma] \setminus \{\sigma\}$ es localmente conexo y por lo tanto sus componentes conexas coinciden con sus cuasi-componentes.

No es difícil percatarse que los conjuntos $p_g((a_i, b_i))$ con $i \in \mathcal{I}$ son cerrados, ya que la proyección p_g es cerrada y las componentes (a_i, b_i) son cerradas, pero $p_g((a_i, b_i))$ con $i \in \mathcal{I}$ también son abiertos, la prueba es similar a cuando se probó que $\mathcal{T}_g[\sigma] \setminus \{\sigma\}$ es abierto. Más aún, por ser imagen continua de intervalos tenemos que son conexos. Es decir, como los conjuntos $p_g(a_i, b_i)$ con $i \in \mathcal{I}$ son abiertos-cerrados conexos y $\mathcal{T}_g[\sigma] \setminus \{\sigma\}$ es localmente conexo, necesariamente deben coincidir con sus componentes conexas.

Finalmente, $\mathcal{T}_g \setminus \mathcal{T}_g[\sigma]$ es conexo, ya que para cualesquiera $\sigma', \sigma'' \in \mathcal{T}_g \setminus \mathcal{T}_g[\sigma]$, $[[\rho, \sigma']] \cup [[\rho, \sigma'']]$ es un conjunto cerrado y conexo contenido en $\mathcal{T}_g \setminus \mathcal{T}_g[\sigma]$. \square

Concluimos la sección con un lema que nos permite comparar árboles codificados por dos funciones distintas g y g' .

Lema 100. *Sean $g, g' : [0, \infty) \rightarrow [0, \infty)$ dos funciones continuas con soporte compacto, tal que $g(0) = g'(0) = 0$. Entonces,*

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

donde $\|g - g'\|$ se refiere a la norma uniforme de $g - g' \in C(\mathbb{R}_0^+, \mathbb{R}_0^+)$.

Demostración. Recordemos la definición de la distancia de Gromov-Hausdorff dada por

$$d_{GH}(\mathcal{T}_g, \mathcal{T}_{g'}) = \frac{1}{2} \inf_{\mathcal{R} \in \mathcal{C}(\mathcal{T}_g, \mathcal{T}_{g'}), (\rho, \rho') \in \mathcal{R}} \text{dis}(\mathcal{R}),$$

donde podemos construir una correspondencia \mathcal{R} entre \mathcal{T}_g y $\mathcal{T}_{g'}$ de la siguiente manera:

$$\mathcal{R} = \{(\sigma, \sigma') : \sigma = p_g(t) \text{ y } \sigma' = p_{g'}(t) \text{ para algún } t \geq 0\}.$$

Para poder acotar superiormente la distorsión de \mathcal{R} , sea $(\sigma, \sigma'), (\eta, \eta') \in \mathcal{R}$. Por la construcción de \mathcal{R} , es posible encontrar $s, t \geq 0$ tal que $p_g(s) = \sigma$, $p_{g'}(s) = \sigma'$ y $p_g(t) = \eta$, $p_{g'}(t) = \eta'$. Ahora, recordemos que

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

de manera que,

$$|d_g(\sigma, \eta) - d_{g'}(\sigma', \eta')| \leq 4 \|g - g'\|.$$

Por lo que tendremos que $\text{dis}(\mathcal{R}) \leq 4\|g - g'\|$, de donde se sigue el resultado. \square

El Lema 100 nos sugiere la siguiente prueba alternativa del Teorema 95. Denotemos por \mathcal{C}_{00} , al conjunto de funciones $g : [0, \infty) \rightarrow [0, \infty)$ tal que

1. g es continua con soporte compacto y $g(0) = 0$.
2. Existe $\epsilon > 0$ y $\alpha > 0$ tal que para todo entero no negativo i , g es lineal con pendiente α ó $-\alpha$ sobre el intervalo $[(i-1)\epsilon, i\epsilon]$.

Entonces no es difícil verificar que \mathcal{T}_g es un árbol real si $g \in \mathcal{C}_{00}$. Efectivamente, vía un reescalamiento espacio-tiempo adecuado, g será la función contorno de un árbol discreto $\mathbf{t} \in \mathcal{A}$, y \mathcal{T}_g coincide (a través de un reescalamiento adecuado) con el árbol real que puede ser construido de \mathbf{t} de manera natural. Entonces, una función en general g puede ser vista como límite uniforme de una sucesión $(g_n)_{n \in \mathbb{N}}$ de elementos de \mathcal{C}_{00} , y el Lema 100 implica que \mathcal{T}_g es el límite bajo la métrica de *Gromov-Hausdorff* de la sucesión \mathcal{T}_{g_n} . Ya que cada \mathcal{T}_{g_n} es un árbol real, \mathcal{T}_g también debe serlo por la completitud del espacio (\mathbb{T}, d_{GH}) (Teorema 91).

OBSERVACIÓN. Recordando que todo árbol real con raíz puede ser representado de la forma \mathcal{T}_g , la separabilidad del espacio (\mathbb{T}, d_{GH}) en el Teorema 91 puede ser obtenida como consecuencia del Lema 100 y la separabilidad del espacio de funciones continuas con soporte compacto sobre \mathbb{R}_+ .

5.3. El Continuum Random Tree (CRT)

Utilizamos nuevamente la notación $\mathbf{e} = (\mathbf{e}_t, 0 \leq t \leq 1)$ para denotar a la excursión Browniana normalizada. Recapitulemos, esta es simplemente una excursión Browniana condicionada a tener duración 1. Al fijarnos en la primera excursión positiva β con duración mayor a 1, escribimos $[G, D]$ para el correspondiente intervalo de tiempo en el que ocurre, y hacemos

$$\mathbf{e}_t = \frac{1}{\sqrt{D-G}} \beta_{(G+(D-G)t) \wedge D} \quad , \quad 0 \leq t \leq 1.$$

Por convención $\mathbf{e}_t = 0$ si $t \geq 1$. Más aún, consideramos $(\mathbf{e}_t)_{t \geq 0}$ como una función continua (aleatoria) sobre el intervalo $[0, 1]$ que cumple las propiedades de la sección anterior.

Definición 101. (Continuum Random Tree) *El continuum random tree es el árbol real aleatorio $\mathcal{T}_{\mathbf{e}}$ codificado por la excursión Browniana normalizada.*

El continuum random tree $\mathcal{T}_{\mathbf{e}}$ es entonces una variable aleatoria que toma valores en el espacio (\mathbb{T}, d_{GH}) . Notemos que la medibilidad de ésta variable aleatoria se sigue del Lema 100.

Es posible enunciar muchos de los resultados obtenidos para los árboles “combinatorios” en términos de convergencia débil para el espacio \mathbb{T} . Sin embargo, daremos un típico ejemplo mostrando que el continuum random tree es límite de árboles combinatorios reescalados.

Recordemos \mathcal{A} denota al conjunto de árboles finitos ordenados con raíz. Introducimos la notación \mathcal{A}_n para denotar al subconjunto de \mathcal{A} que consiste de árboles con n vértices. Podemos y en lo sucesivo veremos a cada elemento $\mathbf{t} \in \mathcal{A}$ como un árbol real con raíz. Simplemente es cuestión de visualizar al árbol \mathbf{t} como una gráfica, unión de segmentos de longitud 1 sobre el plano, equipado con la distancia naturalmente definida como: la distancia entre σ y σ' es la longitud de la trayectoria más corta sobre el árbol.

De manera alternativa, si $C = (C_{\mathbf{t}}(t), t \geq 0)$ es la función contorno del árbol, esto significa que identificamos $\mathbf{t} = \mathcal{T}_C$ (no es una identificación *per se*, porque el árbol \mathbf{t} cuenta con un orden en su estructura que desaparece al considerarle un árbol real).

Para toda $\lambda > 0$ y un árbol $\mathcal{T} \in \mathbb{T}$, el árbol $\lambda\mathcal{T}$ es el “mismo” árbol con todas las distancias multiplicadas por el factor λ (si \mathcal{T} se encuentra encajado sobre el plano

como se sugirió visualizar al árbol anteriormente, esto correspondería a reemplazar el conjunto \mathcal{T} por $\lambda\mathcal{T}$).

Ahora reformulamos los teoremas de convergencia en distribución hacia la excursión browniana en términos de convergencia en distribución de árboles discretos hacia el Continuum Random Tree sobre el espacio (\mathcal{T}, d_{GH}) .

Teorema 102.

Para toda $n \geq 1$, sea $\mathcal{T}_{(n)}$ un árbol aleatorio con distribución uniforme sobre \mathcal{A}_n . Entonces

$$\frac{1}{\sqrt{2n}} \mathcal{T}_{(n)} \xrightarrow[n \rightsquigarrow \infty]{\mathfrak{D}} \mathcal{T}_e,$$

en el sentido de convergencia en distribución para variables aleatorias con valores en (\mathbb{T}, d_{GH}) .

Demostración. Sea, C_k la función contorno de $\mathcal{T}_{(k)}$, y definamos la versión reescalada de esta estableciendo

$$\tilde{C}_k = (2k)^{-1/2} C_k(2kt) \text{ para toda } t \geq 0.$$

Notemos que la función \tilde{C}_k es continua no negativa y se anula en 0 y 1. Por lo tanto es posible definir al árbol $\mathcal{T}_{\tilde{C}_k}$.

Observemos ahora que este último árbol real que acabamos de definir, está relacionado muy cercanamente al árbol $\mathcal{T}_{(k)}$. En efecto, $\mathcal{T}_{\tilde{C}_k}$ es (isométrico a) una unión finita de segmentos de longitud $(2k)^{-1/2}$ en el plano, con estructura genealógica regida por aquella de $\mathcal{T}_{(k)}$ de la forma ya ilustrada por una función contorno. De esta última observación y por la definición de la distancia Gromov-Hausdorff, tendremos que

$$d_{GH}(\mathcal{T}_{\tilde{C}_k}, (2k)^{-1/2} \mathcal{T}_{(k)}) \leq (2k)^{-1/2}. \tag{5.3}$$

Por el Teorema 86, tenemos que

$$(\tilde{C}_n(t), t \geq 0) \xrightarrow[n \rightsquigarrow \infty]{\mathfrak{D}} (\mathbf{e}_t, t \geq 0).$$

Combinando esto último con el Lema 100, obtenemos que

$$d_{GH}(\mathcal{T}_{\tilde{C}_n}, \mathcal{T}_e) \leq 2\|\tilde{C}_n - \mathbf{e}\| \rightarrow 0, \quad n \rightarrow \infty.$$

De donde obtenemos que

$$(\mathcal{T}_{\tilde{C}_n}, d_{\tilde{C}_n}) \xrightarrow[n \rightsquigarrow \infty]{\mathfrak{D}} (\mathcal{T}_e, d_e).$$

El presente teorema se sigue entonces de ésta última convergencia y (5.3). □

OBSERVACIÓN Este último resultado de hecho contiene menos información que los teoremas de convergencia a la excursión browniana antes vistos, ya que el orden lexicográfico inherente a la noción de árbol finito ordenado (y a la codificación de árboles reales por funciones) desaparece cuando consideramos al árbol finito ordenado como un espacio métrico. Aún así es importante pensar al Continuum Random Tree como límite de árboles discretos reescalados.

Existen resultados análogos al teorema anterior para distintas clases de árboles combinatorios, tal y como se comentó en la parte final de la subsección 5.5. En vista de esto, si $\mathcal{T}_{(n)}$ se distribuye de manera uniformemente entre la clase de árboles de Cayley con raíz y n vértices, entonces tendríamos la siguiente convergencia al Continuum Random Tree,

$$\frac{1}{\sqrt{4n}}\mathcal{T}_{(n)} \xrightarrow[n \rightsquigarrow \infty]{\mathfrak{D}} \mathcal{T}_e,$$

sobre el espacio \mathbb{T} . Notemos que los árboles Cayley no están ordenados, pero pueden ser obtenidos de árboles (ordenados) Galton-Watson con distribución de prole Poisson “olvidando” el orden. Aplicando el mismo argumento de la prueba anterior a éstos árboles Galton-Watson (condicionados), obtenemos la convergencia deseada para los árboles de Cayley.

5.4. La Medida de Excursión de Itô.

Nuestro objetivo ahora será obtener ciertas distribuciones explícitas para el Continuum Random Tree, y más específicamente sus respectivas distribuciones finito-dimensionales. Para llevar a cabo estos cálculos, necesitaremos algunas propiedades básicas de las excursiones Brownianas. Antes de lidiar con la excursión Browniana normalizada, consideraremos la medida de Itô.

Denotamos por $(B_t, t \geq 0)$ a un movimiento Browniano en \mathbb{R} , comenzado en x bajo la medida de probabilidad \mathbb{P}_x . Establecemos

$$S_t = \sup_{s \leq t} B_s, \quad I_t = \inf_{s \leq t} B_s$$

y para $a \in \mathbb{R}$. El principio de reflexión nos permite obtener la distribución del vector (S_t, B_t) : Si $a \geq 0$ y $b \in (-\infty, a]$,

$$\mathbb{P}_0[S_t \leq a, B_t \geq b] = \mathbb{P}_0[B_t \geq 2a - b].$$

Tendremos entonces que para toda $t > 0$, la densidad conjunta bajo la medida \mathbb{P}_0 de la pareja (S_t, B_t) es

$$\gamma_t(a, b) = \frac{2(2a - b)}{\sqrt{2\pi t^3}} \exp\left(-\frac{(2a - b)^2}{2t}\right) \mathbf{1}_{\{a \geq 0, b \leq a\}}.$$

El principio de reflexión también implica que S_t y $|B_t|$ se distribuyen idénticamente. Sea $a > 0$, definimos el tiempo de paro $T_a = \inf\{t \geq 0 : B_t = a\}$. Observando que $\{T_a \leq t\} = \{S_t \geq a\}$, obtenemos que la densidad de T_a bajo \mathbb{P}_0 , es la función

$$q_a(t) = \frac{a}{\sqrt{2\pi t^3}} \exp\left(-\frac{a^2}{2t}\right).$$

Notemos que entonces tendremos la siguiente relación,

$$\gamma_t(a, b) = 2q_{2a-b}(t) \quad \text{para } a \geq 0 \text{ y } b < a.$$

Para toda $\epsilon > 0$ denotamos por ϑ_ϵ a la ley de la primera excursión de B afuera de 0 que toca el nivel ϵ . De manera más específica, sea

$$G_\epsilon = \sup\{t < T_\epsilon : B_t = 0\} \quad \text{y} \quad D_\epsilon = \sup\{t > T_\epsilon : B_t = 0\},$$

entonces ϑ_ϵ es la ley del proceso

$$(B_{(G_\epsilon+t) \wedge D_\epsilon}, t \geq 0).$$

La medida ϑ_ϵ así definida es entonces una medida de probabilidad sobre el conjunto $\mathbf{C} = C(\mathbb{R}_+, \mathbb{R}_+)$ de todas las funciones continuas de \mathbb{R}_+ en \mathbb{R}_+ , y tiene soporte en

$$\mathbf{C}_\epsilon = \left\{ e \in \mathbf{C} : \sup_{s \geq 0} e_s \geq \epsilon \right\}.$$

Recordemos que intuitivamente el soporte de una distribución puede pensarse como la cerradura de los posibles valores tomados por una variable aleatoria que tenga dicha distribución.

Si $0 < \epsilon < \epsilon'$, tenemos que

$$\vartheta_\epsilon(\mathbf{C}_{\epsilon'}) = \mathbb{P}_\epsilon[S_{T_0} \geq \epsilon'] = \mathbb{P}_\epsilon[T_{\epsilon'} < T_0] = \frac{\epsilon}{\epsilon'}$$

de donde por definición de probabilidad condicional obtenemos que

$$\vartheta_{\epsilon'} = \vartheta_\epsilon(\cdot | \mathbf{C}_{\epsilon'}) = \frac{\epsilon'}{\epsilon} \vartheta_\epsilon(\cdot \cap \mathbf{C}_{\epsilon'}).$$

Para toda $\epsilon > 0$, establecemos

$$\eta_\epsilon = \frac{1}{2\epsilon} \vartheta_\epsilon.$$

Luego,

$$\eta_{\epsilon'} = \eta_\epsilon(\cdot \cap \mathbf{C}_{\epsilon'})$$

para toda $\epsilon' \in [\epsilon, +\infty)$. Esto conduce a la siguiente definición.

Definición 103. (Medida de Itô). *La medida σ -finita η sobre \mathbf{C} definida por*

$$\eta = \lim_{\epsilon \downarrow 0} \eta_\epsilon,$$

es la medida de Itô de las excursiones positivas de un movimiento Browniano real.

Algunas propiedades de la medida de Itô son las siguientes:

- $\eta(de)$ tiene soporte sobre el conjunto
 $\mathcal{E} = \{e \in \mathbf{C} : \text{Existe } \sigma = \sigma(e) > 0 \text{ tal que } e_t > 0 \text{ si y sólo si } 0 < t < \sigma\}.$

El número $\sigma(e)$ es llamado la longitud o duración de la excursión e .

- Por construcción, η_ϵ es la restricción de la medida η al espacio \mathbf{C}_ϵ y en particular

$$\eta(\mathbf{C}_\epsilon) = \frac{1}{2\epsilon}.$$

- Sea $T_\epsilon(e) = \inf\{t \geq 0 : e_t = \epsilon\}$, entonces la ley del proceso

$$(e_{T_\epsilon(e)+t}, t \geq 0)$$

bajo la medida $\vartheta_\epsilon = \eta(\cdot | T_\epsilon < \infty)$ es idéntica a la ley del proceso

$$(B_{t \wedge T_0}, t \geq 0)$$

bajo la medida \mathbb{P}_ϵ . Esta última propiedad se obtiene de la construcción de ϑ_ϵ y la propiedad fuerte de Markov para el movimiento Browniano al tiempo T_ϵ .

Recordemos que si τ_ℓ es la función inversa continua por la derecha del tiempo local, entonces para cada $\ell \in D_\tau$ (conjunto de discontinuidades de τ), definimos la excursión $e_\ell = (e_\ell(t))_{t \geq 0}$ asociada al intervalo $(\tau_{\ell-}, \tau_\ell)$ por:

$$e_\ell(t) = \begin{cases} \beta_{\tau_{\ell-}+t} & \text{si } 0 \leq t \leq \tau_\ell - \tau_{\ell-}, \\ 0 & \text{si } t > \tau_\ell - \tau_{\ell-}. \end{cases}$$

El siguiente teorema es un resultado básico de la teoría de excursiones y esencial en el contexto del presente trabajo.

Teorema 104. *La medida puntual*

$$\sum_{\ell \in D_\tau} \delta_{(\ell, e_\ell)}(ds, de),$$

es una medida de Poisson puntual sobre $\mathbb{R}_+ \times \mathfrak{E}$, con intensidad

$$ds \otimes \eta(de)$$

donde $\eta(de)$ es la medida de Itô sobre \mathfrak{E} .

El siguiente corolario se sigue de propiedades estándar de medidas de Poisson.

Corolario 105. *Sea A un boreliano subconjunto de \mathfrak{E} tal que $0 < \eta(A) < 1$, y sea*

$$T_A = \inf\{\ell \in D_\tau : e_\ell \in A\}.$$

Entonces, T_A tiene distribución exponencial de parámetro $\eta(A)$, y la distribución de e_{T_A} es la medida condicional

$$\eta(\cdot | A) = \frac{\eta(\cdot \cap A)}{\eta(A)}.$$

Más aún, T_A y e_{T_A} son independientes.

Para una prueba de estos resultados véase [17] Capítulo XII —aquí nuestra medida de Itô η corresponde a la medida n_+ .

Proposición 106. *Sea η la medida de Itô, entonces:*

(i) *Para toda $t > 0$ y toda función medible $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ tal que $g(0) = 0$,*

$$\int \eta(de) g(e_t) = \int_0^\infty dx q_x(t) g(x). \quad (5.4)$$

En particular, $\eta(\sigma > t) = \eta(e_t > 0) = \frac{1}{\sqrt{2\pi t}} < \infty$. Más aún,

$$\eta\left(\int_0^\infty dt g(e_t)\right) = \int_0^\infty dx g(x). \quad (5.5)$$

(ii) Sea $t > 0$ y sean Φ y Ψ dos funciones medibles no-negativas definidas respectivamente sobre $C([0, t], \mathbb{R}_+)$ y \mathbf{C} . Entonces,

$$\begin{aligned} & \int \eta(\mathrm{d}e) \Phi(e_r, 0 \leq r \leq t) \Psi(e_{t+r}, r \geq 0) \\ &= \int \eta(\mathrm{d}e) \Phi(e_r, 0 \leq r \leq t) \mathbb{E}_{e_t} [\Psi(B_{r \wedge T_0}, r \geq 0)], \end{aligned}$$

esto es, el proceso $(e_t, t \geq 0)$ es Markoviano bajo la medida de Itô.

Demostración. (i) Toda función medible y acotada puede ser aproximada por funciones simples y a su vez estas pueden ser aproximadas puntualmente por funciones continuas y acotadas; tendremos entonces que gracias al Teorema de Convergencia Dominada de Lebesgue, basta asumir que g es una función acotada, continua y para la cual existe $\alpha > 0$ tal que $g(x) = 0$ si $x \leq \alpha$. Utilizando la igualdad en ley de los procesos $(B_{T_0 \wedge t})_{t \geq 0}$ y $(e_{T_\epsilon(e) + t})_{t \geq 0}$ respecto a las medidas de probabilidad \mathbb{P}_ϵ y $\vartheta_\epsilon = \eta(\cdot | T_\epsilon < \infty)$ respectivamente, tendremos que:

$$\begin{aligned} \int \eta(\mathrm{d}e) g(e_t) &= \lim_{\epsilon \downarrow 0} \int \eta_\epsilon(\mathrm{d}e) g(e_t) \\ &= \lim_{\epsilon \downarrow 0} \frac{1}{2\epsilon} \int \vartheta_\epsilon(\mathrm{d}e) g(e_t) \\ &= \lim_{\epsilon \downarrow 0} \frac{1}{2\epsilon} \int \eta(\mathrm{d}e) g(e_{T_\epsilon(e) + t}) \mathbf{1}_{\{T_\epsilon(e) < \infty\}} \\ &= \lim_{\epsilon \downarrow 0} \frac{1}{2\epsilon} \mathbb{E}_\epsilon [g(B_{t \wedge T_0})]. \end{aligned}$$

Por otro lado, sea γ la densidad de la pareja (S_t, B_t) como definida al inicio de la sección, tendremos que

$$\begin{aligned} \mathbb{E}_\epsilon [g(B_{t \wedge T_0})] &= \mathbb{E}_\epsilon [g(B_t) \mathbf{1}_{\{t < T_0\}}] \\ &= \mathbb{E}_\epsilon [g(B_t) \mathbf{1}_{\{I_t > 0\}}] \\ &= \mathbb{E}_0 [g(\epsilon + B_t) \mathbf{1}_{\{I_t > -\epsilon\}}] \\ &= \mathbb{E}_0 [g(\epsilon - B_t) \mathbf{1}_{\{S_t \leq \epsilon\}}] \quad (\text{Simetría}) \\ &= \int_0^\epsilon \mathrm{d}y \int_{-\infty}^y \mathrm{d}x g(\epsilon - x) \gamma_t(y, x) \\ &= \int_0^\epsilon \mathrm{d}y \int_{\epsilon - y}^{+\infty} \mathrm{d}x g(x) \gamma_t(y, \epsilon - x). \end{aligned}$$

Así las cosas, la identidad (5.4) del enunciado (i) se seguirá de tomar el límite cuando $\epsilon \downarrow 0$ y considerando la identidad $q_x(t) = \frac{1}{2} \gamma_t(0, -x)$. Recordando ahora que $\sigma = \sigma(e)$ es el tiempo en el cual se cumple:

$$e_t > 0 \text{ si y sólo si } 0 < t < \sigma,$$

tendremos que

$$\eta(\sigma > t) = \eta(e_t > 0) = \frac{1}{\sqrt{2\pi t}} < \infty.$$

Esta última igualdad se obtiene tomando $g(x) = \mathbf{1}_{\{x>0\}}$ de la siguiente manera:

$$\begin{aligned} \eta(e_t > 0) &= \int \eta(de) \mathbf{1}_{\{e_t > 0\}} = \int_0^\infty dx q_x(t) \mathbf{1}_{\{x > 0\}} \\ &= \frac{1}{2} \int_0^\infty dx \gamma_t(0, -x) \\ &= \frac{1}{2} \int_0^\infty dx \frac{2x}{\sqrt{2\pi t^3}} e^{-x^2/2t} = \frac{1}{\sqrt{2\pi t}}, \end{aligned}$$

donde hemos utilizado que $q_x(t) = \frac{1}{2} \gamma_t(0, -x)$ y en la última igualdad un simple cambio de variable. Finalmente, la identidad (5.5) se sigue de (5.4) y recordando que la función $t \rightarrow q_x(t)$ es una densidad de probabilidad, de manera que:

$$\begin{aligned} \int_0^\infty dx g(x) &= \int_0^\infty dx \int_0^\infty dt q_x(t) g(x) \\ &= \int_0^\infty dt \int_0^\infty dx q_x(t) g(x) \\ &= \int_0^\infty dt \int \eta(de) g(e_t) \\ &= \int \eta(de) \int_0^\infty dt g(e_t) = \eta \left(\int_0^\infty dt g(e_t) \right). \end{aligned}$$

(ii) Nuevamente basta con probar el resultado para Φ y Ψ continuas y acotadas, de manera que existe $\alpha \in (0, t)$ tal que $\Phi(\omega(r), 0 \leq r \leq t) = 0$ si $\omega(\alpha) = 0$. La prueba entonces se reduce a la aplicación de la propiedad de Markov del movimiento

Browniano real comenzado en ϵ en el intervalo $[0, T_0]$, escribiendo

$$\begin{aligned}
& \int \eta(de) \Phi(e_r, 0 \leq r \leq t) \Psi(e_{r+t}, r \geq 0) \\
&= \lim_{\epsilon \rightarrow 0} \int \eta(de) \mathbf{1}_{\{T_\epsilon(e) < \infty\}} \Phi(e_{T_\epsilon(e)+r}, 0 \leq r \leq t) \Psi(e_{T_\epsilon(e)+t+r}, r \geq 0) \\
&= \lim_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} \mathbb{E}_\epsilon [\Phi(B_{r \wedge T_0}, 0 \leq r \leq t) \Psi(B_{(t+s) \wedge T_0}, s \geq 0)] \\
&= \lim_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} \mathbb{E}_\epsilon [\mathbb{E}_\epsilon [\Phi(B_{r \wedge T_0}, 0 \leq r \leq t) \Psi(B_{(s+t) \wedge T_0}, s \geq 0) \mid B_{t \wedge T_0}]] \\
&= \lim_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} \mathbb{E}_\epsilon [\Phi(B_{r \wedge T_0}, 0 \leq r \leq t) \mathbb{E}_{B_{t \wedge T_0}} [\Psi(B_{s \wedge T_0}, s \geq 0)]] \\
&= \lim_{\epsilon \rightarrow 0} \int \eta(de) \mathbf{1}_{\{T_\epsilon(e) < \infty\}} \Phi(e_{T_\epsilon(e)+r}, 0 \leq r \leq t) \mathbb{E}_{e(T_\epsilon(e)+t)} [\Psi(B_{s \wedge T_0}, s \geq 0)] \\
&= \int \eta(de) \Phi(e_r, 0 \leq r \leq t) \mathbb{E}_{e_t} [\Psi(B_{r \wedge T_0}, r \geq 0)].
\end{aligned}$$

En la primera e última igualdad, la convergencia dominada es justificada por el hecho de que

$$\mathbf{1}_{\{T_\epsilon(e) < \infty\}} \Phi(e_{T_\epsilon(e)+r}, 0 \leq r \leq t) = 0 \text{ si } \sigma(e) \leq \alpha,$$

y la propiedad $\eta(\sigma > \alpha) < \infty$. □

5.5. Marginales Finito Dimensionales bajo la medida de Itô

Si (\mathcal{T}, d) es un árbol real con raíz ρ , y si $x_1, x_2, \dots, x_p \in \mathcal{T}$, el sub-árbol generado por x_1, x_2, \dots, x_p es simplemente el conjunto

$$\mathcal{T}(x_1, x_2, \dots, x_p) = \bigcup_{i=1}^p [[\rho, x_i]].$$

Es fácil ver que el conjunto $\mathcal{T}(x_1, x_2, \dots, x_p)$, equipado con la métrica d , es nuevamente un árbol real, con una estructura discreta; $\mathcal{T}(x_1, x_2, \dots, x_p)$ puede ser representado por un esqueleto discreto (lo cual se refiere a un árbol discreto con raíz y con sus respectivas p hojas etiquetadas) y a la colección, indexada por los vértices del esqueleto, con longitud la de las “ramas”.

Más que explicitar las definiciones de un árbol real general, nos concentraremos en el caso del árbol real \mathcal{T}_g codificado por g , en el sentido de la subsección 6.2.

Recordemos que $\mathbf{C} = C(\mathbb{R}_+, \mathbb{R}_+)$ y consideremos una función continua arbitraria $g \in \mathbf{C}$, la cual en contraste con las secciones anteriores, ahora no asumiremos que tenga soporte compacto y que cumpla que $g(0) = 0$.

Además recordemos que \mathcal{A} denota al conjunto de árboles finitos ordenados con raíz.

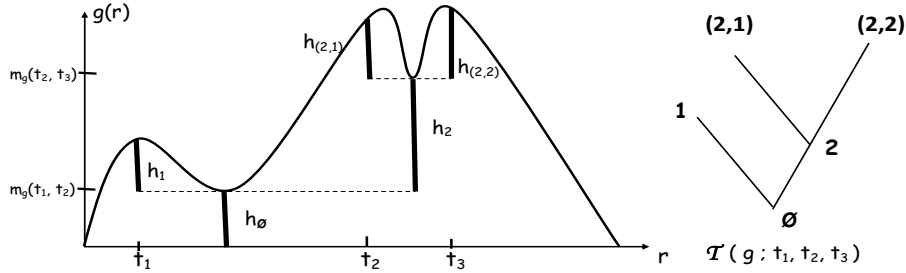
Definición 107. (Árbol Marcado).

Sea $0 \leq t_1 \leq t_2 \leq \dots \leq t_p$, definimos a un árbol marcado como

$$\theta(g; t_1, \dots, t_p) = (\mathcal{T}(g; t_1, \dots, t_p), (h_u)_{u \in \mathcal{T}(g; t_1, \dots, t_p)}),$$

donde $\mathcal{T}(g; t_1, \dots, t_p) \in \mathcal{A}$ y $h_u \geq 0$ para toda $u \in \mathcal{T}(g; t_1, \dots, t_p)$.

Ejemplo. Para $p = 3$, la construcción del árbol marcado correspondiente a los tres puntos elegidos se ilustra en la figura de abajo. En este caso, se tiene que $\mathcal{T}(g; t_1, t_2, t_3) = \{\emptyset, 1, 2, (2,1), (2,2)\}$, como dibujado en la parte derecha de la figura, y donde los números h_u con $u \in \mathcal{T}(g; t_1, t_2, t_3)$ son las longitudes de los segmentos en negrita indicados en la parte izquierda de la figura.



CONSTRUCCIÓN DE UN ÁRBOL MARCADO.

Para ofrecer una definición más precisa de $\theta(g; t_1, t_2, \dots, t_p)$, procederemos por inducción sobre p para su construcción. Para $p = 1$, $\mathcal{T}(g; t_1) = \{\emptyset\}$ y $h_{\emptyset}(g; t_1) = g(t_1)$.

Sea $p \geq 2$. Supongamos que el árbol marcado $\theta(g; t_1, \dots, t_j)$, ha sido construido hasta el orden $p-1$ como hipótesis de inducción. Así las cosas, debe existir un entero $k \in \{1, \dots, p-1\}$ y k enteros $1 \leq i_1 < i_2 < \dots < i_k \leq p-1$, tales que

$$m_g(t_{i-1}, t_i) = m_g(t_1, t_p) \quad \text{si y sólo si} \quad i \in \{i_1, i_2, \dots, i_k\}.$$

Hacemos $i_0 = 0$, $i_{k+1} = p$ por convención. Para todo $\ell \in \{1, 2, \dots, k+1\}$, definimos $g^\ell \in \mathbf{C}$ por medio de la siguiente fórmula:

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

Entonces podemos hacer que $\mathcal{T}(g; t_1, t_2, \dots, t_p)$ sea la concatenación de los árboles

$$\mathcal{T}(g^\ell; t_{i_{\ell-1}+1}, \dots, t_{i_\ell}) \text{ para } 1 \leq \ell \leq k+1.$$

De manera más específica,

$$\mathcal{T}(g; t_1, t_2, \dots, t_p) = \{\emptyset\} \cup \bigcup_{\ell=1}^{k+1} \{\ell u : u \in \mathcal{T}(g^\ell; t_{i_{\ell-1}+1}, \dots, t_{i_\ell})\}.$$

Asimismo, si para $\ell \in \{1, 2, \dots, k+1\}$

$$\theta(g^\ell; t_{i_{\ell-1}+1}, \dots, t_{i_\ell}) = \left(\mathcal{T}(g^\ell; t_{i_{\ell-1}+1}, \dots, t_{i_\ell}), (h_u^\ell)_{u \in \mathcal{T}(g^\ell; t_{i_{\ell-1}+1}, \dots, t_{i_\ell})} \right),$$

definimos las *marcas* $(h_u^\ell)_{u \in \mathcal{T}(g^\ell; t_{i_{\ell-1}+1}, \dots, t_{i_\ell})}$ haciendo,

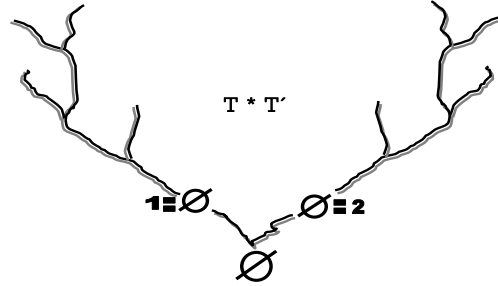
$$h_v = h_u^\ell, \text{ si } v = \ell u, u \in \mathcal{T}(g^\ell; t_{i_{\ell-1}+1}, \dots, t_{i_\ell}), \quad h_\emptyset = m_g(t_1, t_p).$$

Esto completa la construcción del árbol por inducción. Notemos que $k+1$ es el número de hijos de \emptyset en el árbol $\theta(g; t_1, \dots, t_p)$, y $m_g(t_1, t_p)$ es la marca de \emptyset .

Si ahora g satisface las condiciones que suprimimos inicialmente, de tener soporte compacto y cumplir $g(0) = 0$, es fácil ver que $\theta(g; t_1, \dots, t_p)$ corresponde al árbol real $\mathcal{T}_g(p_g(t_1), \dots, p_g(t_p))$ generado por los vértices $p_g(t_1), \dots, p_g(t_p)$ en el árbol \mathcal{T}_g . Siendo más minuciosos en nuestra descripción, si a toda $u \in \mathcal{T}(g; t_1, \dots, t_p)$ le pegamos un segmento de recta en el plano con longitud h_u , de manera que los segmentos de recta adheridos a u y sus hijos compartan un mismo punto extremo (el mismo para todos los hijos de u) y que de otra manera los segmentos de recta no se intersecten, la unión de los respectivos segmentos de recta nos dará un representante de la clase de equivalencia de $\mathcal{T}_g(p_g(t_1), \dots, p_g(t_p))$. (Nótese que la estructura de orden de $\mathcal{T}(g; t_1, \dots, t_p)$ no juega ningún papel en dicha construcción.)

Ahora bien, consideramos al subespacio de \mathcal{A} ,

$$\mathcal{A}_{(p)} = \left\{ \text{Árboles finitos ordenados con raíz y } p \text{ hojas} \right\},$$



y recordemos que una hoja de un árbol $\mathcal{T} \in \mathcal{A}$ es un vértice $u \in \mathcal{T}$ sin descendencia, es decir, tal que $k_u(\mathcal{T}) = 0$ con la notación dada en la primera sección de este capítulo. Denotamos ahora,

$$\Theta_{(p)} = \left\{ \text{Árboles marcados finitos con } p \text{ hojas.} \right\},$$

cuyos elementos son de la forma:

$$\theta = (\mathcal{T}, (h_u)_{u \in \mathcal{T}}) \text{ donde } \mathcal{T} \in \mathcal{A}_{(p)} \text{ y } h_u \geq 0 \text{ para toda } u \in \mathcal{T}.$$

Obviamente consideramos a este conjunto como subespacio de \mathcal{A} y por lo tanto con la topología de subespacio y su respectiva σ -álgebra de Borel. Consideramos también los correspondientes conjuntos de árboles binarios:

$$\mathcal{A}_{(p)}^{\text{bin}} = \left\{ \text{Árboles binarios con raíz y } p \text{ hojas.} \right\}$$

cuyos elementos claramente tienen $2p - 1$ vértices y su subespacio

$$\Theta_{(p)}^{\text{bin}} = \mathcal{A}_{(p)}^{\text{bin}} \cap \Theta_{(p)},$$

en otras palabras, $\theta \in \Theta_{(p)}^{\text{bin}}$ si $\theta = (\mathcal{T}, (h_u)_{u \in \mathcal{T}})$ y su esqueleto \mathcal{T} pertenece a $\mathcal{A}_{(p)}^{\text{bin}}$.

Finalmente, recordemos que

$$\#\mathcal{A}_{(p)}^{\text{bin}} = \frac{(2p - 2)!}{(p - 1)!p!} =: c_p$$

es el número conocido como “Catalán” de orden $p - 1$.

Debería ser claro de nuestra construcción que el árbol $\theta(e; t_1, \dots, t_p)$ sólo depende de los valores $e(t_1), \dots, e(t_p)$ y los mínimos sucesivamente $m_e(t_1, t_2), \dots, m_e(t_{p-1}, t_p)$, de manera que la construcción de un árbol finito se puede representar a través de una función $\Gamma_p : \mathbb{R}_+^{2p-1} \rightarrow \Theta_{(p)}$ medible tal que

$$\Gamma_p(m_e(t_1, t_2), \dots, m_e(t_{p-1}, t_p), e(t_1), \dots, e(t_p)) = \theta(e; t_1, \dots, t_p).$$

Nuestro objetivo ahora será determinar la ley del árbol $\theta(e; t_1, \dots, t_p)$ cuando e es escogido de acuerdo a la medida de excursiones de Itô y (t_1, \dots, t_p) de acuerdo a la medida de Lebesgue sobre $[0, \sigma(e)]_1 \times \dots \times [0, \sigma(e)]_p$. Para simplificar la notación escribiremos $m(s, t)$ en vez de $m_e(s, t)$.

Proposición 108. *Sea f no-negativa medible sobre \mathbb{R}_+^{2p-1} . Entonces,*

$$\eta \left(\int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \sigma\}} dt_1 \dots dt_p f(m(t_1, t_2), \dots, m(t_{p-1}, t_p), e(t_1), \dots, e(t_p)) \right) =$$

$$2^{p-1} \int_{\mathbb{R}_+^{2p-1}} d\alpha_1 \dots d\alpha_{p-1} d\beta_1 \dots d\beta_p \left(\prod_{i=1}^{p-1} \mathbf{1}_{\{\alpha_i \leq \beta_i \wedge \beta_{i+1}\}} \right) f(\alpha_1, \dots, \alpha_{p-1}, \beta_1, \dots, \beta_p).$$

Antes de proceder a probar esta proposición, establecemos el siguiente lema el cual es consecuencia inmediata de conocer la distribución de la pareja (S_t, B_t) . Recordemos que B es un movimiento Browniano que comienza en x bajo la probabilidad \mathbb{P}_x y que $I = (I_t, t \geq 0)$ es el proceso mínimo asociado, es decir $I_t = \inf_{s \in [0, t]} B_s$.

Lema 109. *Si g es una función no-negativa y medible sobre \mathbb{R}^3 y $x \geq 0$,*

$$\mathbb{E}_x \left(\int_0^{T_0} dt g(t, I_t, B_t) \right) = 2 \int_0^x dy \int_y^\infty dz \int_0^\infty dt q_{x+z-2y}(t) g(t, y, z) \quad (5.6)$$

En particular, si h es una función medible no-negativa sobre \mathbb{R}^2 ,

$$\mathbb{E}_x \left(\int_0^{T_0} dt h(I_t, B_t) \right) = 2 \int_0^x dy \int_y^\infty dz h(y, z). \quad (5.7)$$

Demostración del Lema 109.

Intercambiando el orden de integración gracias a Fubini y como consecuencia del

principio de reflexión y la propiedad de simetría de un movimiento browniano, obtenemos que

$$\begin{aligned}
\mathbb{E}_x \left(\int_0^{T_0} dt g(t, I_t, B_t) \right) &= \int_0^\infty dt \mathbb{E}_x (\mathbf{1}_{\{t \leq T_0\}} g(t, I_t, B_t)) \\
&= \int_0^\infty dt \mathbb{E}_x (\mathbf{1}_{\{I_t > 0\}} g(t, I_t, B_t)) = \int_0^\infty dt \mathbb{E}_0 (\mathbf{1}_{\{x+I_t > 0\}} g(t, x+I_t, x+B_t)) \\
&= \int_0^\infty dt \mathbb{E}_0 (\mathbf{1}_{\{x-I_t > 0\}} g(t, x-I_t, x-B_t)) \\
&= \int_{-\infty}^{+\infty} da \int_{-\infty}^{+\infty} db \int_0^\infty dt \mathbf{1}_{\{x-a > 0\}} g(t, x-a, x-b) \gamma_t(a, b) \\
&= \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} dz \int_0^\infty dt \mathbf{1}_{\{y > 0\}} g(t, y, z) \gamma_t(x-y, x-z) \\
&= \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} dz \int_0^\infty dt \mathbf{1}_{\{y > 0\}} g(t, y, z) 2q_{2(x-y)-(x-z)}(t) \mathbf{1}_{\{x \geq y, y \leq z\}} \\
&= 2 \int_0^x dy \int_y^\infty dz \int_0^\infty dt q_{x+z-2y}(t) g(t, y, z),
\end{aligned}$$

donde en la antepenúltima y penúltima igualdad se efectuó un cambio de variable y se utilizó la identidad $2q_{2a-b}(t) = \gamma_t(a, b) \mathbf{1}_{\{a \geq 0, b \leq a\}}$ respectivamente. \square

Demostración de la Proposición 108.

Esto es simple consecuencia del Lema 109. Para $p = 1$, tendríamos que

$$\eta \left(\int_0^{\sigma(e)} f(e_t) dt \right) = \int_0^{+\infty} f(x) dx,$$

que es exactamente el resultado 5.5 de la Proposición 106 ya demostrada.

Ahora procedemos por inducción sobre p . Utilizando la propiedad de Markov bajo η (Proposición 106 (ii)) y posteriormente la relación (5.7), se tiene que:

$$\eta \left(\int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \sigma\}} dt_1 \dots dt_p f(m(t_1, t_2), \dots, m(t_{p-1}, t_p), e(t_1), \dots, e(t_p)) \right)$$

$$\begin{aligned}
&= \int \eta(de) \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \sigma\}} dt_1 \dots dt_p f(m(t_1, t_2), \dots, m(t_{p-1}, t_p), e(t_1), \dots, e(t_p)) \\
&= \int \eta(de) \int dt_1 \dots dt_p \\
&\quad \mathbf{1}_{\{0 \leq t_1 \leq \dots \leq t_p \leq \sigma\}} f(m(t_1, t_2), \dots, m(t_{p-1}, t_p), e(t_1), \dots, e(t_p)) \\
&= \int \eta(de) \int dt_1 \dots dt_{p-1} \mathbf{1}_{\{0 \leq t_1 \leq \dots \leq t_{p-1} \leq \sigma\}} \\
&\quad \left[\int_{t_{p-1}}^{\sigma(e)} dt_p f(m(t_1, t_2), \dots, m(t_{p-1}, t_p), e(t_1), \dots, e(t_p)) \right] \\
&= \int dt_1 \dots dt_{p-1} \int \eta(de) \mathbf{1}_{\{0 \leq t_1 \leq \dots \leq t_{p-1} \leq \sigma\}} \\
&\quad \left[\int dt \mathbf{1}_{[t_{p-1}, \sigma]}(t) f(m^{1,2}, \dots, m^{p-1}(t), e_1, \dots, e_{p-1}, e(t)) \right] \\
&= \int \eta(de) \int dt_1 \dots dt_{p-1} \mathbf{1}_{\{0 \leq t_1 \leq \dots \leq t_{p-1} \leq \sigma\}} \\
&\quad \mathbb{E}_{e(t_{p-1})} \left[\int dt \mathbf{1}_{[0, T_0]} f(m(t_1, t_2), \dots, m(t_{p-2}, t_{p-1}), I_t, e(t_1), \dots, e(t_{p-1}), B_t) \right] \\
&= \eta \left(\int dt_1 \dots dt_{p-1} \mathbf{1}_{\{0 \leq t_1 \leq \dots \leq t_{p-1} \leq \sigma\}} \right. \\
&\quad \left. \mathbb{E}_{e(t_{p-1})} \left[\int dt \mathbf{1}_{[0, T_0]} f(m(t_1, t_2), \dots, m(t_{p-2}, t_{p-1}), I_t, e(t_1), \dots, e(t_{p-1}), B_t) \right] \right)
\end{aligned}$$

y de (5.7) finalmente obtenemos que

$$\begin{aligned}
&= 2 \eta \left(\int_{\{0 \leq t_1 \leq \dots \leq t_{p-1} \leq \sigma\}} dt_1 \dots dt_{p-1} \right. \\
&\quad \left. \int_0^{e(t_{p-1})} d\alpha_{p-1} \int_{\alpha_{p-1}}^{\infty} d\beta_p f(m(t_1, t_2), \dots, m(t_{p-2}, t_{p-1}), \alpha_{p-1}, e(t_1), \dots, e(t_{p-1}), \beta_p) \right).
\end{aligned}$$

Finalmente la prueba es completada utilizando la hipótesis de inducción. \square

Definimos la medida uniforme Λ_p sobre $\Theta_{(p)}^{bin}$ por :

$$\int \Lambda_p(d\theta) F(\theta) = \sum_{\mathcal{T} \in \mathcal{A}_{(p)}^{bin}} \int \prod_{v \in \mathcal{T}} dh_v F(\mathcal{T}, \{h_v, v \in \mathcal{T}\}),$$

para toda $F : \Theta_{(p)}^{bin} \rightarrow \mathbb{R}_+$ medible.

Teorema 110. (Ley del Árbol Marcado)

La ley del árbol $\theta(e; t_1, \dots, t_p)$ bajo la medida

$$\eta(\mathrm{de}) \mathbf{1}_{\{0 \leq t_1 \leq \dots \leq t_p \leq \sigma(e)\}} dt_1 \dots dt_p,$$

es $2^{p-1} \Lambda_p$ donde Λ_p es la medida uniforme sobre $\Theta_{(p)}^{\mathrm{bin}}$ antes mencionada.

Demostración.

Sea $\Gamma_p : \mathbb{R}_+^{2p-1} \rightarrow \Theta_{(p)}$ la función medible dada por

$$\Gamma_p(m(t_1, t_2), \dots, m(t_{p-1}, t_p), e(t_1), \dots, e(t_p)) = \theta(e; t_1, \dots, t_p).$$

La existencia de esta función se sigue fácilmente de nuestra construcción por inducción del árbol marcado $\theta(e; t_1, \dots, t_p)$.

Denotamos por Δ_p la medida sobre \mathbb{R}_+^{2p-1} definida por

$$\Delta_p(\mathrm{d}\alpha_1 \dots \mathrm{d}\alpha_{p-1} \mathrm{d}\beta_1 \dots \mathrm{d}\beta_p) = \left(\prod_{i=1}^{p-1} \mathbf{1}_{[0, \beta_i \wedge \beta_{i+1}]}(\alpha_i) \right) \mathrm{d}\alpha_1 \dots \mathrm{d}\alpha_{p-1} \mathrm{d}\beta_1 \dots \mathrm{d}\beta_p.$$

En vista de la Proposición 108, la prueba del Teorema 110 se reduce a checar que $\Gamma_p(\Delta_p) = \Lambda_p$. Para $p = 1$, esto es obvio.

Sea $p \geq 2$ y supongamos que el resultado es válido hasta el orden de $p - 1$. Para todo $j \in \{1, \dots, p - 1\}$, sea H_j el subconjunto de \mathbb{R}_+^{2p-1} definido por

$$H_j = \{(\alpha_1, \dots, \alpha_{p-1}, \beta_1, \dots, \beta_p) ; \alpha_i > \alpha_j \text{ para toda } i \neq j\},$$

donde al fijar j nos permite asegurar que la colección de dichos conjuntos conforman una partición; tendremos que,

$$\Delta_p = \sum_{j=1}^{p-1} \mathbf{1}_{H_j} \cdot \Delta_p.$$

Por otro lado, es inmediato verificar que $\mathbf{1}_{H_j} \cdot \Delta_p$ es la imagen de la medida

$$\Delta_j(\mathrm{d}\alpha'_1 \dots \mathrm{d}\beta'_j) \otimes \mathbf{1}_{(0, \infty)}(h) \mathrm{d}h \otimes \Delta_{p-j}(\mathrm{d}\alpha''_1 \dots \mathrm{d}\beta''_{p-j})$$

bajo la función $\Psi : (\alpha'_1, \dots, \beta'_j, h, \alpha''_1, \dots, \beta''_{p-j}) \rightarrow (\alpha_1, \dots, \beta_p)$ definida por:

$$\begin{aligned} \alpha_j &= h, \\ \alpha_i &= \alpha'_i + h \quad \text{para } 1 \leq i \leq j - 1, \\ \beta_i &= \beta'_i + h \quad \text{para } 1 \leq i \leq j, \\ \alpha_i &= \alpha''_{i-j} + h \quad \text{para } j + 1 \leq i \leq p - 1, \\ \beta_i &= \beta''_{i-j} + h \quad \text{para } j + 1 \leq i \leq p. \end{aligned}$$

La construcción inductiva del árbol $\theta(e; t_1, \dots, t_p)$ muestra exactamente que, casi en todas partes, para la medida

$$\Delta_j(d\alpha'_1 \dots d\beta'_j) \otimes \mathbf{1}_{(0, \infty)}(h) dh \otimes \Delta_{p-j}(d\alpha''_1 \dots d\beta''_{p-j}),$$

tenemos que

$$\Gamma_p \circ \Psi(\alpha'_1, \dots, \beta'_j, h, \alpha''_1, \dots, \beta''_{p-j}) = \Gamma_{p-j}(\alpha'_1, \dots, \beta'_j) \stackrel{h}{*} \Gamma_{p-j}(\alpha''_1, \dots, \beta''_{p-j}),$$

donde si $\theta \in \Theta_{(j)}$ y $\theta' \in \Theta_{(p-j)}$, el árbol $\theta \stackrel{h}{*} \theta'$ es el que se obtiene a través de la concatenación de los esqueletos discretos de θ y θ' (como anteriormente en la construcción inductiva de $\theta(g; t_1, \dots, t_p)$) y asignando la “marca” h a la raíz \emptyset .

Junto a la hipótesis de inducción, las observaciones previas implican que para cualquier función no-negativa medible f definida sobre $\Theta_{(p)}$,

$$\begin{aligned} & \int \Delta_p(du) \mathbf{1}_{H_j}(u) f(\Gamma_p(u)) \\ &= \int_0^\infty dh \iint \Delta_j(du') \Delta_{p-j}(du'') f(\Gamma_p(\Psi(u', h, u''))) \\ &= \int_0^\infty dh \iint \Delta_j(du') \Delta_{p-j}(du'') f(\Gamma_j(u') \stackrel{h}{*} \Gamma_{p-j}(u'')) \\ &= \int_0^\infty dh \int \Lambda_j \stackrel{h}{*} \Lambda_{p-j}(d\theta) f(\theta) \end{aligned}$$

donde hemos escrito $\Lambda_j \stackrel{h}{*} \Lambda_{p-j}$ para la imagen de $\Lambda_j(d\theta) \Lambda_{p-j}(d\theta')$ bajo el mapeo $(\theta, \theta') \rightarrow \theta \stackrel{h}{*} \theta'$. Para completar la prueba, simplemente notemos que

$$\Lambda_p = \sum_{j=1}^{p-1} \int_0^\infty dh \Lambda_j \stackrel{h}{*} \Lambda_{p-j}.$$

□

OBSERVACIÓN. El hecho de que sólo tengamos árboles binarios en el Teorema 110 proviene de la propiedad de que los mínimos locales del movimiento Browniano sean distintos: Es decir, si $0 < t_1 < \dots < t_p$ y los mínimos locales $m_g(t_i, t_{i+1})$ son distintos, el árbol $\theta(g; t_1, \dots, t_p)$ es claramente binario.

5.6. Marginales Finito Dimensionales del CRT

En esta sección nos proponemos calcular la ley de un árbol $\theta(e; t_1, \dots, t_p)$ cuando e es una excursión Browniana normalizada (la cual denotaremos por \mathbf{e}). Esto corresponde a elegir p vértices de manera independiente e uniforme sobre el Continuum Random Tree (la medida uniforme sobre el Continuum Random Tree $\mathcal{T}_{\mathbf{e}}$ es por definición la imagen de la medida de Lebesgue en $[0, 1]$ bajo la función proyección $p_{\mathbf{e}}$) y determinar la ley del árbol generado por estos vértices.

En contraste con la medida Λ_p del Teorema 110, obtendremos para cada p una medida de probabilidad sobre $\Theta_{(p)}^{\text{bin}}$ a la cual finalmente nos referiremos como la distribución marginal p -dimensional del Continuum Random Tree.

Primero recordamos la conexión entre la medida de Itô y la excursión Browniana normalizada. Nuevamente denotamos por $\sigma = \sigma(e)$ al real positivo tal que $e_t > 0$ si y sólo si $0 < t < \sigma$. Informalmente, la ley de la excursión Browniana normalizada puede entenderse como

$$\eta(\text{de} \mid \sigma(e) = 1).$$

De manera más precisa, utilizando un conocido teorema de desintegración para medidas, junto con la propiedad de escalamiento del movimiento Browniano, es posible demostrar que existe una única colección de medidas de probabilidad $(\eta_{(s)}, s > 0)$ sobre el conjunto \mathcal{E} de excursiones, tal que las siguientes propiedades se mantienen:

- (1) Para toda $s > 0$, $\eta_{(s)}(\sigma = s) = 1$.
- (2) Para toda $\lambda > 0$ y $s > 0$, la ley bajo $\eta_{(s)}(\text{de})$ de $e_{\lambda}(t) = \sqrt{\lambda} e_{t/\lambda}$ es $\eta_{(\lambda s)}$.
- (3) Para todo conjunto de Borel $A \subset \mathcal{E}$,

$$\eta(A) = \frac{1}{2\sqrt{2\pi}} \int_0^{\infty} s^{-3/2} \eta_{(s)}(A) \, ds.$$

Es posible escribir $\eta_{(s)} = \eta(\cdot \mid \sigma = s)$. La medida $\eta_{(1)}$ es la ley de la excursión Browniana normalizada \mathbf{e} .

El primer objetivo es obtener un enunciado más preciso al del Teorema 110 considerando la pareja $(\theta(e; t_1, \dots, t_p), \sigma)$ en vez de $\theta(e; t_1, \dots, t_p)$.

Si $\theta = (\mathcal{T}, \{h_v, v \in \mathcal{T}\})$ es un árbol marcado, la longitud de θ se define de manera natural como

$$L(\theta) = \sum_{v \in \mathcal{T}} h_v.$$

Proposición 111. *La ley de la pareja $(\theta(e; t_1, \dots, t_p), \sigma)$ bajo la medida*

$$\eta(\mathrm{de}) \mathbf{1}_{\{0 \leq t_1 \leq \dots \leq t_p \leq \sigma(e)\}} dt_1 \dots dt_p$$

es

$$2^{p-1} \Lambda_p(\mathrm{d}\theta) q_{2L(\theta)}(s) \mathrm{d}s.$$

Demostración. Recordamos la notación utilizada en la prueba del Teorema 110. Verificaremos que, para toda función no-negativa medible f sobre \mathbb{R}_+^{3p} ,

$$\begin{aligned} & \eta \left(\int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \sigma\}} dt_1 \dots dt_p \right. \\ & \quad \left. f(m(t_1, t_2), \dots, m(t_{p-1}, t_p), e(t_1), \dots, e(t_p), t_1, t_2 - t_1, \dots, \sigma - t_p) \right) \\ &= 2^{p-1} \int \Delta_p(\mathrm{d}\alpha_1 \dots \mathrm{d}\alpha_{p-1} \mathrm{d}\beta_1 \dots \beta_p) \int_{\mathbb{R}_+^{p+1}} \mathrm{d}s_1 \dots \mathrm{d}s_{p+1} q_{\beta_1}(s_1) q_{\beta_1 + \beta_2 - 2\alpha_1}(s_2) \dots \\ & \quad \dots q_{\beta_{p-1} + \beta_p - 2\alpha_{p-1}}(s_p) q_{\beta_p}(s_{p+1}) f(\alpha_1, \dots, \alpha_{p-1}, \beta_1, \dots, \beta_p, s_1, \dots, s_{p+1}). \quad (5.8) \end{aligned}$$

Supongamos que (5.8) es válida. Es posible checar (por inducción sobre p) que

$$2L(\Gamma_p(\alpha_1, \dots, \alpha_{p-1}, \beta_1, \dots, \beta_p)) = \beta_1 + \sum_{i=1}^{p-1} (\beta_i + \beta_{i-1} - 2\alpha_i) + \beta_p.$$

Usando la identidad de la convolución $q_x * q_y = q_{x+y}$ (lo cual es inmediato de la interpretación de q_x como la densidad de T_x bajo \mathbb{P}_0), obtenemos de (5.8) que

$$\begin{aligned} & \eta \left(\int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \sigma\}} dt_1 \dots dt_p f(m(t_1, t_2), \dots, m(t_{p-1}, t_p), e(t_1), \dots, e(t_p), \sigma) \right) \\ &= 2^{p-1} \int \Delta_p(\mathrm{d}\alpha_1 \dots \mathrm{d}\alpha_{p-1} \mathrm{d}\beta_1 \dots \beta_p) \int_0^\infty \mathrm{d}t q_{2L(\Gamma_p(\alpha_1, \dots, \alpha_{p-1}, \beta_1, \dots, \beta_p))}(t) f(\alpha_1, \dots, \beta_p, t). \end{aligned}$$

Como en la prueba del Teorema 110 el enunciado de la presente proposición se sigue de esta última identidad y la igualdad $\Gamma_p(\Delta_p) = \Lambda_p$.

Resta probar (5.8). Para $p = 1$, utilizamos la propiedad de Markov bajo la medida de Itô (Proposición 106 (ii)) y posteriormente la definición de la función q_x .

Finalmente, por la Proposición 106(i), obtenemos que

$$\begin{aligned}
\int \eta(de) \int_0^\sigma dt f(e_t, t, \sigma - t) &= \int \eta(de) \int_0^\sigma dt \mathbb{E}_{e_t} [f(e_t, t, T_0)] \\
&= \int \eta(de) \int_0^\sigma dt \int_0^\infty ds q_{e_t}(s) f(e_t, t, s) \\
&= \int \eta(dx) \int_0^\infty dt q_x(t) \int_0^\infty ds q_x(s) f(x, t, s).
\end{aligned}$$

Sea $p \geq 2$. Aplicando la propiedad de Markov bajo η sucesivamente en t_p y t_{p-1} y posteriormente utilizando la igualdad en (5.6), obtenemos

$$\begin{aligned}
&\eta \left(\int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \sigma\}} dt_1 \dots dt_p \right. \\
&\quad \left. f(m(t_1, t_2), \dots, m(t_{p-1}, t_p), e(t_1), \dots, e(t_p), t_1, t_2 - t_1, \dots, \sigma - t_p) \right) \\
&= \eta \left(\int_{\{0 \leq t_1 \leq \dots \leq t_{p-1} \leq \sigma\}} dt_1 \dots dt_{p-1} \mathbb{E}_{e_{t_{p-1}}} \left[\int_0^{T_0} dt \int_0^\infty ds q_{B_t}(s) \right. \right. \\
&\quad \left. \left. f(m(t_1, t_2), \dots, m(t_{p-2}, t_{p-1}), I_t, e(t_1), \dots, e(t_{p-1}), B_t, t_1, \dots, t_{p-1} - t_{p-2}, t, s) \right] \right) \\
&= 2\eta \left(\int_{\{0 \leq t_1 \leq \dots \leq t_{p-1} \leq \sigma\}} dt_1 \dots dt_{p-1} \right. \\
&\quad \int_0^{e_{t_{p-1}}} dy \int_0^\infty dz \int_0^\infty dt \int_0^\infty ds q_{e_{t_{p-1}} + z - 2y}(t) q_z(s) \\
&\quad \left. f(m(t_1, t_2), \dots, m(t_{p-2}, t_{p-1}), y, e(t_1), \dots, e(t_{p-1}), z, t_1, \dots, t_{p-1} - t_{p-2}, t, s) \right).
\end{aligned}$$

De esta manera, la prueba es completada por inducción sobre p . □

Ahora es posible enunciar el resultado principal de esta sección.

Teorema 112. *La ley del árbol $\theta(e; t_1, \dots, t_p)$ bajo la medida de probabilidad*

$$p! \mathbf{1}_{\{0 \leq t_1 \leq \dots \leq t_p \leq 1\}} dt_1 \dots dt_p \eta_{(1)}(de)$$

es

$$p! 2^{p+1} L(\theta) \exp(-2L(\theta)^2) \Lambda_p(d\theta).$$

Demostración. Sea F una función acotada continua no-negativa sobre $\Theta_{(p)}$ y sea h acotada, no-negativa y medible sobre \mathbb{R}_+ . Por la Proposición 111,

$$\begin{aligned} & \int \eta(\mathrm{d}e) h(\sigma) \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \sigma\}} \mathrm{d}t_1 \dots \mathrm{d}t_p F(\theta(e; t_1, \dots, t_p)) \\ &= 2^{p-1} \int_0^\infty \mathrm{d}s h(s) \int \Lambda_p(\mathrm{d}\theta) \varphi_{2L(\theta)}(s) F(\theta). \end{aligned}$$

Ahora bien, utilizando las propiedades de la definición de las medidas $\eta_{(s)}$ se tiene que

$$\begin{aligned} & \int \eta(\mathrm{d}e) h(\sigma) \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \sigma\}} \mathrm{d}t_1 \dots \mathrm{d}t_p F(\theta(e; t_1, \dots, t_p)) \\ &= \frac{1}{2\sqrt{2\pi}} \int_0^\infty \mathrm{d}s s^{-3/2} h(s) \int \eta_{(s)}(\mathrm{d}e) \\ & \quad \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq s\}} \mathrm{d}t_1 \dots \mathrm{d}t_p F(\theta(e; t_1, \dots, t_p)). \end{aligned}$$

Por comparación con la identidad previa, se obtiene que casi para toda $s > 0$,

$$\begin{aligned} & \int \eta_{(s)}(\mathrm{d}e) \int_{\{0 \leq t_1 \leq \dots \leq t_p \leq \sigma\}} \mathrm{d}t_1 \dots \mathrm{d}t_p F(\theta(e; t_1, \dots, t_p)) \\ &= 2^{p+1} \int \Lambda_p(\mathrm{d}\theta) L(\theta) \exp\left(-\frac{2L(\theta)^2}{s}\right) F(\theta). \end{aligned}$$

Ambos lados de la última ecuación son funciones continuas de s (utilizando la propiedad de escalamiento de $\eta_{(s)}$ para el lado izquierdo). Entonces la igualdad se mantiene para todo $s > 0$, y en particular $s = 1$. Con esto finalizamos la prueba. \square

OBSERVACIONES FINALES. Eligiendo puntos $t_1, \dots, t_p \in [0, 1]$ es posible considerar su arreglo en orden creciente $t'_1 \leq \dots \leq t'_p$ y definir $\theta(e; t_1, \dots, t_p) = \theta(e; t'_1, \dots, t'_p)$. También es posible observar el comportamiento de los puntos con su orden original y considerar al árbol $\tilde{\theta}(e; t_1, \dots, t_p)$ definido como el árbol $\theta(e; t_1, \dots, t_p)$ donde las hojas son etiquetadas $1, \dots, p$, de manera que la hoja correspondiente al tiempo t_i recibe la etiqueta i . (Estas etiquetas no tienen nada que ver con el orden del árbol.) el Teorema 112 implica que la ley del árbol $\tilde{\theta}(e; t_1, \dots, t_p)$ bajo la medida de probabilidad

$$\mathbf{1}_{[0,1]^p}(t_1, \dots, t_p) \mathrm{d}t_1 \dots \mathrm{d}t_p \eta_{(1)}(\mathrm{d}e)$$

tiene densidad

$$2^{p+1} L(\theta) \exp(-2L(\theta)^2)$$

con respecto a $\tilde{\Theta}_{(p)}^{\text{bin}}(d\theta)$, la medida uniforme sobre el conjunto de árboles (*ordenados*) marcados binarios con p hojas.

Es posible entonces olvidar el “ordenamiento”. Supongamos $\tilde{\theta}(e; t_1, \dots, t_p)$ es precisamente el árbol $\tilde{\theta}(e; t_1, \dots, t_p)$ pero sin orden en su estructura. Ya que existen 2^{p-1} posibles ordenamientos para un árbol binario etiquetado con p hojas, obtenemos que la ley (bajo la misma medida) del árbol $\tilde{\theta}(e; t_1, \dots, t_p)$ tiene densidad

$$2^{2p} L(\theta) \exp(-2L(\theta^2))$$

con respecto a $\tilde{\Theta}_{(p)}^{\text{bin}}(d\theta)$, la medida uniforme sobre el conjunto de árboles (*no-ordenados*) marcados binarios con p hojas.

De acuerdo con la normalización del Continuum Random Tree dada por Aldous, reemplazamos la excursión e por $2e$ (esto simplemente indica que todas las marcas h_v son multiplicadas por 2). Obtenemos que la densidad de la ley del árbol $\tilde{\theta}(2e; t_1, \dots, t_p)$ tiene densidad

$$L(\theta) \exp\left(-\frac{L(\theta^2)}{2}\right)$$

con respecto a $\tilde{\Theta}_{(p)}^{\text{bin}}(d\theta)$. Es notable que esta última densidad (aparentemente) no depende de p .

En esta última forma, es posible reconocer las marginales finito-dimensionales del Continuum Random Tree con las que trabajó originalmente Aldous. Dando una descripción más explícita, el esqueleto discreto denotado por $\tilde{T}(2e; t_1, \dots, t_p)$ se distribuye uniformemente sobre el conjunto de árboles binarios con raíz etiquetados y con p hojas. Este conjunto tiene b_p elementos, donde

$$b_p = p! 2^{-(p-1)} c_p = 1 \times 3 \times \dots \times (2p - 3).$$

Entonces, condicionalmente sobre el esqueleto discreto, las marcas h_v se distribuyen con densidad

$$b_p \left(\sum h_v\right) \exp\left(-\frac{(\sum h_v)^2}{2}\right)$$

siendo esta última una densidad de probabilidad sobre \mathbb{R}_+^{2p-1} .

Capítulo 6

Medida de Hausdorff del CRT

Nos proponemos a obtener las medidas de Hausdorff para el Continuum Random Tree y sus conjuntos de nivel y estudiar algunas de sus propiedades. Se seguirá que la medida uniforme sobre el árbol y la medida de tiempo local sobre un conjunto de nivel coinciden con ciertas medidas de Hausdorff.

Antes de dar inicio damos unos preliminares mas generales intentando con esto tener una vista más panorámica de los objetos y definiciones con los que estaremos trabajando.

6.1. Introducción

Sean X, X_1, X_2, \dots v.a.i.i.d. con distribución común F . Consideremos la suma parcial $S_n = X_1 + \dots + X_n$, $n = 1, 2, \dots$, y supongamos que para alguna sucesión de constantes de normalización $A_n > 0$, B_n ($n = 1, 2, \dots$) la sucesión $S_n/A_n - B_n$ tiene un límite en distribución el cuál vamos a considerar Normal.

Establecemos ahora la relación:

$$\frac{S_n}{A_n} - B_n \xrightarrow[n \rightsquigarrow \infty]{\mathcal{D}} N(\mu, \sigma^2)$$

o bien, equivalentemente

$$\lim_{n \rightsquigarrow \infty} \mathbb{P} \left(\frac{S_n}{A_n} - B_n \leq x \right) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (6.1)$$

para todos los puntos de continuidad de la función de distribución de la variable aleatoria $N(\mu, \sigma^2)$. Al conjunto de distribuciones F tal que lo anterior ocurre se le conoce como “dominio de atracción” de la distribución normal. En particular tendremos que la transformada de Fourier de una variable aleatoria $N := N(0, \sigma^2)$, satisface que

$$\mathbb{E} [e^{iuN}] = e^{-\psi(u)},$$

donde

$$\psi(u) = \frac{\sigma^2 u^2}{2}. \quad (6.2)$$

Heurísticamente recordemos que un árbol real es un espacio métrico (\mathcal{T}, d) tal que para cualesquiera dos puntos σ, σ' en \mathcal{T} existe un único arco con puntos extremos σ y σ' ; más aún, este arco es isométrico a un intervalo compacto de la recta real. Los árboles reales que consideramos poseen además un vértice distinguido al cual nos referimos como la raíz.

Escribimos $H(\mathcal{T})$ para la altura de \mathcal{T} , eso es la distancia máxima de la raíz a un vértice en \mathcal{T} . Decimos que dos árboles reales son equivalentes si existe una isometría que preserve la raíz y mapee a uno sobre el otro. Como ya se vió anteriormente el conjunto de estas clases de equivalencias de árboles reales (compactos) con raíz, equipado con la distancia Gromov-Hausdorff conforma un espacio Polaco.

Como ejemplo, tratemos con un simple resultado de aproximación. Sea μ una medida de probabilidad sobre \mathbb{Z}_+ , con $\mu(1) < 1$. Supongamos que μ tiene media uno y esta en el dominio de atracción de una distribución normal, esto último es válido siempre que μ tenga varianza finita. Denotamos por θ a un árbol μ -Galton-Watson, el cual describe la genealogía de un proceso de ramificación Galton-Watson a tiempo discreto con distribución de progenie μ comenzado inicialmente con un ancestro. Si $r > 0$, el árbol a escala $r\theta$ es naturalmente definido requiriendo que la distancia entre dos vértices adyacentes sea r en vez de 1. Aquí $H(\theta)$ representa la máxima generación en θ . Entonces existe una medida σ -finita $\Theta(d\mathcal{T})$ sobre el espacio (\mathbb{T}, d_{GH}) , tal que para toda $a > 0$, la ley condicional del árbol a escala $n^{-1}\theta$ dado que $H(\theta) > an$, converge débilmente cuando $n \rightarrow \infty$ a la medida de probabilidad $\Theta(d\mathcal{T} \mid H(\mathcal{T}) \geq a)$.

El Continuum Random Tree puede interpretarse como el árbol genealógico de un proceso conocido como difusión de Feller el cual puede ser construido como la solución de la ecuación diferencial estocástica

$$dX_t = \sigma \sqrt{X_t} dB_t, \quad X_0 = 1.$$

Esto es, que de cierta manera describe la evolución en tiempo continuo de una

población con valores en la recta real positiva. El proceso X resulta ser en este caso un movimiento browniano cuya ley se caracteriza por la función ψ sobre $[0, \infty)$ dada en (6.2), llamada también “Mecanismo de Ramificación”. Por último, recordemos que la difusión de Feller se extingue en un tiempo finito casi seguramente; esta última condición es necesaria para la compacidad del Continuum Random Tree.

Puede verse en [12] que uno puede asociar una medida σ -finita $\Theta(d\mathcal{T})$ sobre \mathbb{T} la cual identificamos como la ley del árbol real (\mathcal{T}_e, d_e) codificado por e (donde e es una excursión Browniana) bajo la medida de Ito η . La medibilidad de esta variable se tiene gracias al Lema 100. Nótese que η es una medida σ -finita, al igual que Θ ; sin embargo, para toda $\epsilon > 0$

$$\Theta(H(\mathcal{T}_e) > \epsilon) := \eta \left(\sup_{t \in [0, \sigma]} e_t > \epsilon \right) = \frac{1}{2\epsilon} < \infty,$$

en particular denotaremos por $\Theta^{(1)}$ a la ley del Continuum Random Tree \mathcal{T}_e bajo la medida de probabilidad $\eta(de | \sigma(e) = 1) = \eta_{(1)}(de)$.

El Continuum Random Tree cuenta con una importante “Propiedad de Ramificación” análoga al resultado clásico para árboles Galton-Watson: Para toda $a > 0$, bajo la medida de probabilidad $\Theta(\cdot | H(\mathcal{T}) > a)$ y condicionalmente dada la parte de los árboles bajo el nivel a , los subárboles arriba de ese nivel se distribuyen como los átomos de una medida de Poisson puntual cuya intensidad es una variable múltiplo de Θ (el factor aleatorio es la masa total de la medida de tiempo local en el nivel a , $\ell^a(d\sigma) \Theta(d\mathcal{T})$ con notación dada más adelante).

En adición a la propiedad de ramificación, el Continuum Random Tree también posee la siguiente “Propiedad de Escalamiento”: para todo $r > 0$ y para todo árbol $\mathcal{T} \in \mathbb{T}$ denotamos por $r\mathcal{T}$ al “mismo” árbol \mathcal{T} con su métrica d siendo reemplazada por rd . Entonces para toda $r > 0$ la ley de $r\mathcal{T}$ bajo $\Theta(d\mathcal{T})$ esta dada por $r\Theta$.

La construcción de la ley $\Theta^{(1)}$ del Continuum Random Tree es posible obtenerla a partir de la codificación del árbol por su proceso de altura; esta construcción es particularmente conocida para el CRT ya que su proceso de altura es una excursión browniana condicionada a tener duración 1.

Consideremos la siguiente aproximación por árboles discretos. Sea π una distribución de probabilidad sobre $\{0, 1, 2, \dots\}$. Supongamos que π tiene media 1 y es el dominio de atracción de una distribución normal, en el sentido de que existe una sucesión $(A_n)_{n \geq 1}$ de enteros positivos tales que, si ξ_1, ξ_2, \dots son i.i.d. con distribución

π , entonces

$$\frac{\xi_1 + \cdots + \xi_n - n}{A_n} \xrightarrow[n \rightarrow \infty]{\mathfrak{D}} N(0, 1).$$

Sea $c > 0$ una constante y para toda $n \geq 1$ sea θ_n un árbol Galton-Watson con distribución de progenie π condicionada a tener altura mayor a cn . Nótese que θ_n puede ser visto como un árbol real aleatorio al dotar de longitud 1 a cada arista (segmento con vértices adyacentes como puntos extremos). Entonces la distribución de $n^{-1}\theta_n$ converge a la medida de probabilidad $\Theta(\cdot | H(\mathcal{T}) > c)$.

6.2. Medidas del Continuum Random Tree

Antes de enunciar los resultados principales, introducimos medidas aleatorias importantes asociadas al Continuum Random Tree. Para toda $a > 0$, es posible definir Θ c.t.p. (casi en todas partes) una medida finita aleatoria ℓ^a sobre el conjunto de nivel

$$\mathcal{T}(a) := \{\sigma \in \mathcal{T} : d(\rho, \sigma) = a\},$$

la cual de cierta manera se esparce uniformemente sobre el conjunto de nivel: Para toda $\epsilon > 0$ denotamos por $\mathcal{T}_\epsilon(a)$ al subconjunto finito de $\mathcal{T}(a)$ que consiste de aquellos vértices que cuentan con descendencia en el nivel $a + \epsilon$, entonces para toda función continua y acotada φ definida sobre \mathcal{T} ,

$$\langle \ell^a, \varphi \rangle = \lim_{\epsilon \downarrow 0} \frac{1}{\Theta(\sup_{t \in [0, \sigma(e)]} e_t > \epsilon)} \sum_{\sigma \in \mathcal{T}_\epsilon(a)} \varphi(\sigma)$$

donde e denota una excursión browniana.

La medida uniforme $\mathbf{m} = \mathbf{m}_{(\mathcal{T})}$ sobre el árbol \mathcal{T} se define entonces por

$$\mathbf{m} = \int_0^\infty da \ell^a.$$

Ya que el Continuum Random Tree es codificado por una excursión Browniana condicionada a tener duración 1, la ley $\Theta^{(1)}$ del Continuum Random Tree se define informalmente como $\Theta^{(1)} = \Theta(d\mathcal{T} | \mathbf{m}(\mathcal{T}) = 1)$.

En los siguientes resultados, la notación $h - m$ denota la medida de Hausdorff asociada a la función h . Identificamos a continuación con exactitud la función de la medida de Hausdorff para \mathcal{T} y sus conjuntos de nivel.

Teorema 113. *Para toda $r \in (0, 1/2)$, consideremos*

$$h(r) = r^2 \log \log \frac{1}{r}.$$

Entonces existe C_0 , constante positiva tal que Θ c.t.p. para todo boreliano A de \mathcal{T} ,

$$h - m(A) = C_0 \mathbf{m}(A).$$

De acuerdo a este teorema, la medida \mathbf{m} deberá coincidir con cierta medida de Hausdorff sobre \mathcal{T} .

Ya que la excursión normalizada para tener duración 1 y la medida de Itô se relacionan por simples transformaciones de escalamiento, el siguiente corolario es inmediata consecuencia del Teorema 113.

Corolario 114. *El Teorema 113 es válido, con la misma constante C_0 , si Θ es reemplazada por la ley $\Theta^{(1)}$ del CRT.*

Discutamos ahora sobre los conjuntos de nivel. El siguiente teorema nos estaría mostrando que la medida de tiempo local ℓ^a coincide con cierta medida de Hausdorff sobre el conjunto de nivel $\mathcal{T}(a)$.

Teorema 115. *Para todo $r \in (0, 1/2)$, establecemos*

$$\tilde{h}(r) = r \log \log \frac{1}{r}.$$

Entonces existe constante positiva \tilde{C}_0 tal que para toda $a > 0$, uno tiene que Θ c.t.p. para todo boreliano A de $\mathcal{T}(a)$,

$$\tilde{h} - m(A) = \tilde{C}_0 \ell^a(A).$$

A continuación damos resultados básicos de comparación de medidas de Hausdorff que posteriormente utilizaremos para la prueba de los teoremas arriba mencionados.

6.3. Resultados de Comparación para Medidas de Hausdorff

Sea E un espacio métrico compacto. Denotamos por $B_r(x)$ a la bola de radio r centrada en $x \in E$. Sea $c > 1$ fijo, consideraremos al conjunto \mathcal{M}_c de funciones

monótonas crecientes y continuas $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ que cumplan los siguiente:

$$g(0) = 0 \text{ y } g(2r) \leq c g(r).$$

Sea U un subconjunto no vacío de E e I un subconjunto de \mathbb{Z}_+ . Decimos que $\{U_i\}_{i \in I}$ es una δ -cubierta de U si,

$$U \subset \bigcup_{i \in I} U_i \text{ y } 0 < \text{diam}(U_i) \leq \delta.$$

Definimos

$$\mathcal{H}_\delta^g(U) = \inf \left\{ \sum_{i \in I} g(\text{diam}(U_i)) : \{U_i\} \text{ es una } \delta\text{-cubierta de } U \right\}.$$

Entonces observamos sobre todas las posibles δ -cubiertas de U por conjuntos de diámetro a lo más δ buscando minimizar la suma. Al tiempo que δ decrece, la colección de δ -cubiertas permitidas para U se reduce. En consecuencia \mathcal{H}_δ^g aumenta, aproximándose a un límite conforme $\delta \downarrow 0$. Entonces para todo subconjunto U de E , esto nos lleva a definir una medida haciendo

$$g - m(U) = \lim_{\delta \downarrow 0} \mathcal{H}_\delta^g(U),$$

Llamamos a $g - m$ la medida de Hausdorff asociada a g .

Lema 116. *Sea $c > 0$. Existen dos constantes positivas M_1 y M_2 que dependen únicamente de c , tal que la siguiente condición se cumple para toda función $g \in \mathcal{M}_c$. Sea μ una medida de Borel finita sobre E y sea A un boreliano subconjunto de E .*

i.) Si

$$\limsup_{n \rightarrow \infty} \frac{\mu(B_{1/2^n}(x))}{g(1/2^n)} \leq 1, \text{ para todo } x \in A,$$

entonces,

$$g - m(A) \geq M_1 \mu(A).$$

ii.) Si

$$\limsup_{n \rightarrow \infty} \frac{\mu(B_{1/2^n}(x))}{g(1/2^n)} \geq 1, \text{ para todo } x \in A,$$

entonces,

$$g - m(A) \leq M_2 \mu(A).$$

Demostración. i.) Para todo entero $n \geq 1$, hacemos

$$A_n := \left\{ x \in A : \mu \left(B_{1/2^k}(x) \right) \leq 2g \left(1/2^k \right) \text{ para toda } k \geq n \right\}.$$

Es claro que cualquier unión e intersección de la sucesión A_n ($n = 1, 2, \dots$) queda contenida en A , por lo que

$$\limsup_{n \rightarrow \infty} A_n \subseteq A.$$

Si $x \in A$, por hipótesis y el hecho de que el límite inferior es menor igual que el superior tendremos que

$$\liminf_{n \rightarrow \infty} \frac{\mu(B_{1/2^n}(x))}{g(1/2^n)} \leq 1,$$

con esto sabemos que existe $n \in \mathbb{N}$, tal que para toda $k \geq n$

$$\mu(B_{1/2^n}(x)) \leq 2g(1/2^n),$$

es decir, $x \in \liminf_n A_n$ concluyendo con esto que $A \subseteq \liminf_n A_n$.

Con lo anterior hemos probado que $A = \lim_{n \uparrow \infty} A_n$ y por lo tanto $\mu(A) = \lim_{n \uparrow \infty} \mu(A_n)$.

Fijamos ahora $n \geq 1$ y consideramos una 2^{-n} - cubierta $\{U_i\}_{i \in I}$ de A_n . Para cada $i \in I$, denotamos por $r_i > 0$ al diámetro de U_i y escogemos un punto $x_i \in U_i \cap A_n$. Sea $k_i \geq n$ el único entero tal que $1/2^{k_i+1} \leq r_i < 1/2^{k_i}$. Entonces, para toda $i \in I$, tenemos

$$U_i \cap A_n \subset \bar{B}_{r_i}(x_i) \subset B_{1/2^{k_i}}(x_i).$$

Utilizando la propiedad $g(2r) < cg(r)$ y recordando la definición de A_n se sigue que

$$\sum_{i \in I} g(r_i) \geq \frac{1}{c} \sum_{i \in I} g(1/2^{k_i}) \geq \frac{1}{2c} \sum_{i \in I} \mu \left(B_{1/2^{k_i}}(x_i) \right) \geq \frac{1}{2c} \mu(A_n),$$

donde la última desigualdad es por monotonía de la medida, ya que las bolas $B_{1/2^{k_i}}(x_i)$ son de igual manera 2^{-n} - cubierta de A_n . Por la definición de la medida de Hausdorff obtenemos que

$$g - m(A) \geq g - m(A_n) \geq \frac{1}{2c} \mu(A_n)$$

y el resultado deseado se sigue de hacer que $n \uparrow \infty$.

ii.) De la teoría general de medidas de Hausdorff (*Véase* [16] Corolario 2, pág.99), sabemos que

$$g - m(A) = \sup \{ g - m(K) : K \subset A, K \text{ compacto} \}.$$

Sin pérdida de generalidad, supongamos que A es compacto.

Sea $\epsilon > 0$. Por hipótesis, para toda $x \in A$, podemos encontrar $r_x \in (0, \epsilon/8)$ tal que

$$\mu(B_{r_x}(x)) \geq \frac{1}{2}g(r_x).$$

Por compacidad, podemos encontrar $x_1, \dots, x_n \in A$, tal que

$$A \subset \bigcup_{i=1}^n B_{r_{x_i}}(x_i)$$

y podemos asumir a la vez que $r_{x_1} \geq r_{x_2} \geq \dots \geq r_{x_n}$. Con esto, nos será posible construir un subconjunto finito $1 = m_1 < m_2 < \dots < m_\ell$ de $\{1, 2, \dots, n\}$ de manera que si $y_j = x_{m_j}$ tendremos que

$$A \subset \bigcup_{j=1}^{\ell} B_{4r_{y_j}}(y_j)$$

y las bolas $B_{r_{y_j}}(y_j)$ y $B_{r_{y_{j'}}}(y_{j'})$ son disjuntas si $j \neq j'$. De hecho, comenzamos con $m_1 = 1$ y procedemos por inducción. Supongamos que hemos construido $m_1 < m_2 < \dots < m_{p-1}$ de manera que

$$\bigcup_{i=1}^{m_{p-1}} B_{r_{x_i}}(x_i) \subset \bigcup_{j=1}^{p-1} B_{4r_{y_j}}(y_j),$$

y las bolas $B_{r_{y_j}}(y_j)$, $1 \leq j \leq p-1$ son disjuntas. Si

$$A \subset \bigcup_{i=1}^{p-1} B_{4r_{y_i}}(y_i)$$

entonces la construcción está completa. De otra manera, sea $k > m_{p-1}$ el primer entero tal que $B_{r_{x_k}}(x_k)$ no está contenido en la unión de las bolas $B_{4r_{y_j}}(y_j)$ para $j \leq p-1$, y ponemos $m_p = k$. Llanamente, $B_{r_{y_p}}(y_p) \cap B_{r_{y_q}}(y_q) = \emptyset$ si $1 \leq q \leq p-1$ y donde $y_p = x_{m_p} = x_k$, por que de otra manera esto contradice el hecho de que la bola $B_{r_{y_p}}(y_p) = B_{r_{x_k}}(x_k)$ contiene un punto que no pertenece a $B_{4r_{y_q}}(y_q)$. Esto completa la construcción por inducción.

Ahora la bolas $B_{4r_{y_j}}(y_j)$ proveen una cubierta de A por conjuntos de diámetro menor que ϵ , y utilizando la hipótesis obtenemos

$$\sum_{j=1}^{\ell} g(8r_{y_j}) \leq c^3 \sum_{j=1}^{\ell} g(r_{y_j}) \leq 2c^3 \sum_{j=1}^{\ell} \mu(B_{r_{y_j}}(y_j)) \leq 2c^3 \mu(A_\epsilon),$$

donde A_ϵ representa la unión de éstas ϵ -vecindades de A . Dejando que ϵ se vaya a 0 obtenemos el resultado deseado. \square

6.4. Prueba de Teoremas

En esta sección, probamos el Teorema 113 y Teorema 115. Denotamos nuevamente por $\eta(de)$ a la medida de Itô de excursiones positivas del movimiento Browniano normalizada de manera que $\eta(\sup e > \epsilon) = \epsilon^{-1}$ y por $\zeta = \zeta(e)$ la duración de la excursión e . Para todo $s, t \in [0, \zeta]$,

$$d_e(s, t) = e_s + e_t - 2m_e(s, t)$$

donde

$$m_e(s, t) = \inf_{r \in [s \wedge t, s \vee t]} e_r.$$

Recordemos la clase de equivalencia sobre $[0, \zeta]$ dada por: $s \sim t$ si $d_e(s, t) = 0$. Entonces el cociente $\mathcal{T}_e := [0, \zeta]/\sim$ equipado con la métrica d_e es un árbol aleatorio real, cuya raíz es por convención la clase de equivalencia del 0, y la distribución de \mathcal{T}_e bajo $\eta(de)$ es Θ .

Más aun, salvo un factor multiplicativo de 2 el cual será ignorado, la medida uniforme \mathbf{m} sobre \mathcal{T}_e no es otra cosa más que la imagen de la medida de Lebesgue sobre $[0, \zeta]$ bajo la función cociente natural de $[0, \zeta]$ sobre $[0, \zeta]/\sim$, de manera similar la medida de tiempo local ℓ^a es la imagen de la usual medida de tiempo local del movimiento Browniano en el nivel a . Entonces, en la prueba del Teorema 113 y el Teorema 115 trabajaremos con el árbol \mathcal{T}_e bajo $\eta(de)$.

Demostración del Teorema 113. Establecemos primero la existencia de dos constantes positivas c_1 y c_2 tal que, $\eta(de)$ c.t.p. para todo Boreliano A subconjunto de \mathcal{T}_e ,

$$c_1 \mathbf{m}(A) \leq h - m(A) \leq c_2 \mathbf{m}(A). \quad (6.3)$$

Cota Inferior. Haciendo un poco de abuso de notación, con frecuencia identificaremos a un elemento $s \in [0, \zeta]$ con su clase de equivalencia en $\mathcal{T}_e = [0, \zeta]/\sim$. Primero probamos que, $\eta(de)$ c.t.p., para \mathbf{m} -casi toda $s \in \mathcal{T}_e$, se tiene

$$\limsup_{\epsilon \rightarrow 0} \frac{\mathbf{m}(\{t \in \mathcal{T}_e : d_e(s, t) \leq \epsilon\})}{h(\epsilon)} \leq C_1 \quad (6.4)$$

para alguna constante finita C_1 .

Para la prueba de (6.4), necesitamos un lema de descomposición para la excursión Browniana. Asumamos que, sobre cierto espacio de probabilidad, tenemos dos procesos $(B_t, t \geq 0)$, $(B'_t, t \geq 0)$ y para toda $a \geq 0$ una medida de probabilidad Π_a tal que B y B' son bajo Π_a dos movimientos Brownianos independientes que comienzan en a . También establecemos

$$T = \inf \{t \geq 0 : B_t = 0\}, \quad T' = \inf \{t \geq 0 : B'_t = 0\}$$

y escribimos $C(\mathbb{R}_+, \mathbb{R})$ para el espacio de funciones continuas de \mathbb{R}_+ en \mathbb{R} .

Lema 117. *Para toda función medible no-negativa F definida sobre $C(\mathbb{R}_+, \mathbb{R})^2$,*

$$\int \eta(\mathrm{d}e) \int_0^\zeta \mathrm{d}s F((e_{s+t})_{t \geq 0}, (e_{s-t})_{t \geq 0}) = 2 \int_0^\infty \mathrm{d}a \Pi_a[F((B_{t \wedge T})_{t \geq 0}, (B'_{t \wedge T})_{t \geq 0})].$$

Esto es la descomposición de Bismut para la excursión Browniana. Para una prueba, consúltese [13], Lema 1. Notemos que por definición de la distancia d_e y la previa identificación de la medida \mathbf{m} como imagen de la medida de Lebesgue,

$$\mathbf{m}(\{t \in \mathcal{T}_e : d_e(s, t) \leq \epsilon\}) = \int_0^s \mathbf{1}_{\{e(s)+e(t)-2m_e(s,t) \leq \epsilon\}} \mathrm{d}t + \int_s^\zeta \mathbf{1}_{\{e(s)+e(t)-2m_e(s,t) \leq \epsilon\}} \mathrm{d}t.$$

Del Lema 117, observamos que la relación (6.4) se seguirá si es posible probar que para toda $a > 0$,

$$\limsup_{\epsilon \rightarrow 0} \frac{\int_0^\infty \mathrm{d}t \mathbf{1}_{\{a+B_t-2I_t \leq \epsilon\}} + \int_0^\infty \mathrm{d}t \mathbf{1}_{\{a+B'_t-2I'_t \leq \epsilon\}}}{h(\epsilon)} \leq C_1, \quad \Pi_a \text{ c.s.} \quad (6.5)$$

donde

$$I_t = \inf_{s \leq t} B_s, \quad I'_t = \inf_{s \leq t} B'_s.$$

Debido a que la medida de Lebesgue es invariante bajo traslaciones, es suficiente con considerar el caso en que $a = 0$ en (6.5). El Teorema de descomposición de Pitman establece que el proceso $R_t := B_t - 2I_t$ es, bajo Π_0 , un proceso de Bessel tridimensional comenzado en 0, esto es, tiene la misma distribución que el módulo de un movimiento Browniano tridimensional comenzado en el origen. De estimaciones debido a Ciesielski y Taylor [2], existe una constante finita C_2 tal que

$$\limsup_{\epsilon \rightarrow 0} \frac{\int_0^\infty \mathbf{1}_{\{R_t \leq \epsilon\}}}{h(\epsilon)} = C_2, \quad \Pi_0 \text{ casi seguramente.}$$

De esto y el procedimiento análogo para $R'_t := B'_t - 2I'_t$, se deduce (6.5), lo cual a su vez completa la prueba de (6.4). La cota inferior en (6.3) se obtiene entonces del Lema 116. i.)

Cota Superior. Del Lema 116. ii.), la cota superior en (6.3) es posible obtenerla probando la existencia de una constante $K_1 > 0$ tal que, $\eta(de)$ c.t.p.,

$$h - m \left(\left\{ s \in \mathcal{T}_e : \limsup_{\epsilon \rightarrow 0} \frac{\mathbf{m}(\{t \in \mathcal{T}_e : d_e(s, t) \leq \epsilon\})}{h(\epsilon)} \leq K_1 \right\} \right) = 0. \quad (6.6)$$

Para todo entero $n \geq 0$, establecemos $\epsilon_n = 2^{-n}$. Entonces, probaremos la existencia de una constante K_2 tal que, para todo entero $n_0 \geq 0$, $\eta(de)$ c.t.p.,

$$h - m \left(\left\{ s \in \mathcal{T}_e : \begin{array}{l} e(s) > 2^{-n_0} \text{ y} \\ \mathbf{m}(\{t \in \mathcal{T}_e : d_e(s, t) \leq \epsilon_p\}) \leq K_2 h(\epsilon_p), \forall p \geq n_0. \end{array} \right\} \right) = 0. \quad (6.7)$$

No es difícil notar que la relación (6.6) se sigue de (6.7). Para probar entonces (6.7), es necesario introducir cubiertas adecuadas para los conjuntos:

$$F_{n_0}^A := \left\{ s \in \mathcal{T}_e : \begin{array}{l} 2^{-n_0} < e(s) < A \text{ y} \\ \mathbf{m}(\{t \in \mathcal{T}_e : d_e(s, t) \leq \epsilon_p\}) \leq K_2 h(\epsilon_p), \forall p \geq n_0. \end{array} \right\},$$

donde A es un entero positivo. Para toda $n \geq 0$, consideremos la sucesión de tiempos de paro definida inductivamente por:

$$\begin{aligned} T_0^n &= 0, \\ T_1^n &= \inf \{s \geq 0 : e(s) = 2^{-n}\}, \\ T_{k+1}^n &= \inf \{s \geq T_k^n : |e(s) - e(T_k^n)| = 2^{-n}\}, \end{aligned}$$

donde por convención $\inf \emptyset = \infty$. La sucesión $(2^n e(T_k^n) \mathbf{1}_{\{T_k^n < \infty\}}, k \geq 0)$ se distribuye bajo $\eta(de | T_1^n < \infty)$ como una excursión positiva de una caminata aleatoria simple, ya que

$$\begin{aligned} &\eta \left(2^n |e(T_k^n) \mathbf{1}_{\{T_k^n < \infty\}} - e(T_{k-1}^n) \mathbf{1}_{\{T_{k-1}^n < \infty\}}| = 1 \mid T_1^n < +\infty \right) \\ &= \eta \left(2^n e(T_1^n) = 1 \mid T_1^n < +\infty \right) = \mathbb{P}_{1/2^n}(B_{T_0} = 1/2^n) = 1. \end{aligned}$$

En particular, para todo entero $i \geq 1$,

$$\eta \left(\sum_{k=0}^{\infty} \mathbf{1}_{\{T_k^n < \infty, e(T_k^n) = i2^{-n}\}} \right) = 2 \eta(T_1^n < \infty) = 2 \eta \left(\sup_{s \geq 0} e_s > \frac{1}{2^n} \right) = 2^{n+1}. \quad (6.8)$$

Sea $s \in F_{n_0}^A$ y $n \geq n_0 + 2$. Existe un único entero $k > 0$ tal que $s \in [T_k^n, T_{k+1}^n)$. De nuestras definiciones, tenemos que

$$d_e(T_k^n, s) \leq 3 \cdot 2^{-n}.$$

En consecuencia, para toda $p \in \{n_0, n_0 + 1, \dots, n - 2\}$, tenemos que

$$\{t \in \mathcal{T}_e : d_e(s, t) \leq \epsilon_p\} \supset \{t \in \mathcal{T}_e : d_e(T_k^n, t) \leq \epsilon_p/4\}.$$

Se sigue entonces que

$$F_{n_0}^A \subset \bigcup_{k \in I_{n_0, n}} [T_k^n, T_{k+1}^n) \quad (6.9)$$

donde,

$$I_{n_0, n} = \left\{ k \geq 0 : T_k^n < \infty, \quad 2^{-n_0} \leq e(T_k^n) \leq A \quad y \right.$$

$$\left. \int_{T_k^n}^{\zeta} dt \mathbf{1}_{\{d_e(T_k^n, t) \leq \epsilon_p/4\}} \leq K_2 h(\epsilon_p), \forall p \in \{n_0, \dots, n - 2\} \right\}.$$

Para acotar la cardinalidad $\#I_{n_0, n}$, utilizamos la propiedad fuerte de Markov bajo la medida de excursión para escribir, para todo $k \geq 1$,

$$\begin{aligned} \eta(k \in I_{n_0, n}) &= \eta(\mathbf{1}_{\{T_k^n < \infty, 2^{n_0} \leq e(T_k^n) \leq A\}} \\ &\times \Pi_{e(T_k^n)} \left[\int_0^T dt \mathbf{1}_{\{e(T_k^n) + B_t - 2I_t \leq \epsilon_p\}} \leq K_2 h(\epsilon_p), \forall p \in \{n_0, \dots, n - 2\} \right]). \end{aligned}$$

Usando nuevamente el resultado de Pitman citado anteriormente, tenemos que para todo $a \geq 2^{-n_0}$,

$$\begin{aligned} \Pi_a \left[\int_0^T dt \mathbf{1}_{\{a + B_t - 2I_t \leq \epsilon_p/4\}} \leq K_2 h(\epsilon_p), \forall p \in \{n_0, \dots, n - 2\} \right] \\ = \Pi_0 \left[\int_0^\infty dt \mathbf{1}_{\{R_t \leq \epsilon_p/4\}} \leq K_2 h(\epsilon_p), \forall p \in \{n_0, \dots, n - 2\} \right] \end{aligned}$$

donde R es bajo Π_0 un proceso de Bessel tridimensional comenzado en 0. Se sigue del Teorema (1,2) en [14] que, dado que n_0 sea suficientemente grande, es posible elegir K_2 suficientemente pequeño de manera que la última probabilidad que escribimos se encuentre acotada superiormente por

$$\exp(-c(n - n_0)^{1/2})$$

para alguna constante positiva c . Luego,

$$\eta(k \in I_{n_0, n}) \leq \exp(-c(n - n_0)^{1/2}) \eta(T_k^n < \infty, 2^{-n_0} \leq e(T_k^n) \leq A)$$

y sumando sobre k , y utilizando (6.8)

$$\eta(|I_{n_0, n}|) \leq \exp(-c(n - n_0)^{1/2}) A 2^{2n+1}.$$

En particular, por el Lema de Fatou,

$$\liminf_{n \rightarrow \infty} 2^{-2n} \exp(-c(n - n_0)^{1/2}) |I_{n_0, n}| < \infty \quad \eta(\text{de}) \text{ c.t.p.}$$

Ahora recordemos la relación (6.9) y nótese que el diámetro (respecto a la métrica d_e) de cada intervalo $[T_k^n, T_{k+1}^n]$ se encuentra acotado superiormente por $4 \cdot 2^{-n}$. Nuestra afirmación (6.7) se sigue entonces de la definición de una medida de Hausdorff. Esto completa la prueba de (6.6).

Ahora el Teorema 113 se puede deducir de las cotas (6.6) y una apropiada ley cero-uno.

Denotamos p_e a la función cociente natural de $[0, \zeta]$ sobre $\mathcal{T}_e = [0, \zeta]/\sim$. Observamos primero que, para toda $0 \leq s \leq t \leq \zeta$, la cantidad $h - m(p_e([s, t]))$ es una función medible respecto a e . Para verificar esto, consideremos la definición de medida de Hausdorff de $h - m(p_e([s, t]))$, y restringimos nuestra atención a las cubiertas finitas conformadas por bolas (esto es posible considerando la compacidad de los árboles y el hecho de que cualquier subconjunto de un árbol real está contenido en una bola cerrada del mismo diámetro). Más aún, basta con considerar las bolas de radio racional y centro de la forma $p_e(r)$ para algún racional $r \in [0, \zeta]$. Como las bolas abiertas de un espacio topológico son un conjunto que genera a los Borelianos del mismo y el espacio \mathbb{T} es separable, se sigue fácilmente la propiedad de medibilidad.

Definimos un medida finita ν sobre $[0, \zeta]$ estableciendo, para todo $t \in [0, \zeta]$.

$$\nu([0, t]) = h - m(p_e([0, t])).$$

Como el mapeo $t \mapsto h - m(p_e([0, t]))$ es continuo la medida ν es no-atómica. Entonces tendremos también, para todo $0 \leq s \leq t \leq \zeta$,

$$\nu([s, t]) = h - m(p_e([s, t])).$$

En efecto, esto es una consecuencia de las siguiente observación : Si $0 \leq u < v \leq s < t \leq \zeta$, el conjunto $p_e([s, t]) \cap p_e([u, v])$ está contenido en la línea ancestral de $p_e(s)$, por ende tendremos

$$h - m(p_e([s, t]) \cap p_e([u, v])) = 0.$$

Ya que \mathbf{m} se obtiene a partir de la imagen de la medida de Lebesgue bajo p_e , es fácil verificar que $\mathbf{m}(p_e([s, t])) = t - s$ para todo $0 \leq s \leq t \leq \zeta$, $\eta(de)$ c.t.p. De las cotas (6.6), obtenemos $\eta(de)$ c.t.p. para todo $0 \leq s \leq t \leq \zeta$,

$$c_1(t - s) \leq \nu([s, t]) \leq c_2(t - s).$$

De esto obtenemos que la medida ν es absolutamente continua respecto a la medida de Lebesgue sobre $[0, \zeta]$, como ambas son medidas σ -finitas sobre $[0, \zeta]$ tendremos que la derivada de Radon-Nykodin es igual, casi en todas partes a

$$\frac{d\nu}{dt} = \lim_{\epsilon \rightarrow 0} \frac{\nu([t - \epsilon, t + \epsilon])}{2\epsilon} = \lim_{\epsilon \rightarrow 0} \frac{h - m(p_e([t - \epsilon, t + \epsilon]))}{2\epsilon} \quad (6.10)$$

No es difícil verificar que la cantidad $h - m(p_e([t - \epsilon, t + \epsilon]))$ es una función medible respecto de la trayectoria $(e(t + u) - e(t), u \in [-\epsilon, \epsilon])$. Por lo tanto es posible utilizar el Lema 116 y la ley 0 - 1 usual para el movimiento Browniano para obtener que este último límite en (6.10) debe ser igual a una constante $C_0 \in [0, \infty]$, dt c.t.p., $\eta(de)$ c.t.p. Además tendremos que $C_0 \in [c_1, c_2]$ y particularmente $0 < C_0 < \infty$.

Tenemos entonces que $h - m(A) = C_0 \mathbf{m}(A)$ para todo subconjunto de el árbol de la forma $p_e([s, t])$, o cualquier unión finita de dichos conjuntos. De cualquier manera, todo abierto del árbol es un límite creciente de una sucesión de dichas uniones ($p_e^{-1}(U)$ es unión numerable de intervalos abiertos, ya que p_e es una función cociente). Entonces $h - m(A) = C_0 \mathbf{m}(A)$ para todo abierto A de \mathcal{T}_e , lo cual es suficiente para completar la prueba. \square

Demostración del Teorema 115. Ahora nos enfocamos en la medida de Hausssdorf de los conjuntos de nivel de \mathcal{T}_e . Recordemos que

$$\mathcal{T}_e(a) = \{s \in \mathcal{T}_e : d_e(0, s) = a\} = \{s \in \mathcal{T}_e : e(s) = a\}.$$

También denotamos por $(\ell_s^a, s \geq 0)$ al proceso del tiempo local (Browniano) de e en el nivel a . Entonces la medida $\ell^a(ds)$ asociada con la función creciente $s \mapsto \ell_s^a$ puede ser interpretada como una medida sobre $\mathcal{T}_e(a)$ y efectivamente coincide con aquella que fue discutida en la introducción del capítulo (salvo un factor multiplicativo de 2 el cual no es relevante para nuestros propósitos). Es más, para toda función no-negativa y medible $F \in C(\mathbb{R}_+, \mathbb{R})^2$, es posible deducir del Lema 117

$$\int \eta(de) \int_0^\zeta \ell^a(ds) F((e_{s+t})_{t \geq 0}, (e_{s-t^+})_{t \geq 0}) = 2\Pi_a[F((B_{t \wedge T})_{t \geq 0}, (B'_{t \wedge T})_{t \geq 0})]. \quad (6.11)$$

De manera análoga a la prueba del Teorema 113, establecemos primero la existencia de dos constantes positivas \tilde{c}_1, \tilde{c}_2 , de manera que, $\eta(de)$ c.t.p. para todo Boreliano A de $\mathcal{T}(a)$,

$$\tilde{c}_1 \ell^a(A) \leq \tilde{h} - m(A) \leq \tilde{c}_2 \ell^a(A). \quad (6.12)$$

Cota Inferior. Similarmente a como obtuvimos la cota inferior en la prueba anterior, basta mostrar la existencia de una constante C'_1 tal que, $\eta(de)$ c.t.p, para ℓ^a -casi toda $s \in \mathcal{T}_e(a)$,

$$\limsup_{\epsilon \rightarrow 0} \frac{\ell^a(\{t \in \mathcal{T}_e(a) : d_e(s, t) \leq \epsilon\})}{\tilde{h}(\epsilon)} \leq C'_1. \quad (6.13)$$

Si $0 < \epsilon < a$ y $s, t \in \mathcal{T}_e(a)$, tenemos que $d_e(s, t) \leq \epsilon$ si y sólo si $m_e(s, t) \leq \epsilon/2$. De esta observación y (6.11), podemos ver que (6.13) se sigue de verificar que, para toda $a > 0$, Π_a casi seguramente,

$$\limsup_{\epsilon \rightarrow 0} \frac{L_{T_{a-\epsilon}}^a(B) + L_{T_{a-\epsilon}}^a(B')}{\tilde{h}(\epsilon)} \leq C'_1, \quad (6.14)$$

donde $T_{a-\epsilon} = \inf \{t : B_t = a - \epsilon\}$ y $(L_t^a(B), t \geq 0)$ es el proceso de tiempo local de B en el nivel a .

Es sabido que la distribución de $L_{T_{a-\epsilon}}^a(B)$ bajo π_a es exponencial con esperanza 2ϵ . Luego, una aplicación el Lema de Borel-Cantelli nos muestra que, para $\epsilon_n = 2^{-n}$,

$$\limsup_{n \rightarrow \infty} \frac{L_{T_{a-\epsilon_n}}^a(B)}{\tilde{h}(\epsilon_n)} \leq 1.$$

Se seguirá que (6.14) se mantiene con $C'_1 = 8$.

Cota Superior. Bastará probar que existe una constante K'_1 , no dependiente de a , tal que, $\eta(de)$ c.t.p.,

$$\tilde{h} - m(\{s \in \mathcal{T}_e(a) : \limsup_{\epsilon \rightarrow 0} \frac{\ell^a(\{t \in \mathcal{T}_e : d_e(s, t) \leq \epsilon\})}{\tilde{h}(\epsilon)} \leq K'_1\}) = 0 \quad (6.15)$$

Esto requiere encontrar cubiertas adecuadas para los conjuntos

$$G_{n_0} = \{s \in \mathcal{T}_e(a) : \ell^a(\{t \in \mathcal{T}_e(a) : d_e(s, t) \leq \epsilon_p\}) \leq K'_1 \tilde{h}(\epsilon_p), \forall p \geq n_0\},$$

donde K' es una constante positiva. Fijamos $a > 0$ y $n_0 \geq 1$ tal que $2^{-n_0} < a$. Para cubrir a G_{n_0} , introducimos una sucesión de tiempos de paro definidos para todo $n \geq n_0$ como,

$$T_0^n = \inf \{s : e(s) = a\}, T_1^n = \inf \{s > T_0^n : |e(s) - a| = 2^{-n}\}$$

e inductivamente,

$$T_{2k}^n = \inf \{s > T_{2k-1}^n : e(s) = a\}, \quad T_{2k+1}^n = \inf \{s > T_{2k}^n : |e(s) - a| = 2^{-n}\}.$$

Es fácil verificar que

$$\eta \left(\sum_{k=0}^{\infty} \mathbf{1}_{\{T_{2k}^n < \infty\}} \right) \leq C_{(a)} 2^n \quad (6.16)$$

donde la constante $C_{(a)}$ únicamente depende de a .

De manera muy similar a la prueba del Teorema (113), tenemos

$$G_{n_0} \subset \bigcup_{k \in J_{n_0, n}} [T_{2k}^n, T_{2k+1}^n] \quad (6.17)$$

donde

$$J_{n_0, n} = \left\{ k : T_{2k}^n < \infty, \int_{T_{2k}^n}^{\zeta} d\ell_t^a \mathbf{1}_{\{d_e(T_{2k}^n, t) \leq \epsilon_p/4\} \leq K' \tilde{h}(\epsilon_p), \forall p \in \{n_0, \dots, n-2\}} \right\}.$$

Por la propiedad fuerte de Markov al tiempo T_{2k}^n ,

$$\eta(k \in J_{n_0, n}) = \eta \left(T_{2k}^n < \infty, \Pi_a [L_{T_{a-\epsilon_p/8}}^a \leq K' \tilde{h}(\epsilon_p), \forall p \in \{n_0, \dots, n-2\}] \right).$$

Notamos que las variables $L_{T_{a-\epsilon_p/8}}^a - L_{T_{a-\epsilon_p/16}}^a$, $p \in \{n_0, \dots, n-2\}$ son independientes bajo Π_a por la propiedad de Markov y ya que L es una funcional aditiva. Más aún, condicionado al evento de ser estrictamente positivo, lo cual tiene probabilidad de $1/2$, la variable $L_{T_{a-\epsilon_p/8}}^a - L_{T_{a-\epsilon_p/16}}^a$ se distribuye exponencialmente con esperanza $\epsilon_p/4$ (Un resultado debido a Ray-Knight. Véase [17]). Se sigue que

$$\begin{aligned} & \Pi_a [L_{T_{a-\epsilon_p/8}}^a \leq K' \tilde{h}(\epsilon_p), \forall p \in \{n_0, \dots, n-2\}] \\ & \leq \Pi_a [L_{T_{a-\epsilon_p/8}}^a - L_{T_{a-\epsilon_p/16}}^a \leq K' \tilde{h}(\epsilon_p), \forall p \in \{n_0, \dots, n-2\}] \\ & = \prod_{p=n_0}^{n-2} 1 - \frac{1}{2} \exp(-4K' \log \log 2^p). \end{aligned}$$

Si $K' < 1/8$, la última igualdad es acotada superiormente por $\exp(-n^{1/2})$ para n grande. Entonces para toda n suficientemente grande,

$$\eta(k \in J_{n_0, n}) \leq \exp(-n^{1/2}) \eta(T_{2k}^n < \infty).$$

Combinando esto último con (6.16) y haciendo uso del Lema de Fatou, llegamos a

$$\liminf_{n \rightarrow \infty} 2^{-n} \exp(n^{1/2}) \#J_{n_0, n} < \infty, \quad \eta(\text{de}) \text{ c.t.p.} \quad (6.18)$$

Como por construcción el diámetro de cada intervalo $[T_{2^k}^n, T_{2^{k+1}}^n]$ es acotado por $4 \cdot 2^{-n}$, (6.17) y (6.18) implican $\tilde{h} - m(G_{n_0}) = 0$. La estimación en (6.15) se obtiene de una adecuada elección de K'_1 , y esto completa la prueba de las cotas (6.12).

El final de la prueba es ahora similar al del teorema anterior. Introducimos la medida aleatoria $\tilde{\nu}$ sobre $[0, \zeta]$ definida por

$$\tilde{\nu} = \tilde{h} - m(p_e([0, t]) \cap \mathcal{T}(a)).$$

Las cotas en (6.12) implican que $\tilde{\nu}(dt)$ es absolutamente continua respecto a $\ell^a(dt)$, y que su densidad es acotada inferior y superiormente por \tilde{c}_1 y \tilde{c}_2 respectivamente. Un argumento por una ley cero-uno, aunado a (6.11), muestra que esta densidad es igual a una constante \tilde{C}_0 , $\ell^a(dt)$ c.t.p., $\eta(\text{de})$ c.t.p. Más aún, esta constante no depende de a . \square

Bibliografía

- [1] Billingsley Patrick, “*Convergence of Probability Measures*”, John Wiley & Sons, Inc., 1968.
- [2] Cielsielski, Taylor, S.J., (1962) “*First Passage Times and Soujourn Times for Brownian motion in spaces and the exact Hausdorff measure of the sample path.*”, Trans. Amer. Math. Soc. **103**, pgs. 369-383.
- [3] Dugundji J., “*Topology*”, Allyn and Bacon Series in Advanced Mathematics, 1978.
- [4] Duquesne Thomas, Le Gall J. F., (2005) “*Probabilistic and Fractal Aspects of Levy Trees.*”. Probability Th. Rel. Fields **131** pgs. 553-603.
- [5] Durrett Rick, “*Probability: Theory and Example*”, Cambridge Series in Statistical and Probabilistic Mathematics, 1978.
- [6] Engelking R., “*General Topology*”, Sigma Series in Pure Mathematics; Vol. 6 Heldermann Verlag Berlin , 1989.
- [7] Ethier S.N. and T. Kurtz, “*Markov Processes, Characterization and Convergence*”, John Wiley & Sons Inc., 2005.
- [8] Evans Steven N., Pitman Jim, Winter Anita, (2005) “*Rayleigh processes, real trees, and root growth with re-Grafting*”, Probability Th. Rel. Fields
- [9] Geluk J.L. and Hann, “*Stable probability distributions and their domains of attractions: A direct approach. En: Probability and Mathematical Statistics*”, pg. 169-188, 2000.
- [10] Kaigh W.D., “*An invariance principle for random walk conditioned by a late return to zero*”, Ann. Probab. 4 (1976), 115-121.
- [11] Le Gall J. F., (2005) “*Random Trees and Applications.*”, Probability Surveys, 2005.

- [12] Le Gall J. F., (1999) “*Spatial Branching Processes Random Snakes and Partial Differential Equations.*”, Lectures in Mathematics ETH Zürich. Birkhäuser, Boston. pg. 1-13.
- [13] Le Gall J. F., (1993) “*The Uniform Random Tree in a Brownian Excursion.*”, Probability Th. Rel. Fields **96** pgs. 369-383.
- [14] Le Gall, J.F., (1985) “*Sur la mesure de Hausdorff de la courbe brownienne. En: Seminaire de Probabilities XIX. Lecture Notes Math.*”, pp. 297-313.
- [15] Marckert J.F., “*Limit of Random Walks and Random Trees*”, Graz, 2009.
- [16] Rogers C.A., “*Hausdorff Measures*”, Cambridge University Press, 1970.
- [17] Revuz Daniel, Yor Marc , “*Continuous Martingales and Brownian Motion*”, Springer, 2005.
- [18] Spitzer F., “*Principles of Random Walk*”, Van Nostrand, 1963.