

**Benemérita Universidad Autónoma de
Puebla**

Facultad de Ciencias Físico Matemáticas

Posgrado en Ciencias Matemáticas

**Estudio Probabilístico de Índices
Topológicos para grafos aleatorios.**

Tesis

Presentada para obtener el grado de
Maestro en Ciencias Matemáticas

Presenta

Saylé Caridad Sigarreta Ricardo

Director de Tesis
Hugo Adán Cruz Suárez

Puebla, Puebla. 2022



BUAP

DR. SEVERINO MUÑOZ AGUIRRE
SECRETARIO DE INVESTIGACIÓN Y
ESTUDIOS DE POSGRADO, FCFM-BUAP
P R E S E N T E:

Por este medio le informo que la C:

SAYLÉ CARIDAD SIGARRETA RICARDO

estudiante de la Maestría en Ciencias (Matemáticas), ha cumplido con las indicaciones que el Jurado le señaló en el Coloquio que se realizó el día 14 de noviembre de 2022, con la tesis titulada:

Estudio probabilístico de índices Topológicos para grafos aleatorios

Por lo que se le autoriza a proceder con los trámites y realizar el examen de grado en la fecha que se le asigne.

A T E N T A M E N T E.
H. Puebla de Z. a 22 de noviembre de 2022

DRA. PATRICIA DOMÍNGUEZ SOTO
COORDINADORA DEL POSGRADO
EN MATEMÁTICAS.



Facultad
de Ciencias
Físico Matemáticas

Av. San Claudio y 18 Sur, edif. FM1
Ciudad Universitaria, Col. San
Manuel, Puebla, Pue. C.P. 72570
01 (222) 229 55 00 Ext. 7550 y 7552

A mi abuela ...

Agradecimientos

Quiero expresar mi gratitud a Dios, quien con su bendición llena siempre mi vida.

Mi profundo agradecimiento a mis padres y hermana por su apoyo incondicional en mi formación académica.

De igual manera mis agradecimientos a mis sinodales, Dra. Hortensia Reyes Cervantes, Dr. Carlos Guillén Galván y Dr. Rei Israel Ortega Gutiérrez por sus atentas lecturas y sugerencias, que sin duda mejoraron en gran parte este trabajo.

Quiero expresar mi agradecimiento al Dr. Hugo Adán Cruz Suárez, principal colaborador durante todo este proceso, quien con su dirección, conocimiento y enseñanza permitió el desarrollo de este trabajo.

Finalmente, al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico otorgado para la realización de esta maestría.

Introducción

Iniciada en 1736 por Euler y desarrollada en el siglo XIX por los ingleses A. Cayley y J.J. Silvester [51], la Teoría de Grafos se ha convertido en una herramienta práctica y teórica muy potente. Un grafo G está determinado por dos conjuntos (V, E) , el conjunto de nodos (V) y aristas (E). Las aristas y los nodos se interpretan en función del problema a modelar. Sin duda alguna, la posibilidad de representar a los grafos mediante diagramas permite que sean utilizados como modelos estructurales en la ciencia. En particular, es muy frecuente su utilidad en Química [6, 38], por ejemplo, para contar isómeros [56]. De hecho, el término grafo fue sugerido por Silvester para referirse a la fórmula estructural de un compuesto. Obteniendo así los grafos moleculares, donde los vértices corresponden a los átomos y las aristas a los enlaces de una molécula (véase Figura 1).

Una cantidad numérica TI asociada a un grafo G que satisface la igualdad $TI(G) = TI(G')$ para todo grafo G' isomorfo a G se llama invariante del grafo. En la teoría química de grafos, los

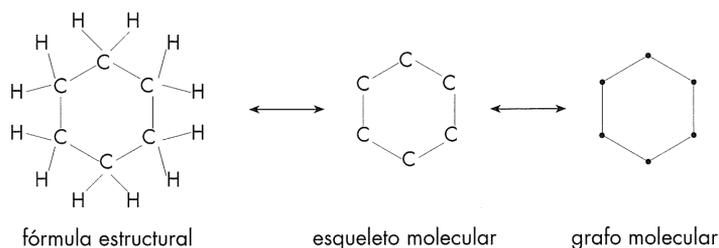


Figura 1: Esqueleto del grafo molecular ciclohexano.

invariantes de grafos que se aplican en investigaciones químicas se conocen como índices topológicos. La información estructural de un grafo puede representarse de diferentes maneras: matrices, polinomios, índices topológicos, etc. Los índices topológicos cuantifican la información estructural contenida en el grafo y son independientes de la numeración de los nodos y aristas, de ahí el adjetivo topológico. El objetivo de definir un índice topológico es asociar cada estructura química con un valor numérico y así investigar sus propiedades. El interés teórico y práctico de los índices topológicos ha experimentado un crecimiento explosivo desde su introducción, dando lugar a innumerables trabajos publicados que logran posicionarlos como una herramienta útil en múltiples problemas prácticos. De hecho, los índices topológicos han encontrado aplicaciones en Química [43, 65], Lingüística Computacional [50], Ecología [55], entre otras. A modo de resumen, las primeras investigaciones en esta área aparecieron en [74], dando lugar al ya conocido índice de Wiener para analizar y correlacionar las propiedades fisicoquímicas de los alcanos. En 1971, Haruo Hosoya continuó la investigación sobre los índices topológicos introduciendo el índice Z de Hosoya [33]. Por otra parte, el índice de Zagreb apareció por primera vez en [30]. Posteriormente se define en [57] el índice de Randić, considerado posiblemente, el índice topológico más estudiado y aplicado en la actualidad. Dando paso a generalizaciones como los "índices de conectividad molecular", introducidos en [39].

Hoy en día, existen un gran número de índices topológicos en la literatura [25]. De hecho, recientemente, en [26], se introdujo un nuevo índice topológico mediante un enfoque geométrico, denominado índice de Sombor, definido como

$$SO(G) = \sum_{uv \in E(G)} \sqrt{(d_u)^2 + (d_v)^2},$$

donde d_v es el grado de un vértice v . Actualmente, se han presentado varios invariantes de grafos relacionados con el índice de Sombor. Por ejemplo, en [40], Kulli introdujo el índice de Nirmala de un grafo G de la siguiente manera

$$N(G) = \sum_{uv \in E(G)} \sqrt{d_u + d_v}. \quad (1)$$

Trabajos recientes sobre el índice de Nirmala pueden consultarse en [28,29,41]. En particular, los índices topológicos basados en grados se han convertido en una poderosa herramienta, por ejemplo, el índice olvidado mide las propiedades físico-químicas de las estructuras moleculares [20,21], el índice GA puede utilizarse como herramienta de predicción en las investigaciones QSPR/QSAR [70] y el índice de conectividad de los enlaces atómicos ha demostrado ser un valioso índice de predicción en el estudio del calor de formación en los alcanos [15]. Hoy en día, el estudio de índices topológicos para grafos deterministas y aleatorios representa una parte importante de la investigación actual en química matemática y teoría química de grafos.

Finalmente, en el desarrollo de aplicaciones, se ha llegado a la conclusión de que los grafos aleatorios son una herramienta apropiada y útil para analizar fenómenos que evolucionan en el tiempo, ya que muchas características importantes son difíciles de capturar utilizando modelos deterministas. En este sentido, es importante mencionar que algunos trabajos realizan estudios de índices topológicos sobre redes y estructuras aleatorias. Para un mejor tratamiento remitimos a los lectores interesados a [1, 45, 49, 54, 80].

Basados en lo anterior la estructura de la tesis consta de tres capítulos. En el Capítulo 1 abordaremos el estudio de algunos índices topológicos para un modelo de árboles araña aleatorio. En el Capítulo 2 los índices topológicos basados en grados serán estudiados para las cadenas aleatorias de espiros concluyendo en el Capítulo 3 con el mismo estudio para las cadenas aleatorias de poliominós.

Índice general

Agradecimientos	4
Introducción	5
1. Árboles araña aleatorios	10
1.1. Preliminares	10
1.2. Árboles araña aleatorios	12
1.2.1. Hojas	13
1.2.2. Una clase de RST que evoluciona de manera preferencial	14
1.3. Índices Topológicos sobre árboles araña aleatorios .	16
1.3.1. Índice de Zagreb generalizado	17
1.3.1.1. Índice de Zagreb	23
1.3.1.2. Índice de Gordon-Scantlebury . .	26
1.3.1.3. Índice de Platt	29
1.3.1.4. Índice olvidado	30
1.3.2. Índice de Gini basado en grados	31
1.3.3. Índice Hoover basado en grados	35
2. Cadenas aleatorias de espiros	38
2.1. Preliminares	38
2.2. Índices topológicos sobre cadenas aleatorias de espiros	40
2.3. Interpretación de los resultados y ejemplos	47
3. Cadenas aleatorias de poliominós	54
3.1. Preliminares	54
3.2. Índices topológicos sobre cadenas aleatorias de poliominós	56
3.3. Cadena de poliominós	62

ÍNDICE GENERAL

4. Conclusiones	69
Bibliografía	70

Capítulo 1

Árboles araña aleatorios

1.1. Preliminares

Una clase de árboles relevante para los estudios químicos son los árboles con un número determinado de hojas. Un nodo se llama hoja si tiene grado 1. En [14] los autores demostraron que los árboles con n hojas ($n \geq 3$) que maximizan el primer modificado índice de conexión de Zagreb deben ser árboles araña (STs) o estrellas dobles. Por otro lado, en [66] se utilizaron STs para estudiar sistemas hexagonales que modelan moléculas de bencenoides y moléculas de bencenoides catacondensadas no ramificadas.

En particular, motivados por el aumento sustancial del interés en los modelos de árboles aleatorios Ren, Zhang y Dey investigaron dos clases de árboles aleatorios langosta que evolucionan según diferentes reglas y una clase de árboles araña aleatorios ($RSTs$) que crecen de forma preferencial [61]. En este importante trabajo, los autores obtuvieron resultados muy útiles, de hecho, para la clase de árboles araña aleatorios caracterizaron la estructura del modelo determinando la distribución exacta y asintótica del número de hojas. Inspirados en [61] y considerando los argumentos expuestos en los párrafos anteriores, en este capítulo, consideramos una clase de árboles araña que se incorporan con aleatoriedad, llamados árboles araña aleatorios ($RSTs$) e investigamos varios índices topológicos útiles de esta clase aleatoria, incluyendo el índice de Gini basado en grados, el índice de Hoover basado en grados, el

índice de Zagreb generalizado y otros índices asociados a estos. En concreto, se desarrolla un teorema del límite central para la distribución asintótica del número de hojas de un *RST*. Finalmente, es importante mencionar que lo desarrollado en esta sección se encuentra basado en el artículo [69] por aparecer en la revista Main Group Metal Chemistry.

Notación:

- \mathbb{R} denota el conjunto de los números reales.
- El valor esperado y la varianza de una variable aleatoria X definida en $(\Omega, \mathcal{F}, \mathbb{P})$ se denotan como $\mathbb{E}(X)$ y $\mathbb{V}(X)$, respectivamente.
- $X \sim F$ significa que la variable aleatoria X tiene una función de distribución F .
- $M_X(\cdot)$ denota la función generadora de momentos de la variable aleatoria X .
- $Bin(n-1, p)$ representa una variable aleatoria con distribución binomial con parámetros $n-1 \geq 0$ y $p \in [0, 1]$
- $Ber(p)$ representa una variable aleatoria con distribución Bernoulli con el parámetro $p \in [0, 1]$.
- $N(\mu, \sigma^2)$ representa una variable aleatoria con distribución normal donde $\mu \in \mathbb{R}$ es la media y $\sigma^2 > 0$ es la varianza de la variable aleatoria.
- $\chi^2(\lambda, k)$ representa una variable aleatoria con distribución chi-cuadrada no central donde $k > 0$ son los grados de libertad y $\lambda > 0$ es el parámetro de no centralidad.
- \xrightarrow{P} denota la convergencia en probabilidad y \xrightarrow{D} denota la convergencia en distribución.
- Si $r > 0$, $\xrightarrow{L_r}$ denota la convergencia en r -media.
- Dadas dos funciones de valor real $f(x)$ y $g(x) \neq 0$, llamamos $f(x) = o(g(x))$ si tenemos $\frac{f(x)}{g(x)} \rightarrow 0$ cuando $x \rightarrow \infty$.

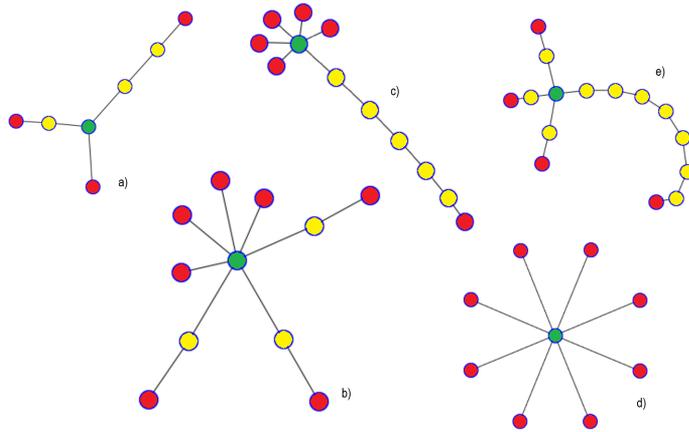


Figura 1.1: Ejemplos de STs .

- Dadas dos funciones de valor real $f(x)$ y $g(x)$, llamamos a $f(x) = O(g(x))$ si existe un número real positivo L y un número real x_0 tal que $|f(x)| \leq Lg(x)$ para todo $x \geq x_0$.

1.2. Árboles araña aleatorios

Un árbol araña es un árbol conectado con un centroide de grado al menos 3. Todos los nodos restantes se clasifican en dos categorías: nodos internos de grado 2 y hojas de grado 1. Por lo tanto, excepto el centroide, todos los nodos de un árbol araña tienen grados de a lo más 2. En Figura 1.1 se presentan ejemplos de STs . El centro estará representado por el color verde, los nodos internos por el color amarillo y las hojas por el color rojo.

Entrando en materia, la clase de RST considerada en este trabajo evoluciona de la siguiente manera: en el momento 1, se comienza con un grafo semilla que contiene un centroide y tres hojas. En cada etapa subsiguiente, las hojas y el centroide podrán reclutar nuevos nodos (en el tiempo $n \geq 2$):

1. El centroide será seleccionado con probabilidad p , $0 \leq p \leq 1$.
2. Se seleccionará una hoja con probabilidad $\frac{1-p}{L_n}$ donde L_n denota el número de hojas en el RST en el momento n .

Observe que las siguientes identidades son válidas:

$$p + \sum_{i=1}^{L_n} \frac{1-p}{L_n} = p + \frac{(1-p)L_n}{L_n} = 1.$$

Nótese que sólo el centroide y las hojas se califican para reclutar nuevos nodos. Si se selecciona el centroide, se le adjunta una nueva hoja; si se selecciona una hoja, se adjunta una nueva hoja a la hoja seleccionada y el reclutador se convierte en un nodo interno. Finalmente, tenemos que en cada etapa el grafo generado es un árbol araña con $n + 3$ nodos.

1.2.1. Hojas

En lo que sigue, L_n denota el número de hojas en un RST en el momento n , con $n \geq 1$.

Proposición 1.2.1. *Para $n \geq 1$ y $0 \leq p \leq 1$, se cumplen las siguientes afirmaciones:*

- 1) $L_n = 3 + \text{Bin}(n-1, p)$.
- 2) $\mathbb{E}(L_n) = 3 + (n-1)p$ y $\mathbb{V}(L_n) = (n-1)(1-p)p$.
- 3) $M_{L_n}(t) = (1-p + pe^t)^{n-1} e^{3t}$, $t \in \mathbb{R}$.

Demostración. Por la construcción del modelo se verifica que

$$\mathbb{P}(L_n = L_{n-1} + 1 \mid L_{n-1}) = p.$$

Así, dado $t \in \mathbb{R}$ se obtiene la siguiente relación de recurrencia

$$\begin{aligned} M_{L_n}(t) &= \sum_{i=3}^{n+2} e^{ti} \mathbb{P}(L_n = i) \\ &= (1-p) \sum_{i=3}^{n+1} e^{ti} \mathbb{P}(L_{n-1} = i) + p \sum_{i-1=3}^{n+1} e^{ti} \mathbb{P}(L_{n-1} = i-1) \\ &= (1-p + pe^t) M_{L_{n-1}}(t). \end{aligned}$$

Resolviendo la fórmula recursiva con $L_1 = 3$ obtenemos que

$$M_{L_n}(t) = (1 - p + pe^t)^{n-1} e^{3t}, t \in \mathbb{R}.$$

Indicándonos que $L_n = 3 + \text{Bin}(n - 1, p)$. En consecuencia, las afirmaciones 1) y 2) se verifican. \square

Teorema 1.2.2. *Para L_n cuando $n \rightarrow \infty$, tenemos que*

$$\frac{L_n - 3 - (n - 1)p}{\sqrt{p(1 - p)(n + k)}} \xrightarrow{D} N(0, 1),$$

para cada $k \in \mathbb{R}$.

Demostración. Por la Proposición 1.2.1 1), L_n puede expresarse como $L_n = 3 + \sum_{i=1}^{n-1} \text{Ber}(p)$ así $L_n - 3 = \sum_{i=1}^{n-1} \text{Ber}(p)$, aplicando el Teorema Central del Límite [24] obtenemos $\frac{L_n - 3 - (n - 1)p}{\sqrt{(1 - p)p(n - 1)}} \xrightarrow{D} N(0, 1)$. Ahora note que para cada $k \in \mathbb{R}$,

$$\frac{L_n - 3 - (n - 1)p}{\sqrt{(1 - p)p(n + k)}} = \frac{L_n - 3 - (n - 1)p}{\sqrt{(1 - p)p(n - 1)}} \frac{\sqrt{(1 - p)p(n - 1)}}{\sqrt{(n + k)(1 - p)p}} \quad (1.1)$$

y

$$\frac{\sqrt{(1 - p)p(n - 1)}}{\sqrt{(1 - p)p(n + k)}} \rightarrow 1, \text{ cuando } n \rightarrow \infty. \quad (1.2)$$

Finalmente por (1.1) y (1.2) la demostración es concluida. \square

1.2.2. Una clase de RST que evoluciona de manera preferencial

En un artículo muy reciente [61], los autores inspirados por [7] introdujeron una clase de RST que evoluciona de manera preferencial como sigue: en el momento 1, se inicia con un grafo semilla que contiene un centroide de grado 3 y tres hojas. En cada punto posterior, la probabilidad de que un nodo calificado reclute a un recién llegado es proporcional a su grado. Si se selecciona el centroide, se le une una nueva hoja; si se selecciona una hoja, se

adjunta una nueva hoja a la hoja seleccionada y el reclutador se convierte en un nodo interno. En consecuencia, para $n \geq 2$

$$\mathbb{P}(\mathbf{1}_{v,n}) = \frac{\deg_{v,n-1}}{\sum_{u \in Q_{n-1}} \deg_{u,n-1}},$$

donde v es un nodo calificado en el tiempo n , $\mathbf{1}_{v,n}$ indica el evento de que el nodo v es elegido como reclutador en el momento n , $\deg_{i,n-1}$ es el grado de un nodo i en el tiempo $n-1$ y Q_{n-1} denota el conjunto de nodos calificados en el tiempo $n-1$. Así, para $n \geq 2$, se deduce que:

1. La probabilidad de que el centroide reclute a un recién llegado en un momento n es $\frac{L_{n-1}}{2L_{n-1}} = \frac{1}{2}$.
2. La probabilidad de que una hoja reclute a un recién llegado en un momento n es $\frac{1}{2L_{n-1}}$.

Por lo tanto, podemos concluir que la clase de *RSTs* que evoluciona de manera preferencial (modelo preferencial) es el modelo presentado en la sección anterior para $p = 1/2$. Así, tomando $p = 1/2$ en los resultados previos, obtenemos como caso particular la Proposición 3 y el Teorema 1 de [61]. La siguiente proposición enlista dichos resultados.

Proposición 1.2.3. *Dado $n \geq 1$, las siguientes propiedades se verifican en el modelo preferencial*

- 1) $L_n = 3 + \text{Bin}(n-1, \frac{1}{2})$,
- 2) $\mathbb{E}(L_n) = \frac{n+5}{2}$ y $\mathbb{V}(L_n) = \frac{n-1}{4}$,
- 3) $M_{L_n}(t) = (\frac{1+e^t}{2})^{n-1} e^{3t}$, $t \in \mathbb{R}$.
- 4) Para toda $k \in \mathbb{R}$, $\frac{L_n - \frac{n+5}{2}}{\frac{\sqrt{n+k}}{2}} \xrightarrow{D} N(0,1)$, cuando $n \rightarrow \infty$.

1.3. Índices Topológicos sobre árboles araña aleatorios

Una cantidad numérica TI asociada a un grafo G que satisface la ecuación $TI(G) = TI(G^*)$ para todo grafo G^* isomorfo a G se denomina invariante del grafo. En la teoría química de grafos, los invariantes de grafos que se aplican en las investigaciones químicas se conocen como índices topológicos. El objetivo de definir un índice topológico es asociar cada estructura química con un valor numérico y así investigar sus propiedades, como la regularidad y la ramificación. Sea $G = (V, E)$, muchos índices topológicos importantes ($TI(G)$) se pueden definir como

$$TI(G) = \sum_{v \in V} h(deg_v), \quad (1.3)$$

donde $h : \{1, 2, \dots\} \rightarrow (0, \infty)$ y deg_v es el grado de un nodo $v \in V$. En el presente capítulo, estudiaremos algunos índices que satisfacen (1.3) en el modelo introducido en el presente capítulo. Ahora, nota que en cada etapa el árbol generado tiene tres tipos de nodos, centroide, hoja e interno, con grados $L_n, 1$ y 2 , respectivamente.

Proposición 1.3.1. *Sea TI_n el valor del índice topológico en la etapa n . Para cada $n \geq 1$, obtenemos*

$$\mathbb{E}(TI_n) = \mathbb{E}(h(L_n)^\alpha) + (h(1)^\alpha - h(2)^\alpha)\mathbb{E}(L_n) + h(2)^\alpha(n + 2)$$

y

$$\mathbb{V}(TI_n) = \mathbb{V}(h(L_n)^\alpha + (h(1)^\alpha - h(2)^\alpha)L_n).$$

Demostración. Observe que $I_n + L_n + 1 = n + 3$, donde I_n es el número de nodos internos del árbol en la etapa n , se deduce que:

$$\begin{aligned} TI_n &= h(L_n)^\alpha + h(1)^\alpha L_n + h(2)^\alpha(n + 2 - L_n) \\ &= h(L_n)^\alpha + (h(1)^\alpha - h(2)^\alpha)L_n + h(2)^\alpha(n + 2). \end{aligned} \quad (1.4)$$

Aplicando (1.4), obtenemos inmediatamente la media y la varianza de TI_n . \square

Como consecuencia de la Proposición 1.2.1 2) se sigue que.

Proposición 1.3.2. *Si $h(L_n) = aL_n + b$ con $a, b \in \mathbb{R}$ entonces $M_{h(L_n)}(t) = (1 - p + pe^{at})^{n-1}e^{(3a+b)t}$, $t \in \mathbb{R}$ y $n \geq 1$.*

1.3.1. Índice de Zagreb generalizado

El índice de Zagreb fue introducido por los químicos Gutman y Trinajstić [30]. Posteriormente, se señalaron algunas de sus propiedades matemáticas generales y se mostró su relación con otras magnitudes de interés en la teoría química de grafos [27]. De hecho, el índice de Zagreb y sus variantes se han utilizado en los estudios de las relaciones cuantitativas estructura-propiedad/actividad (QSPR/QSAR) [11, 37, 62]. Hoy en día, como indicador de su importancia, las ideas esbozadas en el documento inicial son exploradas por otros numerosos estudiosos [5, 19, 52].

Al tiempo $n \geq 1$, tomando $h(x) = x$ y $\alpha \in \mathbb{R}$ en (1.3) se obtiene el índice de Zagreb generalizado (Z_n^g). De acuerdo a (1.4),

$$Z_n^g = L_n^\alpha + (1 - 2^\alpha)L_n + 2^\alpha(n + 2). \quad (1.5)$$

Antes de continuar notemos que dado $n \in \mathbb{N}$ fijo, Z_n^g queda en términos de L_n , es decir, en el cálculo del índice de Zagreb generalizado no es determinante el número de nodos internos asociado a cada hoja. Indicándonos que para los índices topológicos que verifican la ecuación (1.3) lo importante del modelo es si en cada etapa se elige o no al centro (no le es de interés cuál hoja fue seleccionada), ya que, dos *RSTs* con el mismo número de hojas hasta la etapa n (por ende el mismo número de nodos internos) coincidirán en el valor del *TI* sin importar el número de nodos internos asociado a cada hoja.

Proposición 1.3.3. *Sea $\alpha \in \{1, 2, \dots\}$ y $t \in \mathbb{R}$, se tiene*

$$\frac{d^\alpha M_{L_n}(t)}{dt} = \sum_{i=1}^{\alpha} C_{\alpha,i} p^i e^{it} \frac{d^i M_{L_n}(u(t))}{du},$$

donde $u(t) = 1 - p + pe^t$, $C_{\alpha,\alpha} = C_{\alpha,1} = 1$ y $C_{\alpha,i} = C_{\alpha-1,i-1} + iC_{\alpha-1,i}$ para $i \in \{2, 3, \dots, \alpha - 1\}$.

Demostración. Abordaremos la demostración mediante inducción matemática sobre α . Sustituyendo $1 - p + pe^t$ por $u(t)$ en la Proposición 1.2.1 3), obtenemos $M_{L_n}(t) = M_{L_n}(u(t)) = u(t)^{n-1} \left(\frac{u(t)+p-1}{p}\right)^3$. Así, para la base, que es $\alpha = 3$, tenemos

$$\frac{d^3 M_{L_n}}{dt}(t) = p^3 e^{3t} \frac{d^3 M_{L_n}}{du}(u(t)) + 3p^2 e^{2t} \frac{d^2 M_{L_n}}{du}(u(t)) + pe^t \frac{dM_{L_n}}{du}(u(t)),$$

con $C_{3,1} = C_{3,3} = 1$ y $C_{3,2} = C_{2,1} + 2C_{2,2} = 3$. Suponemos que la afirmación es válida para todo α , es decir

$$\frac{d^\alpha M_{L_n}}{dt}(t) = \sum_{i=1}^{\alpha} C_{\alpha,i} p^i e^{it} \frac{d^i M_{L_n}}{du}(u(t)), \quad (1.6)$$

con $C_{\alpha,\alpha} = C_{\alpha,1} = 1$ y $C_{\alpha,i} = C_{\alpha-1,i-1} + iC_{\alpha-1,i}$ para $i \in \{2, 3, \dots, \alpha - 1\}$. Así, para cada $i \in \{1, 2, \dots, \alpha\}$ obtenemos que

$$\begin{aligned} & \frac{d}{dt} \left(C_{\alpha,i} p^i e^{it} \frac{d^i M_{L_n}}{du}(u(t)) \right) \\ &= i C_{\alpha,i} p^i e^{it} \frac{d^i M_{L_n}}{du}(u(t)) + C_{\alpha,i} p^{i+1} e^{(i+1)t} \frac{d^{i+1} M_{L_n}}{du}(u(t)). \end{aligned} \quad (1.7)$$

Por (1.6) y (1.7), hemos demostrado el siguiente resultado

$$\frac{d^{\alpha+1} M_{L_n}}{dt}(t) = \sum_{i=1}^{\alpha+1} C_{\alpha+1,i} p^i e^{it} \frac{d^i M_{L_n}}{du}(u(t)),$$

con $C_{\alpha+1,\alpha+1} = C_{\alpha+1,1} = 1$ y $C_{\alpha+1,i} = C_{\alpha,i-1} + iC_{\alpha,i}$ para $i \in \{2, 3, \dots, \alpha\}$, lo que completa la prueba. \square

Un caso especial de la Proposición 1.3.3 es el siguiente resultado, que es válido cuando $t = 0$ en (1.6).

Corolario 1.3.4. *Para $\alpha \in \{1, 2, \dots\}$, se verifica que*

$$\frac{d^\alpha M_{L_n}}{dt}(0) = \sum_{i=1}^{\alpha} C_{\alpha,i} p^i \frac{d^i M_{L_n}}{du}(1),$$

con $C_{\alpha,\alpha} = C_{\alpha,1} = 1$ y $C_{\alpha,i} = C_{\alpha-1,i-1} + iC_{\alpha-1,i}$ para $i \in \{2, 3, \dots, \alpha - 1\}$.

Teorema 1.3.5. Para $n \geq 1$, $p \in [0, 1]$ y $\alpha \in \{1, 2, \dots\}$ la siguiente identidad se verifica

$$\frac{d^\alpha M_{L_n}(0)}{dt} = p^\alpha n^\alpha + \frac{\alpha}{2}(\alpha(1-p) - p + 5)p^{\alpha-1}n^{\alpha-1} + O(n^{\alpha-2}).$$

Demostración. Primero, observe que

$$M_{L_n}(u(t)) = \frac{1}{p^3}(u(t)^{n+2} + 3(p-1)u(t)^{n+1} + 3(p-1)^2u(t)^n + (p-1)^3u(t)^{n-1}),$$

$t \in \mathbb{R}$. Luego para cada $i \in \{1, 2, \dots\}$, obtenemos

$$\begin{aligned} \frac{d^i M_{L_n}(u(0))}{du} &= \\ \sum_{k=-1}^2 \frac{b_k}{p^3} (n+k)(n+k-1) \dots (n+k-(i-1)) u(0)^{n+k-i}, \end{aligned}$$

con $b_{-1} = (p-1)^3$, $b_0 = 3(p-1)^2$, $b_1 = 3(p-1)$ y $b_2 = 1$.

Por consiguiente, se deduce que:

$$\begin{aligned} \frac{d^i M_{L_n}(1)}{du} &= \frac{1}{p^3} \sum_{k=-1}^2 b_k n^i + \frac{1}{2p^3} \sum_{k=-1}^2 b_k i(2k+1-i) n^{i-1} + O(n^{i-2}) \\ &= n^i - \frac{i(ip+p-6)}{2p} n^{i-1} + O(n^{i-2}). \end{aligned}$$

Por Corolario 1.3.4, se concluye que para cada $\alpha \in \{1, 2, \dots\}$

$$\begin{aligned} \frac{d^\alpha M_{L_n}(0)}{dt} &= \\ p^\alpha n^\alpha - \frac{\alpha}{2p}(\alpha p + p - 6)p^\alpha n^{\alpha-1} + C_{\alpha,\alpha-1} p^{\alpha-1} n^{\alpha-1} + O(n^{\alpha-2}), \end{aligned}$$

como $C_{\alpha,\alpha-1} = C_{\alpha-1,\alpha-2} + (\alpha-1)C_{\alpha-1,\alpha-1} = C_{\alpha-1,\alpha-2} + \alpha - 1 = \frac{\alpha(\alpha-1)}{2}$. Entonces

$$\frac{d^\alpha M_{L_n}}{dt}(0) = p^\alpha n^\alpha + \frac{\alpha}{2}(\alpha(1-p) - p + 5)p^{\alpha-1}n^{\alpha-1} + O(n^{\alpha-2}).$$

□

En consecuencia, por el Teorema 1.3.5, los dos primeros momentos de Z_n^g para $\alpha \in \{3, 4, \dots\}$ vienen dados por:

$$\mathbb{E}(Z_n^g) = p^\alpha n^\alpha + \frac{\alpha}{2}(\alpha(1-p) - p + 5)p^{\alpha-1}n^{\alpha-1} + O(n^{\alpha-2}) \text{ y}$$

$$\mathbb{E}((Z_n^g)^2) = p^{2\alpha}n^{2\alpha} + \alpha(2\alpha(1-p) - p + 5)p^{2\alpha-1}n^{2\alpha-1} + O(n^{2\alpha-2}).$$

Entonces,

$$\mathbb{V}(Z_n^g) = \alpha^2(1-p)p^{2\alpha-1}n^{2\alpha-1} + O(n^{2\alpha-2}).$$

Teorema 1.3.6. *Dado $\alpha \in \{3, 4, \dots\}$, se verifica que $\frac{Z_n^g}{n^\alpha} \xrightarrow{P} p^\alpha$, cuando n tiende a infinito.*

Demostración. Sea $X_n = \frac{1}{n^\alpha}(Z_n^g - \frac{\alpha}{2}(\alpha(1-p) - p + 5)p^{\alpha-1}n^{\alpha-1} - O(n^{\alpha-2}))$ por la desigualdad de Chebyshev [24] obtenemos

$$\mathbb{P}(|X_n - p^\alpha| \geq \epsilon) \leq \frac{1}{n^{2\alpha}\epsilon^2}(\alpha^2(1-p)p^{2\alpha-1}n^{2\alpha-1} + O(n^{2\alpha-2})),$$

para cualquier $\epsilon > 0$. Si $n \rightarrow \infty$ entonces

$$\frac{1}{n^{2\alpha}\epsilon^2}(\alpha^2(1-p)p^{2\alpha-1}n^{2\alpha-1} + O(n^{2\alpha-2})) \rightarrow 0,$$

así $X_n \xrightarrow{P} p^\alpha$. Por otro lado, $X_n = \frac{Z_n^g}{n^\alpha} - x_n$ con $x_n = \frac{1}{n^\alpha}(\frac{\alpha}{2}(\alpha(1-p) - p + 5)p^{\alpha-1}n^{\alpha-1} + O(n^{\alpha-2}))$ luego $x_n \rightarrow 0$ cuando $n \rightarrow \infty$. Por lo tanto, el resultado se verifica. □

Corolario 1.3.7. *Para cualquier $\alpha \in \{3, 4, \dots\}$ y $r > 0$, se verifica que $\frac{Z_n^g}{n^\alpha} \xrightarrow{L_r} p^\alpha$, cuando n tiende a infinito.*

Demostración. Usando el Teorema 1.3.6 tenemos $\frac{Z_n^g}{n^\alpha} \xrightarrow{P} p^\alpha$ cuando n tiende a infinito. Por otro lado, $\frac{Z_n^g}{n^\alpha} \geq 0$ para todos los $n \geq 1$, luego $|\frac{Z_n^g}{n^\alpha}| = \frac{Z_n^g}{n^\alpha}$. Para cada $r > 0$ existe $N \in \{1, 2, \dots\}$ tal que $N > r$. Por el Teorema 4.2 en el Capítulo 5 de [24] obtenemos $\left\{ \left(\frac{Z_n^g}{n^\alpha} \right)^r, n \geq 1 \right\}$ es uniformemente integrable. Luego, aplicando el Teorema 5.4 en el Capítulo 5 de [24] con $X_n = \frac{Z_n^g}{n^\alpha}$, obtenemos la convergencia en L_r y se completa la demostración. \square

Observación 1.3.8. De acuerdo al Teorema 5.4 en el Capítulo 5 de [24] tenemos que $\mathbb{E} \left(\left(\frac{Z_n^g}{n^\alpha} \right)^r \right) \rightarrow p^{\alpha r}$ cuando $n \rightarrow \infty$.

Corolario 1.3.9. Para $n \geq 1$ y $\alpha \in \{3, 4, \dots\}$, en el modelo preferencial se verifican las siguientes propiedades:

- 1) $\mathbb{E}(Z_n^g) = \frac{n^\alpha}{2^\alpha} + \frac{\alpha(\alpha+9)}{2^{\alpha+1}} n^{\alpha-1} + O(n^{\alpha-2})$,
- 2) $\mathbb{E}((Z_n^g)^2) = \frac{n^{2\alpha}}{4^\alpha} + \frac{\alpha(2\alpha+9)}{4^\alpha} n^{2\alpha-1} + O(n^{2\alpha-2})$,
- 3) $\mathbb{V}(Z_n^g) = \frac{\alpha^2}{4^\alpha} n^{2\alpha-1} + O(n^{2\alpha-2})$.
- 4) Para todo $r > 0$, $\frac{Z_n^g}{n^\alpha} \xrightarrow{L_r} \frac{1}{2^\alpha}$ cuando $n \rightarrow \infty$.

Por último, dado $n \in \mathbb{N}$ fijo, tenemos que $L_n \in [3, n+2]$. Si nos concentramos en todos los posibles RST s generados hasta la etapa n , se tendría que, $L_n = 3$ es el caso donde hasta la etapa n no fue elegido el centro, es decir, son todos los ST s con $n-1$ nodos internos divididos en 3 hojas (coincidiendo con el caso $p = 0$) y $L_n = n+2$ es el caso donde hasta la etapa n sólo fue elegido el centro y el ST coincidiría con la estrella S_{n+3} (coincidiendo con el caso $p = 1$). El objetivo es estudiar los valores extremos del índice de Zagreb generalizado sobre los RST s generados hasta la etapa n . Estudiando la función $Z_n^g := f(L_n)$ con $L_n \in [3, n+2]$ se concluye que:

- Para $\alpha = 0$, $f(L_n) = n + 3$.
- Para $\alpha = 1$, $f(L_n) = 2n + 4$.

- Para $\alpha > 1$ ó $\alpha < 0$, $f(3) \leq f(L_n) \leq f(n+2)$.
- Para $0 < \alpha < 1$, $f(n+2) \leq f(L_n) \leq f(3)$.

Además las igualdades se verifican si y sólo si $L_n = 3$ ó $L_n = n+2$, respectivamente. Para comprobar lo anterior veamos que: Primero, x^α con $\alpha \in \mathbb{R}$ y $x \in \mathbb{R}^+$ es una función derivable, además es estrictamente convexa para $\alpha > 1$ ó $\alpha < 0$ y estrictamente cóncava para $0 < \alpha < 1$ en \mathbb{R}^+ . Por ende para $x, y \in \mathbb{R}^+$ con $x \neq y$ se verifica que

- $x^\alpha > y^\alpha + \alpha \cdot y^{\alpha-1} \cdot (x - y)$ para $\alpha > 1$ ó $\alpha < 0$.
- $x^\alpha < y^\alpha + \alpha \cdot y^{\alpha-1} \cdot (x - y)$ para $0 < \alpha < 1$.

Así, tomando $x = 2$ y $y = 3$ se sigue que

- $2^\alpha > 3^\alpha + \alpha 3^{\alpha-1}(2 - 3)$ para $\alpha > 1$ ó $\alpha < 0$.
- $2^\alpha < 3^\alpha + \alpha 3^{\alpha-1}(2 - 3)$ para $0 < \alpha < 1$.

Por otro lado, usando nuevamente las propiedades de convexidad y concavidad de la función x^α con $\alpha \in \mathbb{R}$ y $x \in \mathbb{R}^+$ tenemos que

- $\frac{3^\alpha}{2} + \frac{1^\alpha}{2} > \left(\frac{3}{2} + \frac{1}{2}\right)^\alpha$ para $\alpha > 1$ ó $\alpha < 0$.
- $\frac{3^\alpha}{2} + \frac{1^\alpha}{2} < \left(\frac{3}{2} + \frac{1}{2}\right)^\alpha$ para $0 < \alpha < 1$.

Por lo tanto usando lo anterior se concluye que

- $2^\alpha - 1 < \alpha \cdot 3^{\alpha-1}$ para $\alpha > 1$ ó $\alpha < 0$.
- $2^\alpha - 1 > \alpha \cdot 3^{\alpha-1}$ para $0 < \alpha < 1$.

Ahora, $f'(L_n) = \alpha \cdot L_n^{\alpha-1} + 1 - 2^\alpha$ con $L_n \in [3, n+2]$. Usando lo antes visto y la monotonía de x^α con $\alpha \in \mathbb{R}$ y $x \in \mathbb{R}^+$, tenemos que

- $2^\alpha - 1 < \alpha \cdot 3^{\alpha-1} < \alpha \cdot L_n^{\alpha-1}$ para $\alpha > 1$.
- $1 - 2^\alpha > -\alpha \cdot 3^{\alpha-1} > -\alpha \cdot L_n^{\alpha-1}$ para $\alpha < 0$.
- $2^\alpha - 1 > \alpha \cdot 3^{\alpha-1} > \alpha \cdot L_n^{\alpha-1}$ para $0 < \alpha < 1$.

Así, para $\alpha > 1$ ó $\alpha < 0$ se tiene que $f(L_n)$ con $L_n \in [3, n + 2]$ es estrictamente creciente y para $0 < \alpha < 1$ se tiene que $f(L_n)$ con $L_n \in [3, n + 2]$ es estrictamente decreciente. Concluyendo lo requerido. Cabe destacar que en [46], se demostró que si nos fijamos en todos los árboles de n vértices el índice de Zagreb generalizado con $\alpha > 1$ ó $\alpha < 0$ alcanza su máximo y mínimo en S_n y P_n respectivamente. Por otro lado, para $0 < \alpha < 1$, el índice de Zagreb generalizado alcanza su máximo y mínimo en P_n y S_n respectivamente. Lo que coincide con lo obtenido anteriormente, sin incluir el camino ya que este no es un ST .

1.3.1.1. Índice de Zagreb

En el momento $n \geq 1$, tomando $h(x) = x$ y $\alpha = 2$ en (1.3) obtenemos el índice de Zagreb (Z_n). Según (1.5) tenemos

$$Z_n = L_n^2 - 3L_n + 4(n + 2). \quad (1.8)$$

Así, podemos obtener los momentos de Z_n mediante (1.8). A modo de ejemplo, en Figura 1.2 se presenta el índice de Zagreb para algunos STs .

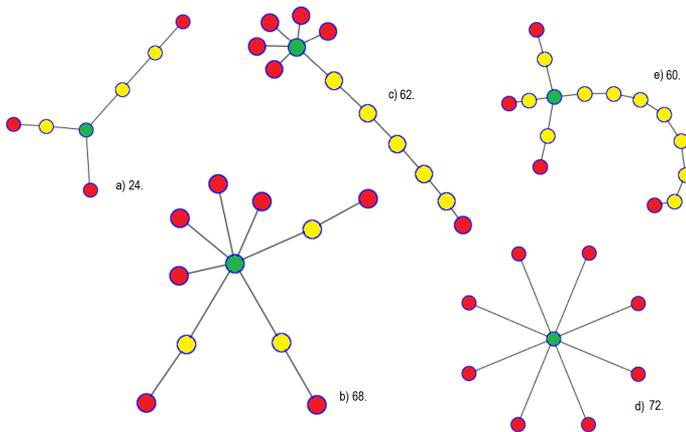


Figura 1.2: Índice de Zagreb para los STs mostrados en Figura 1.1.

Claramente, para $n \geq 1$

$$\mathbb{E}(Z_n) = n^2 p^2 + (-3p^2 + 4p + 4)n + 2p^2 - 4p + 8 \text{ y}$$

$$\mathbb{V}(Z_n) = (-4p^4 + 4p^3)n^3 + (22p^4 - 40p^3 + 18p^2)n^2 + (-38p^4 + 92p^3 - 70p^2 + 16p)n + 20p^4 - 56p^3 + 52p^2 - 16p.$$

Proposición 1.3.10. *Dado $p \in [0, 1]$, cuando n tiende a infinito se verifica que*

$$1) \text{ Para todos los } k \in \mathbb{R}, \frac{Z_n - n^2 p^2}{2\sqrt{p^3(1-p)(n+k)^3}} \xrightarrow{D} N(0, 1).$$

$$2) \text{ Para todos los } r > 0, \frac{Z_n}{n^2} \xrightarrow{Lr} p^2.$$

Demostración. 1) Por (1.8) $Z_n = (L_n - \frac{3}{2})^2 + \frac{16n+23}{4}$ con $n \geq 1$. Tenga en cuenta que el Teorema 1.2.2 implica que

$$\frac{L_n - \frac{3}{2}}{\sqrt{p(1-p)(n+k)}},$$

es equivalente a una variable aleatoria normal con media

$$\frac{\frac{3}{2} + (n-1)p}{\sqrt{p(1-p)(n+k)}}$$

y varianza 1 en distribución. Por lo tanto,

$$\left(\frac{L_n - \frac{3}{2}}{\sqrt{p(1-p)(n+k)}} \right)^2 \sim \chi^2 \left(\frac{(\frac{3}{2} + (n-1)p)^2}{p(1-p)(n+k)}, 1 \right).$$

Indicándonos que:

$$\frac{Z_n}{p(1-p)(n+k)} - \frac{16n+23}{4p(1-p)(n+k)} \sim \chi^2 \left(\frac{(\frac{3}{2} + (n-1)p)^2}{p(1-p)(n+k)}, 1 \right).$$

Por la aproximación normal de la distribución chi-cuadrada no central (véase [64]), se obtiene que

$$\frac{\frac{Z_n}{p(1-p)(n+k)} - \frac{16n+23}{4p(1-p)(n+k)} - (1 + \frac{(\frac{3}{2} + (n-1)p)^2}{p(1-p)(n+k)})}{\sqrt{2(1 + \frac{2(\frac{3}{2} + (n-1)p)^2}{p(1-p)(n+k)})}} \xrightarrow{D} N(0, 1)$$

cuando $n \rightarrow \infty$. En particular,

$$\begin{aligned}
 & \frac{\frac{Z_n}{p(1-p)(n+k)} - \frac{16n+23}{4p(1-p)(n+k)} - 1 - \frac{(\frac{3}{2} + (n-1)p)^2}{p(1-p)(n+k)}}{\sqrt{2(1 + \frac{2(\frac{3}{2} + (n-1)p)^2}{p(1-p)(n+k)})}} \\
 &= \frac{Z_n - \frac{16n+23}{4} - p(1-p)(n+k) - (\frac{3}{2} + (n-1)p)^2}{\sqrt{p(1-p)(n+k)}\sqrt{2(p(1-p)(n+k) + 2(\frac{3}{2} + (n-1)p)^2)}} \\
 &= \frac{Z_n - \frac{16n+23}{4} - p(1-p)(n+k) - (\frac{3}{2} + (n-1)p)^2}{\sqrt{4p^3(1-p)(n+k)^3 + o(n^3)}} \\
 &= a_n \frac{Z_n - n^2 p^2}{\sqrt{4p^3(1-p)(n+k)^3}} + b_n.
 \end{aligned}$$

donde

$$\begin{aligned}
 a_n &= \frac{\sqrt{4p^3(1-p)(n+k)^3}}{\sqrt{4p^3(1-p)(n+k)^3 + o(n^3)}}, \\
 b_n &= \frac{o(n^{\frac{3}{2}})}{\sqrt{4p^3(1-p)(n+k)^3 + o(n^3)}}.
 \end{aligned}$$

Se verifica que $a_n \rightarrow 1$ y $b_n \rightarrow 0$ cuando $n \rightarrow \infty$, luego

$$\frac{Z_n - n^2 p^2}{2\sqrt{p^3(1-p)(n+k)^3}} \xrightarrow{D} N(0, 1).$$

2) La prueba se puede hacer de manera similar a la del Corolario 1.3.7. \square

Corolario 1.3.11. *El índice de Zagreb en el modelo preferencial verifica las siguientes identidades para $n \geq 1$:*

$$\begin{aligned}
 1) \quad \mathbb{E}(Z_n) &= \frac{n^2}{4} + \frac{21n}{4} + \frac{13}{2}, \\
 2) \quad \mathbb{V}(Z_n) &= \frac{n^3}{4} + \frac{7n^2}{8} - \frac{3n}{8} - \frac{3}{4}.
 \end{aligned}$$

3) Para toda $k \in \mathbb{R}$, $\frac{Z_n - \frac{n^2}{4}}{\frac{(n+k)^{\frac{3}{2}}}{2}} \xrightarrow{D} N(0, 1)$ cuando $n \rightarrow \infty$.

4) Para toda $r > 0$, $\frac{Z_n}{n^2} \xrightarrow{Lr} \frac{1}{4}$, cuando $n \rightarrow \infty$.

Usando el lenguaje Python, realizamos un experimento numérico para complementar la Proposición 1.3.10 1) con $k = 0$, desarrollada en esta sección. Dado un p fijo en $(0, 1)$, vía el siguiente pseudocódigo generamos de forma independiente 5, 000 araña aleatorios después de $n = 10, 000$ pasos evolutivos.

Algorithm 1 Generar *RSTs*.

Input: $p \in (0, 1)$: parámetro del modelo, M : número de *RSTs*,
 n : pasos evolutivos de los *RST*.

Output: Una muestra de M *RSTs* con parámetro p al tiempo n .

Sample = \emptyset

for $i = 1$ **to** M **do**

 Simular una $Bin(n - 1, p)$.

$L_n = 3 + Bin(n - 1, p)$.

Sample = *Sample* \cup $\{L_n\}$.

end for

Para cada *RST* simulado, ya que conocemos el número de hojas se calcula el índice de Zagreb. El histograma de los datos de la muestra con una normal aproximada vía el método del núcleo se da en las Figuras 1.3, 1.4 y 1.5. Además, confirmamos la conclusión a través de la prueba de normalidad Shapiro-Wilk con un nivel de significancia de 0.05, obteniendo un p -value igual a 0.070, 0.365 y 0.469 para $p = 0.3, 0.5$ y 0.7 , respectivamente.

1.3.1.2. Índice de Gordon-Scantlebury

Definiendo S_n como el índice de Gordon-Scantlebury en el tiempo $n \geq 1$, sabemos que se verifica $Z_n = 2(S_n + E_n)$ [53] donde E_n es el número de aristas en el tiempo n . Por otro lado, el árbol generado por el modelo en el tiempo n tiene $n + 3$ nodos, entonces tiene $n + 2$ aristas, por lo tanto, $S_n = \frac{Z_n}{2} - n - 2$. Para $n \geq 1$, obtenemos

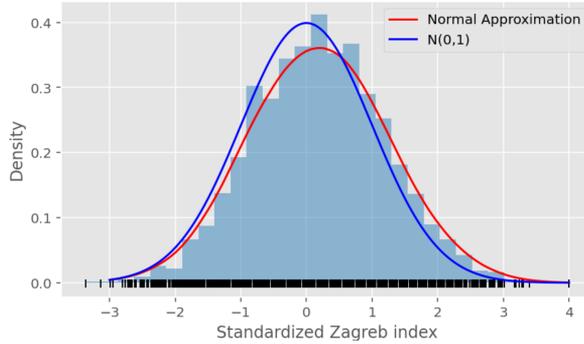


Figura 1.3: Histograma de los índices estandarizados de Zagreb; la curva roja es la densidad estimada de la muestra ($p = 0.3$).

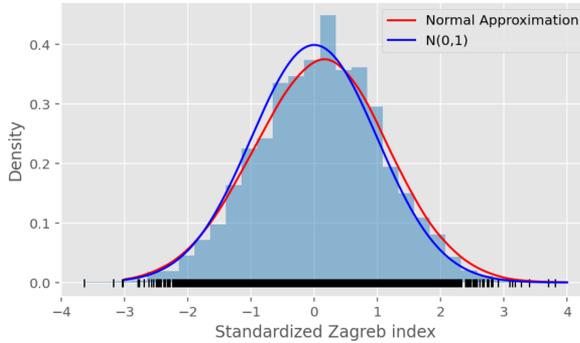


Figura 1.4: Histograma de los índices estandarizados de Zagreb; la curva roja es la densidad estimada de la muestra ($p = 0.5$).

$$\mathbb{E}(S_n) = \frac{n^2 p^2}{2} + \left(-\frac{3}{2}p^2 + 2p + 1\right)n + p^2 - 2p + 2 \text{ y}$$

$$\mathbb{V}(S_n) = (-p^4 + p^3)n^3 + \left(\frac{11}{2}p^4 - 10p^3 + \frac{9}{2}p^2\right)n^2 + \left(-\frac{19}{2}p^4 + 23p^3 - \frac{35}{2}p^2 + 4p\right)n + 5p^4 - 14p^3 + 13p^2 - 4p.$$

Proposición 1.3.12. Para $p \in [0, 1]$, cuando n tiende a infinito se verifica que

$$1) \text{ Para toda } k \in \mathbb{R}, \frac{S_n - \frac{n^2 p^2}{2}}{\sqrt{p^3(1-p)(n+k)^3}} \xrightarrow{D} N(0, 1).$$

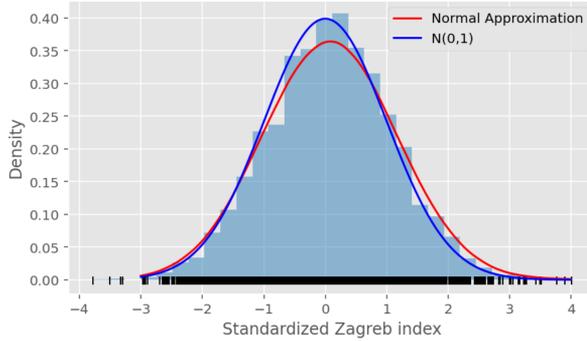


Figura 1.5: Histograma de los índices estandarizados de Zagreb; la curva roja es la densidad estimada de la muestra ($p = 0.7$).

2) Para toda $r > 0$, $\frac{S_n}{n^2} \xrightarrow{Lr} \frac{p^2}{2}$.

Demostración. 1) La proposición 1.3.10 1) implica que

$$\frac{Z_n - n^2 p^2}{2\sqrt{p^3(1-p)(n+k)^3}} \xrightarrow{D} N(0, 1).$$

Además,

$$\begin{aligned} \frac{Z_n - n^2 p^2}{2\sqrt{p^3(1-p)(n+k)^3}} &= \frac{\frac{Z_n}{2} - \frac{n^2 p^2}{2}}{\sqrt{p^3(1-p)(n+k)^3}} \\ &= \frac{S_n - \frac{n^2 p^2}{2} + n + 2}{\sqrt{p^3(1-p)(n+k)^3}} \xrightarrow{D} N(0, 1). \end{aligned}$$

Tenga en cuenta que

$$\frac{n+2}{\sqrt{p^3(1-p)(n+k)^3}} \rightarrow 0 \text{ cuando } n \rightarrow \infty.$$

Por lo tanto, la demostración se concluye.

2) Se demuestra directamente por la Proposición 1.3.10 2). \square

Corolario 1.3.13. *El índice de Gordon-Scantlebury en el modelo preferencial verifica las siguientes identidades para $n \geq 1$:*

$$1) \mathbb{E}(S_n) = \frac{n^2}{8} + \frac{13n}{8} + \frac{5}{4},$$

$$2) \mathbb{V}(S_n) = \frac{n^3}{16} + \frac{7n^2}{32} - \frac{3n}{32} - \frac{3}{16}.$$

$$3) \text{ Para toda } k \in \mathbb{R}, \frac{S_n - \frac{n^2}{8}}{\frac{(n+k)^{3/2}}{4}} \xrightarrow{D} N(0, 1), \text{ cuando } n \rightarrow \infty.$$

$$4) \text{ Para toda } r > 0, \frac{S_n}{n^2} \xrightarrow{Lr} \frac{1}{8}, \text{ cuando } n \rightarrow \infty.$$

1.3.1.3. Índice de Platt

Dejemos que P_n denote el índice de Platt en el tiempo $n \geq 1$, verificando que $P_n = 2S_n$ [53]. Así obtenemos que

$$\mathbb{E}(P_n) = n^2 p^2 + (-3p^2 + 4p + 2)n + 2p^2 - 4p + 4 \text{ y}$$

$$\mathbb{V}(P_n) = \mathbb{V}(Z_n).$$

Proposición 1.3.14. *Para $p \in [0, 1]$, cuando n tiende a infinito se verifica que*

$$1) \text{ Para toda } k \in \mathbb{R}, \frac{P_n - n^2 p^2}{2\sqrt{p^3(1-p)(n+k)^3}} \xrightarrow{D} N(0, 1).$$

$$2) \text{ Para toda } r > 0, \frac{P_n}{n^2} \xrightarrow{Lr} p^2.$$

Demostración. 1) Según la Proposición 1.3.12 1) obtenemos

$$\frac{S_n - \frac{n^2 p^2}{2}}{\sqrt{p^3(1-p)(n+k)^3}} \xrightarrow{D} N(0, 1).$$

Además,

$$\frac{S_n - \frac{n^2 p^2}{2}}{\sqrt{p^3(1-p)(n+k)^3}} = \frac{P_n - n^2 p^2}{2\sqrt{p^3(1-p)(n+k)^3}} \xrightarrow{D} N(0, 1).$$

La prueba está completa.

2) La prueba es fácilmente verificada por la Proposición 1.3.12 2). \square

Finalmente, tomando $p = 1/2$ en los resultados antes vistos obtenemos el siguiente corolario.

Corolario 1.3.15. *El índice de Platt del modelo preferencial verifica las siguientes identidades para $n \geq 1$:*

$$1) \mathbb{E}(P_n) = \frac{n^2}{4} + \frac{13n}{4} + \frac{5}{2},$$

$$2) \mathbb{V}(P_n) = \frac{n^3}{4} + \frac{7n^2}{8} - \frac{3n}{8} - \frac{3}{4}.$$

$$3) \text{ Para toda } k \in \mathbb{R}, \frac{P_n - \frac{n^2}{4}}{\frac{(n+k)^{\frac{3}{2}}}{2}} \xrightarrow{D} N(0, 1), \text{ cuando } n \rightarrow \infty.$$

$$4) \text{ Para toda } r > 0, \frac{P_n}{n^2} \xrightarrow{Lr} \frac{1}{4}, \text{ cuando } n \rightarrow \infty.$$

1.3.1.4. Índice olvidado

En el momento $n \geq 1$, tomando $h(x) = x$ y $\alpha = 3$ en (1.3) obtenemos el índice Olvidado (F_n). Según (1.5) tenemos

$$F_n = L_n^3 - 7L_n + 8(n + 2).$$

A modo de ejemplo, en Figura 1.6 se presenta el índice Olvidado para algunos STs .

Nuestra siguiente tarea es calcular el primer momento y la varianza de F_n ,

$$\mathbb{E}(F_n) = n^3 p^3 + (12p^2 - 6p^3)n^2 + (11p^3 - 36p^2 + 30p + 8)n - 6p^3 + 24p^2 - 30p + 22$$

y

$$\mathbb{V}(F_n) = 9p^5(1-p)n^5 - 9p^4(1-p)(13p-18)n^4 + 3p^3(1-p)(197p^2 - 490p + 302)n^3 - 9p^2(1-p)(159p^3 - 530p^2 + 572p - 192)n^2 - 6p(1-p)^2(272p^3 - 803p^2 + 714p - 150)n + 36p(1-p)^2(19p^3 - 64p^2 + 71p - 25).$$

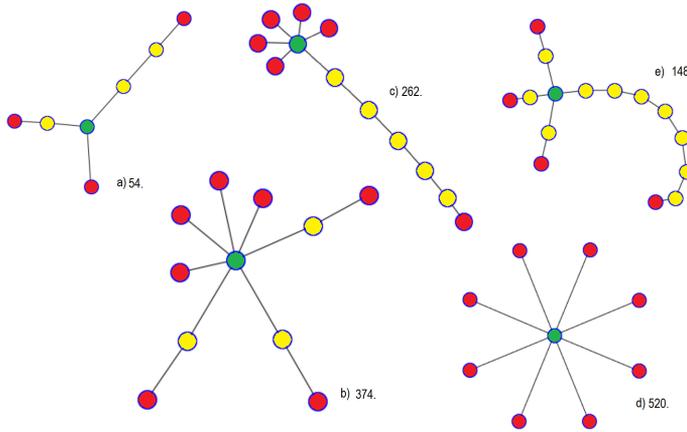


Figura 1.6: Índice Olvidado para los STs mostrados en Figura 1.1.

Según el Corolario 1.3.7, obtenemos el siguiente resultado.

Corolario 1.3.16. *Para todo $r > 0$ se verifica que, $\frac{F_n}{n^3} \xrightarrow{L_r} p^3$ cuando $n \rightarrow \infty$.*

Corolario 1.3.17. *El índice olvidado en el modelo preferencial verifica las siguientes identidades para $n \geq 1$:*

- 1) $\mathbb{E}(F_n) = \frac{n^3}{8} + \frac{9n^2}{4} + \frac{123n}{8} + \frac{49}{4},$
- 2) $\mathbb{V}(F_n) = \frac{9n^5}{64} + \frac{207n^4}{64} + \frac{1275n^3}{64} + \frac{1341n^2}{64} - \frac{483n}{16} - \frac{225}{16}.$
- 3) *Para todo $r > 0$, $\frac{F_n}{n^3} \xrightarrow{L_r} \frac{1}{8}$ cuando $n \rightarrow \infty$.*

1.3.2. Índice de Gini basado en grados

Recientemente, un índice de Gini basado en grados para grafos generales fue propuesto por [12]. Este último índice es una medida topológica en un grafo que captura la proximidad a grafos regulares. En [61] los autores consideraron el índice de Gini basado en grados introducido por [12], con ligeras modificaciones. En esta sección, estudiamos el índice de Gini basado en grados introducido en [61]. Por definición, el índice de Gini basado en grados de un

grafo dentro de la clase de RST en el momento $n \geq 1$ viene dado por

$$G_n = \frac{\sum_{i < j \in V_n} |deg_i - deg_j|}{(n+3)^2 \mathbb{E}(deg_{v^*})},$$

donde v^* es un vértice arbitrario (elegido uniformemente) de un grafo seleccionado aleatoriamente de la clase de RST y V_n denota el conjunto de vértices al tiempo n . Tomamos $\mathbb{E}(G_n)$ como el índice de Gini basado en grados de la clase. Realizando la suma sin repetir pares de vértices, debido a las características del modelo, obtenemos

$$\begin{aligned} & \sum_{i < j \in V_n} |deg_i - deg_j| \\ &= |L_n - 1|L_n + |L_n - 2|(n + 2 - L_n) + L_n(n + 2 - L_n) \\ &= -L_n^2 + (2n + 5)L_n - 2n - 4, \end{aligned}$$

ya que $L_n \geq 3$ para todos los $n \geq 1$. Finalmente

$$\mathbb{E}(d_{v^*}) = \frac{2(n+2)}{n+3},$$

así,

$$G_n = \frac{-L_n^2 + (2n + 5)L_n - 2n - 4}{2(n+3)(n+2)}. \quad (1.9)$$

A modo de ejemplo, en Figura 1.7 se presenta el índice de Gini para algunos STs .

Siguiendo, de acuerdo a (1.9) se deduce que

$$\mathbb{E}(G_n) = \frac{(2p-p^2)n^2 + (3p^2-4p+4)n - 2p^2 + 2p + 2}{2(n+3)(n+2)}.$$

A continuación, obtenemos una propiedad asintótica del índice de Gini basado en grados de la clase de RST en el tiempo n .

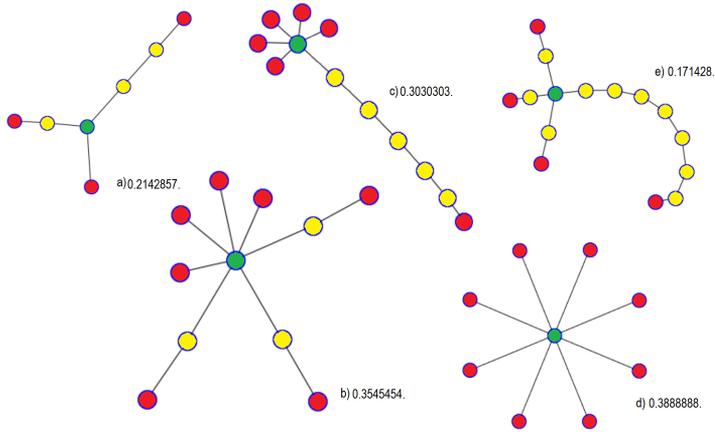


Figura 1.7: Índice de Gini para los ST s mostrados en Figura 1.1.

Proposición 1.3.18. *Cuando $n \rightarrow \infty$, tenemos $\mathbb{E}(G_n) \rightarrow \frac{p(2-p)}{2}$.*

Vemos en (1.9) que

$$G_n = \frac{-(L_n - \frac{2n+5}{2})^2 + (\frac{2n+5}{2})^2 - 2n - 4}{2(n+3)(n+2)}.$$

Por lo tanto,

$$\mathbb{V}(G_n) = \frac{4p(1-p)^3 n^3 + 2p(11p-6)(1-p)^2 n^2 + 2p(1-p)(19p^2 - 23p + 6)n - 4p(1-p)(5p^2 - 5p + 1)}{4(n+3)^2(n+2)^2}.$$

Teorema 1.3.19. *Se verifica que $G_n \xrightarrow{P} \frac{p(2-p)}{2}$, cuando $n \rightarrow \infty$.*

Demostración. Por la desigualdad de Chebyshev [24] tenemos

$$\mathbb{P}(|G_n - \mathbb{E}(G_n)| \geq \epsilon) \leq \frac{\mathbb{V}(G_n)}{\epsilon^2}.$$

para cualquier $\epsilon > 0$. Si $n \rightarrow \infty$ entonces $\frac{\mathbb{V}(G_n)}{\epsilon^2} \rightarrow 0$, así $G_n - \mathbb{E}(G_n) \xrightarrow{P} 0$. Por lo tanto, la Proposición 1.3.18 completa la prueba. \square

Corolario 1.3.20. *Para todo $r > 0$, tenemos $G_n \xrightarrow{L_r} \frac{p(2-p)}{2}$, cuando n tiende a infinito.*

Demostración. Argumentado de manera similar al Corolario 1.3.7 vía el Teorema 1.3.19, el resultado es válido. \square

Una consecuencia directa de la Proposición 1.3.18 y el Corolario 1.3.20 es el siguiente resultado.

Corolario 1.3.21. *En el modelo preferencial para todo $r > 0$, se verifica que $G_n \xrightarrow{L_r} \frac{3}{8}$ y $\mathbb{E}(G_n) \rightarrow \frac{3}{8}$, cuando $n \rightarrow \infty$.*

Observación 1.3.22. 1. *En vista de la Proposición 1.3.18 y el Teorema 1.3.19 definimos $f(p) = \frac{p(2-p)}{2}$, $p \in [0, 1]$. Nótese que f es estrictamente creciente, en consecuencia, si queremos una clase más regular debemos elegir valores más pequeños de p , ya que en [12] se muestra que un valor menor del índice de Gini basado en grados sugiere más regularidad de un grafo o una clase de grafos; lo cual tiene sentido en este caso ya que el centro tendría un grado menor. De hecho la clase menos regular se obtiene cuando $p = 0$, donde el grado del centroide sería 3 y la clase más irregular se obtiene con $p = 1$ donde el grado del centroide al tiempo n sería $n + 2$. Específicamente para $p \in [0, \frac{1}{2})$ tenemos que la clase de RST que evolucionan de manera preferencial es relativamente menos regular que la clase de RST estudiada en este trabajo (ver Figura 1.8).*

2. *En [12] los autores compararon la regularidad de dos clases de árboles binarios. Los autores mostraron que la clase de árboles binarios uniformes tiene un índice de Gini basado en grados $\frac{3}{16}$ ([12], Sección 6.1), mientras que la de la clase de árboles de búsqueda binarios es $\frac{2}{9}$ ([12], Sección 6.2). Tenga en cuenta que para $p \leq 1 - \frac{\sqrt{10}}{4} \approx 0.209$, tenemos $f(p) \leq \frac{3}{16} < \frac{2}{9}$, entonces la clase de árboles de búsqueda binarios y binarios uniformes son relativamente menos regulares que la clase de RST estudiada en este trabajo para $p \in [0, 1 - \frac{\sqrt{10}}{4}]$ (ver Figura 1.8)*

3. *En ([80], Sección 2), los autores concluyeron que el índice de Gini basado en grados para la clase de orugas aleatorias es $\frac{1}{2}$. Dado que $f(p) < \frac{1}{2}$ para $p \in [0, 1)$, concluimos que la clase de RST es más regular que la clase de orugas aleatorias de [80] (ver Figura 1.8).*

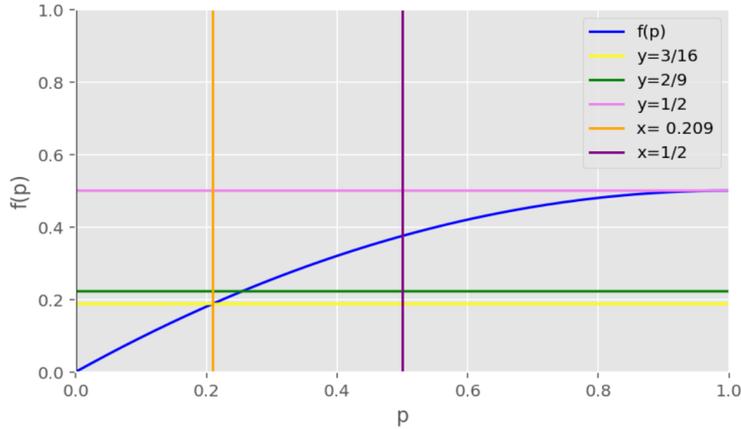


Figura 1.8: Representación gráfica de las funciones definidas en la Observación 1.3.22 donde $f(p) = \frac{p(2-p)}{2}$.

Observación 1.3.23. *Los resultados expuestos en las Secciones 3.2.1 y 3.2.2 de [61] se obtienen como un caso especial de los resultados demostrados en las Secciones 1.3.1 y 1.3.2, respectivamente. Por último, note que dado $n \in \mathbb{N}$ fijo, el índice de Gini sobre todos los RST generados hasta la etapa n alcanza el mínimo y el máximo en $L_n = 3$ y $L_n = n + 2$, respectivamente.*

1.3.3. Índice Hoover basado en grados

En [80] los autores propusieron, como una medida competitiva para evaluar la regularidad de los grafos, un índice de Hoover basado en grados análogo al índice de Gini basado en grados introducido en [12]. En nuestro contexto, en el momento $n \geq 1$ el índice de Hoover basado en grados de un grafo dentro de la clase de RSTs (H_n) se define de la siguiente manera:

$$H_n = \frac{\sum_{i \in V_n} |(n+3)deg_i - 2(n+2)|}{4(n+2)(n+3)},$$

donde V_n denota el conjunto de vértices al tiempo n . A modo de ejemplo, en Figura 1.9 se presenta el índice de Hoover para algunos

STs.

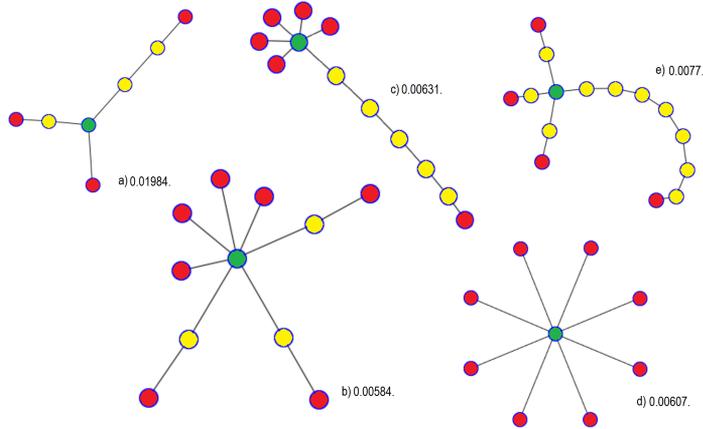


Figura 1.9: Índice de Hoover para los *STs* mostrados en Figura 1.1.

De manera similar, tomamos $\mathbb{E}(H_n)$ como el índice Hoover basado en grados de la clase. En esta sección se utiliza el mismo análisis aplicado en la Sección 1.3.2 y obtenemos los siguientes resultados.

Proposición 1.3.24. *Para $n \geq 1$, tenemos*

- 1) $H_n = \frac{(n+1)L_n}{2(n+3)(n+2)}$,
- 2) $\mathbb{E}(H_n) = \frac{pn^2+3n+3-p}{2(n+3)(n+2)}$,
- 3) $\mathbb{V}(H_n) = \frac{p(1-p)(n^2-1)(n+1)}{4(n+3)^2(n+2)^2}$.
- 4) *Para todo $r > 0$, $H_n \xrightarrow{L_r} \frac{p}{2}$, cuando $n \rightarrow \infty$.*

Una consecuencia directa de la Proposición 1.3.24 es el siguiente corolario.

Corolario 1.3.25. *En el modelo preferencial para todo $r > 0$, se verifica que $H_n \xrightarrow{L_r} \frac{1}{4}$ y $\mathbb{E}(H_n) \rightarrow \frac{1}{4}$ cuando n tiende a infinito.*

- Observación 1.3.26.** 1. En vista de la Proposición 1.3.24 definimos $f_1(p) = \frac{p}{2}, p \in [0, 1]$, luego f_1 es estrictamente creciente. Dado que se muestra en [80] que un valor más cercano a 0 sugiere que los grafos de la clase tienden a ser más regulares, por un argumento similar a la Observación 1.3.22 1) obtenemos el mismo comportamiento que el índice de Gini basado en grados estudiado en la Sección 1.3.2.
2. Además, se observa que $f_1(p) < f(p)$ para $p \in (0, 1)$. Entonces, el índice de Hoover basado en grados de la clase de RST presentada en este capítulo es menor que el índice de Gini basado en grados de la misma clase cuando n tiende a infinito.
3. En ([80], Sección 3), los autores concluyeron que el índice de Hoover basado en grados para la clase de orugas aleatorias es $\frac{1}{2}$ como $n \rightarrow \infty$. Dado que $f_1(p) < \frac{1}{2}$ para $p \in [0, 1)$, concluimos que la clase de RST es más regular que la clase de orugas aleatorias de [80], vía el índice de Hoover basado en grados.
4. Por último, nota que dado $n \in \mathbb{N}$ fijo, el índice de Hoover sobre todos los RST generados hasta la etapa n alcanza el mínimo y el máximo en $L_n = 3$ y $L_n = n + 2$, respectivamente.

Capítulo 2

Cadenas aleatorias de espiros

2.1. Preliminares

La Teoría de Martingalas es una herramienta matemática muy potente. El concepto fue introducido por Paul Lévy en 1934, y recibió su nombre en 1939 por J. André Ville. El desarrollo de toda una teoría se debe a Joseph L. Doob. Hoy en día, el concepto de martingala es bien conocido. En particular, existen teoremas del límite central de martingalas, que dan condiciones bajo las cuales la distribución del proceso se aproxima a la normal. De hecho, en [18] y [36] los autores utilizaron un enfoque de martingala para estudiar índices topológicos, como los índices de Zagreb, Gordon-Scantlebury y Platt.

Por otro lado, en [10, 16, 35, 45, 58–60, 71, 81] los autores estudiaron los índices topológicos en cadenas aleatorias. En particular, los compuestos espiros son compuestos químicos en los que podemos identificar anillos unidos por un sólo átomo. Los derivados de los espiros son productos químicos que se ven con bastante frecuencia y que pueden utilizarse en la síntesis orgánica, la síntesis de fármacos intercambiadores de calor, etc. Motivados por la información anterior, en este apartado estudiamos índices topológicos en las cadenas aleatorias de espiros (véase la Definición 2.1.1). Nuestro objetivo es asociar una martingala al índice topológico, de forma que, las propiedades que se pueden deducir de la martingala sean

útiles para mostrar las del índice topológico. Finalmente, es importante mencionar que lo desarrollado en esta sección se encuentra basado en el artículo [67].

Definición 2.1.1. *La cadena aleatoria de espiros, denotada por $RSC_n = RSC(n, p_1, p_2, p_3)$ con n hexágonos se construye de la siguiente manera:*

- RSC_1 es un hexágono RSC_2 contiene dos hexágonos, ver Figura 2.1.
- Para cada $n > 2$, RSC_n se construye uniendo un hexágono a RSC_{n-1} de tres maneras, dando lugar a $RSC_n^1, RSC_n^2, RSC_n^3$ con probabilidad p_1, p_2 y p_3 respectivamente, donde $0 < p_i < 1$ y $p_1 + p_2 + p_3 = 1$, ver Figura 2.2.

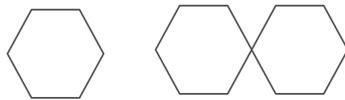


Figura 2.1: Los grafos de RSC_1 y RSC_2 .

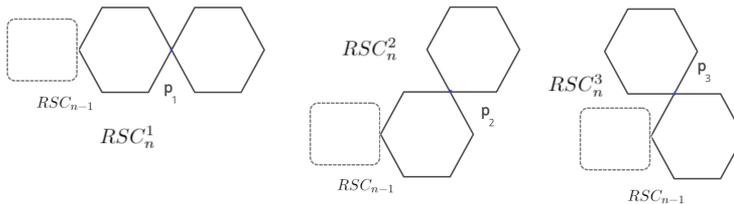


Figura 2.2: Los tres links para $RSC_n (n > 2)$.

2.2. Índices topológicos sobre cadenas aleatorias de espiros

Sea $G = (V(G), E(G))$, muchos índices topológicos importantes (basados en grados) pueden ser definidos como

$$TI(G) = \sum_{vu \in E(G)} f(d_v, d_u), \quad (2.1)$$

donde $f : \{1, 2, \dots\} \times \{1, 2, \dots\} \rightarrow (0, \infty)$ es cualquier función con la propiedad de que $f(x, y) = f(y, x)$ para $x, y \in \{1, 2, \dots\}$ (simétrica) y d_v es el grado de un nodo $v \in V(G)$. Los principales índices topológicos de la forma (2.1) son:

1. Si $f(x, y) = x + y$ entonces $TI(G)$ es el primer índice de Zagreb.
2. Si $f(x, y) = x^{-2} + y^{-2}$ entonces $TI(G)$ es el índice de grado inverso.
3. Si $f(x, y) = x^2 + y^2$ entonces $TI(G)$ es el índice Olvidado.
4. Si $f(x, y) = xy$ entonces $TI(G)$ es el segundo índice de Zagreb.
5. Si $f(x, y) = 1/\sqrt{xy}$ entonces $TI(G)$ es el índice de Randić habitual.
6. Si $f(x, y) = \frac{xy}{x+y}$ entonces $TI(G)$ es el índice indeg de la suma inversa.
7. Si $f(x, y) = \frac{x^2+y^2}{xy}$ entonces $TI(G)$ es el índice de diferencia simétrica.
8. Si $f(x, y) = 2\frac{\sqrt{xy}}{x+y}$ entonces $TI(G)$ es el índice geométrico-aritmético.
9. Si $f(x, y) = \sqrt{x^2 + y^2}$ entonces $TI(G)$ es el índice de Sombor.
10. Si $f(x, y) = \frac{1}{\sqrt{x+y}}$ entonces $TI(G)$ es el índice de suma-conectividad.

Observación 2.2.1. *Nótese que el índice de Nirmala (1) es la versión inversa del índice de suma-conectividad. Además, el índice de Nirmala es el índice de suma-conectividad variable, para $a = 1/2$.*

Teorema 2.2.2. *Sea $RSC_n = RSC(n, p_1, p_2, p_3)$ con $n \geq 2$ una cadena aleatoria de espiros. Entonces*

$$\mathbb{E}(TI_n) = TI_2 + \alpha(n - 2),$$

$$\mathbb{V}(TI_n) = (\beta - \alpha^2)(n - 2),$$

donde $i = 1, 2, 3$, $TI_n = TI(RSC_n)$, $TI_{n,i} = TI(RSC_n^i)$, $\alpha_i = TI_{3,i} - TI_2$, $\alpha = \sum_{i=1}^3 \alpha_i p_i$ y $\beta = \sum_{i=1}^3 \alpha_i^2 p_i$.

Demostración. Sea L_n con $n \geq 3$ una variable aleatoria con rango $\{1, 2, 3\}$ tal que $p_i = \mathbb{P}(L_n = i)$ y L_2 denota el enlace inicial, es decir, L_n denota el enlace seleccionado en el tiempo n . Obsérvese que, en el tiempo $n - 1$ tenemos

$$\underbrace{\underbrace{HL_2H L_3HL_4HL_5 \dots L_{n-1}H}_{RSC_2}}_{RSC_{n-1}}$$

Entonces, en el tiempo n , obtenemos

$$\underbrace{\underbrace{\underbrace{HL_2H L_3HL_4HL_5 \dots L_{n-1}H L_nH}_{RSC_2}}_{RSC_{n-1}}}_{RSC_n}$$

Por lo tanto, debemos prestar atención al cambio en el cálculo del índice topológico al unir H con H a través de L_n . Sea $n \geq 3$ y $i = 1, 2, 3$, entonces a partir de este planteamiento, por la definición de cadena aleatoria de espiros y $TI(G)$ en la ecuación

(2.1) obtenemos la siguiente relación recursiva casi segura entre TI_{n-1} y TI_n , condicionada a L_n y \mathbb{F}_{n-1} .

$$TI_{n,L_n} - TI_{n-1} = TI_{3,L_n} - TI_2,$$

donde \mathbb{F}_{n-1} denota el σ -álgebra generada por la historia de la cadena aleatoria espiro en las primeras $n-1$ etapas. Ahora, tomamos la esperanza con respecto a L_n para obtener,

$$\begin{aligned} \mathbb{E}(TI_n | \mathbb{F}_{n-1}) &= \sum_{i=1}^3 (TI_{n-1} + \alpha_i) p_i \\ &= TI_{n-1} + \sum_{i=1}^3 \alpha_i p_i, \end{aligned}$$

donde, $\alpha_i = TI_{3,i} - TI_2$. Entonces, tomando la esperanza, obtenemos una relación de recurrencia para $\mathbb{E}(TI_n)$,

$$\mathbb{E}(TI_n) = \mathbb{E}(TI_{n-1}) + \sum_{i=1}^3 \alpha_i p_i. \quad (2.2)$$

Resolvemos la ecuación (2.2) con el valor inicial $\mathbb{E}(TI_2) = TI_2$ y obtenemos el resultado planteado inicialmente,

$$\mathbb{E}(TI_n) = TI_2 + \alpha(n-2),$$

donde $\alpha = \sum_{i=1}^3 \alpha_i p_i$. La expresión para $\mathbb{E}(TI_n^2)$ se sigue de forma similar,

$$\begin{aligned} \mathbb{E}(TI_n^2 | \mathbb{F}_{n-1}) &= \sum_{i=1}^3 (TI_{n-1} + \alpha_i)^2 p_i \\ &= \sum_{i=1}^3 TI_{n-1}^2 p_i + 2TI_{n-1} \alpha_i p_i + \alpha_i^2 p_i \\ &= TI_{n-1}^2 + 2TI_{n-1} \alpha + \beta, \end{aligned}$$

donde $\beta = \sum_{i=1}^3 \alpha_i^2 p_i$, entonces

$$\begin{aligned} \mathbb{E}(TI_n^2) &= \mathbb{E}(TI_{n-1}^2) + 2\alpha\mathbb{E}(TI_{n-1}) + \beta \\ &= \mathbb{E}(TI_{n-1}^2) + 2\alpha TI_2 + 2\alpha^2(n-3) + \beta, \end{aligned}$$

con $\mathbb{E}(TI_2^2) = TI_2^2$, entonces iterando, se obtiene que

$$\mathbb{E}(TI_n^2) = TI_2^2 + (2\alpha TI_2 + \beta)(n-2) + (n-3)(n-2)\alpha^2.$$

La varianza de TI_n se obtiene inmediatamente tomando la diferencia entre $\mathbb{E}(TI_n^2)$ y $\mathbb{E}(TI_n)^2$,

$$\begin{aligned} \mathbb{V}(TI_n) &= \beta(n-2) + ((n-2)(n-3) - (n-2)^2)\alpha^2 \\ &= (\beta - \alpha^2)(n-2), \end{aligned}$$

demostrando lo requerido. □

Nota que $\beta - \alpha^2 = 0$ si y sólo si $\alpha_1 = \alpha_2 = \alpha_3$ equivalentemente $TI_n = TI_2 + \alpha(n-2)$ c.s. con $n \geq 2$ (una sucesión determinista). Recordemos que c.s. significa casi seguro, es decir, si una propiedad se cumple c.s. entonces el conjunto donde no se cumple está contenido en un conjunto de probabilidad cero. Ahora, con una formulación vía martingala investigaremos el comportamiento asintótico de TI_n cuando $\beta - \alpha^2 > 0$. La idea clave es considerar una transformación M_n y requerimos que las variables aleatorias transformadas formen una martingala en la siguiente proposición.

Proposición 2.2.3. *Para $n \geq 2$, $\{M_n = TI_n - \alpha(n-2)\}_n$ es una martingala con respecto a \mathbb{F}_n .*

Demostración. Primero, observemos que $\mathbb{E}(|M_n|) < +\infty$. Entonces, por Teorema 2.2.2,

$$\begin{aligned}\mathbb{E}(TI_n - \alpha(n-2) \mid \mathbb{F}_{n-1}) &= \mathbb{E}(TI_n \mid \mathbb{F}_{n-1}) - \alpha(n-2) \\ &= TI_{n-1} + \alpha - \alpha(n-2) \\ &= TI_{n-1} - \alpha(n-3).\end{aligned}$$

La demostración es concluida. □

Utilizamos la notación \xrightarrow{D} para denotar la convergencia en la distribución y \xrightarrow{P} para denotar la convergencia en la probabilidad. La variable aleatoria $N(\mu, \sigma^2)$ aparece en el siguiente teorema para denotar a la normal distribuida con media μ y varianza σ^2 .

Teorema 2.2.4. *Cuando $n \rightarrow \infty$,*

$$\frac{TI_n - (n-2)\alpha}{\sqrt{n}} \xrightarrow{D} N(0, \beta - \alpha^2).$$

Demostración. Nota que, para $j \geq 3$ y $i = 1, 2, 3$, tenemos que

$$|\nabla M_j| = |\nabla TI_j - \alpha| \leq 2 \max_i \{|\alpha_i|\},$$

donde $\nabla M_j = M_j - M_{j-1}$ y $\nabla TI_j = TI_j - TI_{j-1}$. Entonces, cuando n tiende a ∞

$$\lim_{n \rightarrow \infty} \frac{|\nabla M_j|}{\sqrt{n}} = 0.$$

Así, dado $\varepsilon > 0$, existe un $n_0(\varepsilon) > 0$ tal que, los conjuntos $\{|\nabla M_j| > \varepsilon\sqrt{n}\}$ son vacíos para todo $n > n_0(\varepsilon)$. Por lo tanto

$$U_n := \frac{1}{n} \sum_{j=3}^n \mathbb{E} \left((\nabla M_j)^2 \mathbb{I}_{\{|\nabla M_j| > \varepsilon\sqrt{n}\}} \mid \mathbb{F}_{j-1} \right),$$

converge a cero casi seguramente y así $U_n \xrightarrow{P} 0$. Como resultado, se verifica la condición de Lindeberg. Por otro lado, la condición de varianza condicional viene dada por

$$V_n := \frac{1}{n} \sum_{j=3}^n \mathbb{E} \left((\nabla M_j)^2 \mid \mathbb{F}_{j-1} \right) \xrightarrow{P} \beta - \alpha^2.$$

Notando que,

$$\begin{aligned} \frac{1}{n} \sum_{j=3}^n \mathbb{E} \left((\nabla M_j)^2 \mid \mathbb{F}_{j-1} \right) &= \frac{1}{n} \sum_{j=3}^n \mathbb{E} \left((\nabla T I_j - \alpha)^2 \mid \mathbb{F}_{j-1} \right) \\ &= \frac{1}{n} \sum_{j=3}^n \sum_{i=1}^3 (\alpha_i - \alpha)^2 p_i \\ &= \frac{n-2}{n} \sum_{i=1}^3 (\alpha_i - \alpha)^2 p_i. \end{aligned}$$

Finalmente, por el Teorema del Límite Central de Martingalas [31], obtenemos el resultado indicado, ya que

$$\begin{aligned} \sum_{i=1}^3 (\alpha_i - \alpha)^2 p_i &= \sum_{i=1}^3 \alpha_i^2 p_i - 2\alpha \sum_{i=1}^3 \alpha_i p_i + \alpha^2 \\ &= \beta - \alpha^2. \end{aligned}$$

□

Entonces podemos utilizar el Teorema 2.2.4 para obtener el siguiente resultado.

Corolario 2.2.5. *Cuando $n \rightarrow \infty$,*

$$\frac{TI_n - \mathbb{E}(TI_n)}{\sqrt{\mathbb{V}(TI_n)}} \xrightarrow{D} N(0, 1).$$

El siguiente teorema ofrece más detalles sobre la distribución de los índices topológicos en las cadenas aleatorias de espiros. Aquí, $M_R(\cdot)$ denota la función generadora de momentos de una variable aleatoria R .

Teorema 2.2.6. *Sea RSC_n con $n \geq 2$ una cadena aleatoria de espiros. Entonces,*

$$TI_n = TI_2 + a^T X,$$

donde $a^T = (\alpha_1, \alpha_2, \alpha_3)$ y $X = (X_1, X_2, X_3)$ es una variable aleatoria multinomial con parámetros $n - 2$ y (p_1, p_2, p_3) .

Demostración. Sea $t \in \mathbb{R}$, observe que,

$$\begin{aligned} \mathbb{E}(e^{tTI_n} \mid \mathbb{F}_{n-1}) &= \sum_{i=1}^3 e^{tTI_{n-1}} e^{t\alpha_i p_i} \\ &= e^{tTI_{n-1}} \sum_{i=1}^3 e^{t\alpha_i p_i}. \end{aligned}$$

Así, podemos concluir que

$$M_{TI_n}(t) = M_{TI_{n-1}}(t) \sum_{i=1}^3 e^{t\alpha_i p_i}.$$

Por lo tanto, podemos escribir,

$$\begin{aligned} M_{TI_n}(t) &= M_{TI_2}(t) \left(\sum_{i=1}^3 e^{t\alpha_i p_i} \right)^{n-2} \\ &= M_{TI_2}(t) M_X(\alpha_1 t, \alpha_2 t, \alpha_3 t) \\ &= M_{TI_2}(t) M_{a^T X}(t), \end{aligned}$$

lo que completa la prueba. □

Es útil observar que la aproximación dada en Corolario 2.2.5 es idéntica a la obtenida por el siguiente método. Sea $n \geq 2$, por el Teorema 2.2.6 tenemos que

$$TI_n = TI_2 + a^T X.$$

Del Teorema Central del Límite [63], se deduce para el caso de vectores aleatorios que X se distribuye asintóticamente según una distribución normal multivariante con media $\mathbb{E}(X)$ y matriz de covarianza $V(X)$. En consecuencia, $a^T X$ se distribuye asintóticamente según una distribución normal con media $a^T \mathbb{E}(X)$ y varianza $a^T V(X)a$. Entonces, a medida que n va a ∞

$$\frac{TI_n - TI_2 - a^T \mathbb{E}(X)}{\sqrt{a^T V(X)a}} = \frac{TI_n - \mathbb{E}(TI_n)}{\sqrt{\mathbb{V}(TI_n)}} \xrightarrow{D} N(0, 1).$$

2.3. Interpretación de los resultados y ejemplos

A modo de resumen, sobre lo visto en la sección anterior, tenemos el siguiente resultado.

Teorema 2.3.1. *Sea RSC_n con $n \geq 2$ una cadena aleatoria espiro. Entonces*

$$TI_n = A + BX + Cn,$$

$$\mathbb{E}(TI_n) = A + (Bp_1 + C)n - 2Bp_1,$$

$$\mathbb{V}(TI_n) = B^2 p_1(1 - p_1)(n - 2),$$

$$\frac{TI_n - \mathbb{E}(TI_n)}{\sqrt{\mathbb{V}(TI_n)}} \xrightarrow{D} N(0, 1),$$

donde $A = 4f(2, 2) - 4f(2, 4)$, $B = f(2, 2) - 2f(2, 4) + f(4, 4)$, $C = 2f(2, 2) + 4f(2, 4)$ y X tiene una distribución binomial con parámetros $n - 2$ y p_1 .

Observación 2.3.2. 1. *Vía la definición, se verifica que $\alpha_2 = \alpha_3 = 2f(2, 2) + 4f(2, 4)$ y $\alpha_1 = 3f(2, 2) + 2f(2, 4) + f(4, 4)$.*

2. *Sea $G = (V(G), E(G))$ un grafo finito, simple y conectado y $h : \{1, 2, \dots\} \rightarrow (0, \infty)$, debido a la siguiente identidad [13]*

$$\sum_{vu \in E(G)} \frac{h(d_v)}{d_v} + \frac{h(d_u)}{d_u} = \sum_{v \in V(G)} h(d_v).$$

Tenemos que, si

$$\begin{aligned} TI(RSC_n) &= \sum_{v \in V(RSC_n)} h(d_v) \\ &= \sum_{vu \in E(RSC_n)} \frac{h(d_v)}{d_v} + \frac{h(d_u)}{d_u}. \end{aligned}$$

entonces $TI(RSC_n) = TI_2 + \alpha(n - 2)$ (una sucesión determinista), ya que $\alpha_1 = \alpha_2$.

Ahora, buscando aplicar el Teorema 2.3.1, presentamos los siguientes corolarios.

Corolario 2.3.3. *Sea $RSC_n = RSC(n, p_1, p_2, p_3)$ una cadena aleatoria espiro y N_n el índice de Nirmala de una RSC_n , con $n \geq 2$. Entonces*

$$N_n = 8 - 4\sqrt{6} + (2 - 2\sqrt{6} + 2\sqrt{2})X + (4 + 4\sqrt{6})n,$$

$$\mathbb{E}(N_n) = 8 - 4\sqrt{6} + \left((2 - 2\sqrt{6} + 2\sqrt{2})p_1 + 4 + 4\sqrt{6} \right) n - 2(2 - 2\sqrt{6} + 2\sqrt{2})p_1,$$

$$\mathbb{V}(N_n) = (2 - 2\sqrt{6} + 2\sqrt{2})^2 p_1 (1 - p_1) (n - 2),$$

$$\frac{N_n - \mathbb{E}(N_n)}{\sqrt{\mathbb{V}(N_n)}} \xrightarrow{D} N(0, 1),$$

donde X tiene una distribución binomial con parámetros $n - 2$ y p_1 .

Corolario 2.3.4. *Sea $RSC_n = RSC(n, p_1, p_2, p_3)$ una cadena aleatoria espiro y $M1_n$ el primer índice de Zagreb de una RSC_n , con $n \geq 2$. Entonces*

$$M1_n = 32n - 8. \tag{2.3}$$

Corolario 2.3.5. *Sea $RSC_n = RSC(n, p_1, p_2, p_3)$ una cadena aleatoria espiro y R_n el índice de Randić de una RSC_n , con $n \geq 2$. Entonces*

$$R_n = 2 - \sqrt{2} + (3/4 - \sqrt{2}/2)X + (1 + \sqrt{2})n,$$

$$\mathbb{E}(R_n) = 2 - \sqrt{2} + \left((3/4 - \sqrt{2}/2)p_1 + 1 + \sqrt{2} \right) n + (\sqrt{2} - 3/2)p_1, \quad (2.4)$$

$$\mathbb{V}(R_n) = (3/4 - \sqrt{2}/2)^2 p_1 (1 - p_1) (n - 2),$$

$$\frac{R_n - \mathbb{E}(R_n)}{\sqrt{\mathbb{V}(R_n)}} \xrightarrow{D} N(0, 1),$$

donde X tiene una distribución binomial con parámetros $n - 2$ y p_1 .

Corolario 2.3.6. *Sea $RSC_n = RSC(n, p_1, p_2, p_3)$ una cadena aleatoria espiro y S_n el índice de Sombor de una RSC_n , con $n \geq 2$. Entonces*

$$S_n = 8\sqrt{2} - 8\sqrt{5} + (6\sqrt{2} - 4\sqrt{5})X + (4\sqrt{2} + 8\sqrt{5})n,$$

$$\mathbb{E}(S_n) = 8\sqrt{2} - 8\sqrt{5} + \left((6\sqrt{2} - 4\sqrt{5})p_1 + 4\sqrt{2} + 8\sqrt{5} \right) n - 2(6\sqrt{2} - 4\sqrt{5})p_1,$$

$$\mathbb{V}(S_n) = (6\sqrt{2} - 4\sqrt{5})^2 p_1 (1 - p_1) (n - 2),$$

$$\frac{S_n - \mathbb{E}(S_n)}{\sqrt{\mathbb{V}(S_n)}} \xrightarrow{D} N(0, 1),$$

donde X tiene una distribución binomial con parámetros $n - 2$ y p_1 .

Corolario 2.3.7. *Sea $RSC_n = RSC(n, p_1, p_2, p_3)$ una cadena aleatoria espiro y $M2_n$ el segundo índice de Zagreb de una RSC_n , con $n \geq 2$. Entonces*

$$M2_n = 4X + 40n - 16,$$

$$\mathbb{E}(M2_n) = (4p_1 + 40)n - 8p_1 - 16, \quad (2.5)$$

$$\mathbb{V}(M2_n) = 16p_1(1 - p_1)(n - 2),$$

$$\frac{M2_n - \mathbb{E}(M2_n)}{\sqrt{\mathbb{V}(M2_n)}} \xrightarrow{D} N(0, 1),$$

donde X tiene una distribución binomial con parámetros $n - 2$ y p_1 .

Observación 2.3.8. *De hecho, podemos ver que (2.3) y (2.4) se obtienen en [35]. También, podemos ver que el Corolario 2.3.6 y (2.5) se obtienen en [81] y [58], respectivamente.*

Observación 2.3.9. *Para $n \geq 2$ y $p_1 \in (0, 1)$, se deducen de los Corolarios 2.3.3, 2.3.4, 2.3.5, 2.3.6 y 2.3.7 que*

$$\mathbb{E}(R_n) \leq \mathbb{E}(N_n) \leq \mathbb{E}(S_n) \leq \mathbb{E}(M1_n) \leq \mathbb{E}(M2_n) \text{ (ver Figura 2.3),}$$

$$\mathbb{V}(R_n) \leq \mathbb{V}(N_n) \leq \mathbb{V}(S_n) \leq \mathbb{V}(M2_n).$$

Finalmente, usando el lenguaje Python, realizamos un experimento numérico para apoyar los comportamientos asintóticos desarrollados en los Corolarios 2.3.3, 2.3.5, 2.3.6 y 2.3.7. Dado un $p_1 \in (0, 1)$, vía el siguiente pseudocódigo generamos de forma independiente 5 000 réplicas de una cadena aleatoria de espiros después de $n = 10,000$ pasos evolutivos.

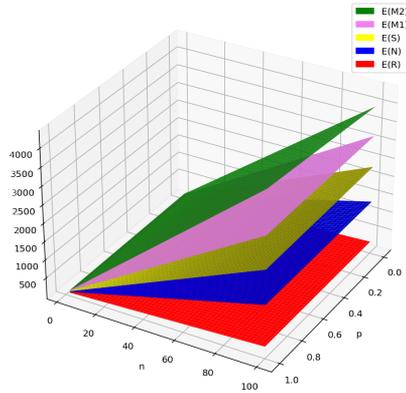


Figura 2.3: Diferencia entre $\mathbb{E}(R_n)$, $\mathbb{E}(N_n)$, $\mathbb{E}(S_n)$, $\mathbb{E}(M1_n)$ y $\mathbb{E}(M2_n)$.

Algorithm 2 Generar *RSCs*.

Input: $p_1 \in (0, 1)$: parámetro del modelo, M : número de *RSCs*,
 n : pasos evolutivos de las *RSCs*.

Output: Una muestra de M *RSCs* con parámetro p_1 al tiempo n .

$Sample = \emptyset$

for $i = 1$ **to** M **do**

 Simular una $Bin(n - 2, p_1)$.

$X = Bin(n - 2, p_1)$.

$Sample = Sample \cup \{X\}$.

end for

Para cada cadena simulada, ya que conocemos la información de interés (variable aleatoria X), se calcula su índice topológico. Los histogramas de los datos de la muestra con una curva de aproximación normal se dan en las Figuras 2.4, 2.5, 2.6 y 2.7.

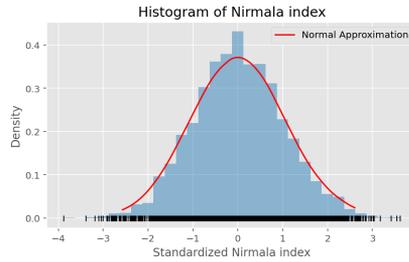


Figura 2.4: Histograma del índice de Nirmala estandarizado; la curva roja gruesa es la densidad estimada de la muestra.

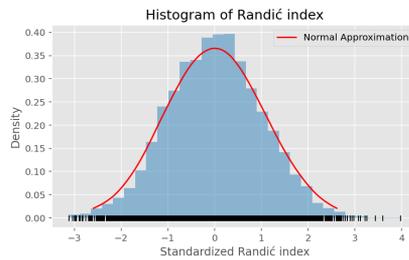


Figura 2.5: Histograma del índice de Randić estandarizado; la curva roja gruesa es la densidad estimada de la muestra.

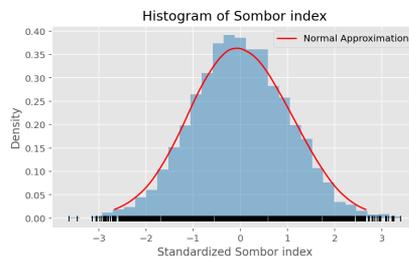


Figura 2.6: Histograma del índice de Sombor estandarizado; la curva roja gruesa es la densidad estimada de la muestra.

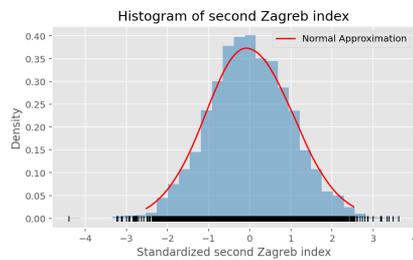


Figura 2.7: Histograma del segundo índice de Zagreb estandarizado; la curva roja gruesa es la densidad estimada de la muestra.

Capítulo 3

Cadenas aleatorias de poliominós

3.1. Preliminares

De forma natural, nos preguntamos si el tratamiento discutido anteriormente puede realizarse en otras cadenas aleatorias. Así, de forma similar, en este capítulo, estudiaremos los índices topológicos basados en grados, ahora, sobre cadenas aleatorias de poliominós.

Un sistema de poliominós es un grafo plano finito 2-conectado tal que cada cara interior (digamos una celda) está rodeada por un cuadrado regular de longitud uno. En un sistema de poliominós, se dice que dos cuadrados son adyacentes si comparten un lado. Una cadena de poliominós es un sistema de poliominós en el que la unión de los centros de sus celdas adyacentes forma un camino $c_1c_2 \dots c_n$, donde c_i es el centro de la celda i -ésima. Por lo tanto, en una cadena de poliominós cada cuadrado es adyacente a lo sumo con otros dos cuadrados. Si un cuadrado tiene sólo un cuadrado adyacente, se llama terminal, si tiene dos cuadrados adyacentes que no tienen ningún vértice de grado 2, se llama medial, y si tiene dos cuadrados adyacentes de forma tal que tiene un vértice de grado 2, se denomina kink. Una cadena de poliominós sin kinks se llama cadena lineal Li_n . Una cadena de poliominós formada sólo por kinks y cuadrados terminales se denomina cadena en zigzag Z_n (ver Figura 3.1). Una cadena lineal máxima (que contiene los

cuadrados terminales y los kinks en su extremo) en las cadenas de poliomino se denomina segmento de la cadena de poliomino.

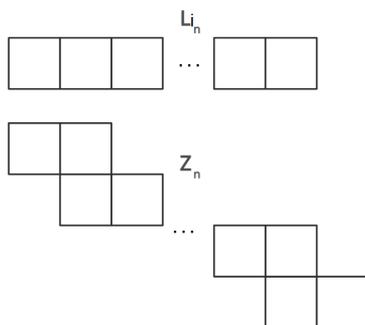


Figura 3.1: La cadena lineal y la cadena zigzag.

El nombre de poliomino fue introducido en 1953 por analogía con el dominó por Solomon W. Golomb [23] y desde entonces los sistemas de poliomino han sido ampliamente estudiados, por ejemplo, en química orgánica, especialmente en compuestos aromáticos policíclicos. En la actualidad, los trabajos recientes sobre las cadenas de poliomino incluyen los emparejamientos perfectos [47, 82], la búsqueda de fórmulas para calcular varios índices topológicos [9, 34, 48, 54] y los problemas extremos [2, 3, 79]. En concreto, las cadenas aleatorias de poliomino han atraído una gran atención de los investigadores en los últimos años [44, 72, 73, 75, 76].

Una cadena aleatoria de poliomino ($RPC_n = RPC(n, p_1, p_2)$) puede construirse de la siguiente manera: para $n = 1$ y $n = 2$, los RPC_n se muestran en la Figura 3.2. Para $n \geq 3$, se puede adjuntar un nuevo cuadrado de dos maneras, lo que resulta en RPC_n^1 y RPC_n^2 con probabilidad p_1 y p_2 respectivamente, donde $0 < p_1, p_2 < 1$ y $p_1 + p_2 = 1$, véase la Figura 3.3. Para una cadena aleatoria de poliomino en el tiempo n , el valor de un índice topológico es una variable aleatoria. Teniendo en cuenta los argumentos expuestos en los párrafos anteriores y utilizando un enfoque similar al realizado en el capítulo anterior, en lo siguiente estudiaremos los índices topológicos basados en grado de una cadena aleatoria de poliomino. Finalmente, es importante men-

cionar que lo desarrollado en esta sección se encuentra basado en el artículo [68].



Figura 3.2: Los grafos de RPC_1 y RPC_2 .

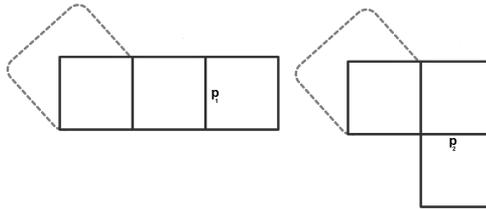


Figura 3.3: Los dos links para $RPC_n (n \geq 3)$.

3.2. Índices topológicos sobre cadenas aleatorias de poliominós

En esta sección, exponemos y demostramos nuestros resultados principales. En primer lugar, L_n denota el enlace seleccionado en el momento $n \geq 3$, es decir, L_n denota una variable aleatoria con rango $\{1, 2\}$ donde $p_i = \mathbb{P}(L_n = i)$. Para $i, j \in \{1, 2\}$, $TI_n = TI(RPC_n)$, RPC_n^i denota una cadena aleatoria de poliominós en el tiempo $n \geq 3$ tal que $L_n = i$, $TI_{n,i} = TI(RPC_n^i)$, $RPC_n^{j,i}$ denota una cadena de poliominós aleatoria en el tiempo $n \geq 4$ tal que $L_{n-1} = j$ y $L_n = i$, $TI_{n,j,i} = TI(RPC_n^{j,i})$, $\alpha_{j,i} = TI_{4,j,i} - TI_{3,j}$, $\alpha_i = TI_{3,i} - TI_2$, $\alpha = \sum_{j=1}^2 \sum_{i=1}^2 \alpha_{j,i} p_j p_i$ y $\beta = \sum_{j=1}^2 \sum_{i=1}^2 \alpha_{j,i}^2 p_j p_i$.

Observación 3.2.1. *Nota que, por definición:*

1. $\alpha_{1,1} = \alpha_1 = 3f(3, 3)$,
2. $\alpha_{1,2} = 3f(3, 4) + f(2, 4) + f(2, 3) - 2f(3, 3)$,

$$3. \alpha_{2,1} = f(3, 4) - f(2, 4) + f(2, 3) + 2f(3, 3),$$

$$4. \alpha_{2,2} = f(4, 4) + 2f(2, 4),$$

$$5. \alpha_2 = 2f(3, 4) + 2f(2, 4) - f(3, 3).$$

Entonces, $\alpha_{2,1} - \alpha_{1,2} = \alpha_{1,1} - \alpha_2$. En particular, cuando $f(x, y) = x^a + y^a$ con $a \in \mathbb{R}$ y $x, y \in \{1, 2, \dots\}$ las siguientes condiciones se verifican:

$$1. \alpha_{1,1} = \alpha_{2,1},$$

$$2. \alpha_{2,2} = \alpha_{1,2} = \alpha_2.$$

Además, en este caso, $TI_n = \sum_{v \in V(RPC_n)} (d_v)^{a+1}$, debido a la siguiente identidad

$$\sum_{vu \in E(G)} (d_v)^a + (d_u)^a = \sum_{v \in V(G)} (d_v)^{a+1},$$

se puede consultar la validez de la expresión anterior, por ejemplo, en [13].

Teorema 3.2.2. Sea $RPC_n = RPC(n, p_1, p_2)$ una cadena aleatoria de poliomínos, entonces para $n \geq 3$

$$\mathbb{E}(TI_n) = \mathbb{E}(TI_3) + \alpha(n - 3),$$

$$\mathbb{V}(TI_n) = \mathbb{V}(TI_3) + (\beta - \alpha^2)(n - 3),$$

donde

$$\mathbb{E}(TI_3) = TI_2 + \sum_{i=1}^2 \alpha_i p_i,$$

$$\mathbb{V}(TI_3) = \sum_{i=1}^2 \alpha_i^2 p_i - \left(\sum_{i=1}^2 \alpha_i p_i \right)^2.$$

Demostración. Para $n \geq 4$, se deduce de la definición de una cadena aleatoria de poliominós y por la definición de $TI(G)$ en la ecuación (2.1) la siguiente relación recursiva casi segura de TI_n condicionada a \mathbb{F}_{n-1} y al vector aleatorio (L_{n-1}, L_n)

$$TI_{n,L_{n-1},L_n} - TI_{n-1} = TI_{4,L_{n-1},L_n} - TI_{3,L_{n-1}},$$

donde \mathbb{F}_{n-1} denota la σ -álgebra generada por la historia de la cadena aleatoria de poliominós en las primeras $n-1$ etapas. Ahora para $n \geq 4$, tomamos la esperanza con respecto a (L_{n-1}, L_n) para obtener

$$\begin{aligned} \mathbb{E}(TI_n | \mathbb{F}_{n-1}) &= \sum_{j=1}^2 \sum_{i=1}^2 (TI_{n-1} + \alpha_{j,i}) p_j p_i \\ &= TI_{n-1} + \sum_{j=1}^2 \sum_{i=1}^2 \alpha_{j,i} p_j p_i, \end{aligned}$$

donde, $\alpha_{j,i} = TI_{4,j,i} - TI_{3,j}$. Entonces, tomando esperanza en ambos miembros, obtenemos una relación de recurrencia para $\mathbb{E}(TI_n)$ con $n \geq 4$,

$$\mathbb{E}(TI_n) = \mathbb{E}(TI_{n-1}) + \sum_{j=1}^2 \sum_{i=1}^2 \alpha_{j,i} p_j p_i. \quad (3.1)$$

Resolvemos la ecuación (3.1) con el valor inicial $\mathbb{E}(TI_3)$ y obtenemos el resultado planteado, es decir,

$$\mathbb{E}(TI_n) = \mathbb{E}(TI_3) + \alpha(n-3),$$

donde $\alpha = \sum_{j=1}^2 \sum_{i=1}^2 \alpha_{j,i} p_j p_i$. Dado $n \geq 4$, la expresión para $\mathbb{E}(TI_n^2)$ se sigue de forma similar,

$$\begin{aligned}
 \mathbb{E}(TI_n^2 \mid \mathbb{F}_{n-1}) &= \sum_{j=1}^2 \sum_{i=1}^2 (TI_{n-1} + \alpha_{j,i})^2 p_j p_i \\
 &= \sum_{j=1}^2 \sum_{i=1}^2 TI_{n-1}^2 p_j p_i + 2TI_{n-1} \alpha_{j,i} p_j p_i + \alpha_{j,i}^2 p_j p_i \\
 &= TI_{n-1}^2 + 2\alpha TI_{n-1} + \beta,
 \end{aligned}$$

donde $\beta = \sum_{j=1}^2 \sum_{i=1}^2 \alpha_{j,i}^2 p_j p_i$, así

$$\begin{aligned}
 \mathbb{E}(TI_n^2) &= \mathbb{E}(TI_{n-1}^2) + 2\alpha \mathbb{E}(TI_{n-1}) + \beta \\
 &= \mathbb{E}(TI_{n-1}^2) + 2\alpha \mathbb{E}(TI_3) + 2\alpha^2(n-4) + \beta,
 \end{aligned}$$

entonces iterando, para $n \geq 3$ se obtiene que

$$\mathbb{E}(TI_n^2) = \mathbb{E}(TI_3^2) + (2\alpha \mathbb{E}(TI_3) + \beta)(n-3) + \alpha^2(n-3)(n-4).$$

Dado $n \geq 3$, la varianza de TI_n se obtiene inmediatamente tomando la diferencia entre $\mathbb{E}(TI_n^2)$ y $\mathbb{E}(TI_n)^2$,

$$\begin{aligned}
 \mathbb{V}(TI_n) &= \mathbb{V}(TI_3) + \beta(n-3) + ((n-3)(n-4) - (n-3)^2) \alpha^2 \\
 &= \mathbb{V}(TI_3) + (\beta - \alpha^2)(n-3).
 \end{aligned}$$

Finalmente, observe que

$$\begin{aligned}
 \mathbb{E}(TI_3) &= \mathbb{E}(\mathbb{E}(TI_3 \mid L_3)) \\
 &= \sum_{i=1}^2 (TI_2 + \alpha_i) p_i \\
 &= TI_2 + \sum_{i=1}^2 \alpha_i p_i,
 \end{aligned}$$

donde $\alpha_i = TI_{3,i} - TI_2$. Del mismo modo, tenemos

$$\mathbb{V}(TI_3) = \sum_{i=1}^2 \alpha_i^2 p_i - \left(\sum_{i=1}^2 \alpha_i p_i \right)^2,$$

demostrando el resultado. □

Obsérvese que las siguientes afirmaciones son equivalentes

1. $\beta - \alpha^2 = 0$.
2. $\alpha_{1,1} = \alpha_{1,2} = \alpha_{2,2} = \alpha_{2,1}$.
3. Para $n \geq 2$, $TI_n = TI_2 + \alpha(n - 2)$ casi seguramente.
4. $f(3, 4) = (2f(4, 4) + f(2, 4))/3$, $f(3, 3) = (f(4, 4) + 2f(2, 4))/3$
y $f(2, 3) = (-f(4, 4) + 4f(2, 4))/3$.

En consecuencia, cuando $\beta - \alpha^2 = 0$, $\frac{TI_n}{n}$ converge casi seguramente a α a medida que $n \rightarrow \infty$. Cabe destacar que utilizando las equivalencias anteriores podemos concluir que, TI_n es una sucesión determinista casi seguramente si y sólo si $\alpha_{1,1} = \alpha_{1,2} = \alpha_{2,2} = \alpha_{2,1}$. Por lo tanto, por la Observación 3.2.1 si $f(x, y) = x^a + y^a$ con $a \in \mathbb{R}$ tenemos que TI_n es una sucesión determinista casi seguramente si y sólo si $2 \cdot 3^{a+1} = 4^{a+1} + 2^{a+1}$, $a \in \mathbb{R}$. La última ecuación tiene dos soluciones únicas $a = 0, -1$, ya que para $a \in (-1, 0)$, x^{a+1} es una función estrictamente cóncava sobre \mathbb{R}^+ por lo que $(\frac{4+2}{2})^{a+1} > \frac{4^{a+1} + 2^{a+1}}{2}$ y para $a > 0$ o $a < -1$, x^{a+1} es una función estrictamente convexa en \mathbb{R}^+ por lo que $(\frac{4+2}{2})^{a+1} < \frac{4^{a+1} + 2^{a+1}}{2}$. Por lo tanto, TI_n es una sucesión determinista casi seguramente si y sólo si $a \in \{0, -1\}$. Este hecho tiene sentido ya que

$$\begin{aligned} \sum_{vu \in E(RPC_n)} (d_v)^0 + (d_u)^0 &= \sum_{v \in V(RPC_n)} (d_v)^1 \\ &= 2|E(RPC_n)| \\ &= 2 + 6n. \end{aligned}$$

y

$$\begin{aligned} \sum_{vu \in E(RPC_n)} (d_v)^{-1} + (d_u)^{-1} &= \sum_{v \in V(RPC_n)} (d_v)^0 \\ &= |V(RPC_n)| \\ &= 2 + 2n. \end{aligned}$$

Ahora, una formulación de martingalas nos permitirá estudiar el comportamiento asintótico de TI_n cuando $\beta - \alpha^2 > 0$.

Proposición 3.2.3. *Para $n \geq 3$, $\{M_n = TI_n - \alpha(n - 3)\}_n$ es una martingala con respecto a \mathbb{F}_n .*

Demostración. Observa que $\mathbb{E}(|M_n|) < +\infty$. Para $n \geq 4$, por Teorema 3.2.2,

$$\begin{aligned} \mathbb{E}(M_n | \mathbb{F}_{n-1}) &= \mathbb{E}(TI_n - \alpha(n - 3) | \mathbb{F}_{n-1}) \\ &= \mathbb{E}(TI_n | \mathbb{F}_{n-1}) - \alpha(n - 3) \\ &= TI_{n-1} + \alpha - \alpha(n - 3) \\ &= TI_{n-1} - \alpha(n - 4) \\ &= M_{n-1}. \end{aligned}$$

La demostración es completada. □

Utilizamos la notación \xrightarrow{D} para denotar la convergencia en la distribución y \xrightarrow{P} para denotar la convergencia en la probabilidad. Aquí, $N(\mu, \sigma^2)$ denota una variable aleatoria con distribución normal con media μ y varianza σ^2 .

Teorema 3.2.4. *Cuando $n \rightarrow \infty$,*

$$\frac{TI_n - (n-3)\alpha}{\sqrt{n}} \xrightarrow{D} N(0, \beta - \alpha^2).$$

Demostración. Para $k \geq 4$ y $j, i \in \{1, 2\}$, tenemos que

$$|\nabla M_k| = |\nabla TI_k - \alpha| \leq 2 \max_{(j,i)} \{|\alpha_{j,i}|\},$$

donde $\nabla M_k = M_k - M_{k-1}$ y $\nabla TI_k = TI_k - TI_{k-1}$. Esto es, dado $\varepsilon > 0$, existe $N_0(\varepsilon) > 0$ tal que, los conjuntos $\{|\nabla M_k| > \varepsilon\sqrt{n}\}$ son vacíos para todo $n > N_0(\varepsilon)$. En lo que sigue, concluimos que

$$U_n := \frac{1}{n} \sum_{k=4}^n \mathbb{E} \left((\nabla M_k)^2 \mathbb{I}_{\{|\nabla M_k| > \varepsilon\sqrt{n}\}} | \mathbb{F}_{k-1} \right),$$

converge a cero casi seguramente, de ahí, $U_n \xrightarrow{P} 0$. Entonces, se verifica la condición de Lindeberg. Por otro lado, la condición de varianza condicional viene dada por

$$V_n := \frac{1}{n} \sum_{k=4}^n \mathbb{E} \left((\nabla M_k)^2 \mid \mathbb{F}_{k-1} \right) \xrightarrow{P} \beta - \alpha^2.$$

Ya que,

$$\begin{aligned} \frac{1}{n} \sum_{k=4}^n \mathbb{E} \left((\nabla M_k)^2 \mid \mathbb{F}_{k-1} \right) &= \frac{1}{n} \sum_{k=4}^n \mathbb{E} \left((\nabla T I_k - \alpha)^2 \mid \mathbb{F}_{k-1} \right) \\ &= \frac{1}{n} \sum_{k=4}^n \sum_{j=1}^2 \sum_{i=1}^2 (\alpha_{j,i} - \alpha)^2 p_j p_i \\ &= \frac{n-3}{n} \sum_{j=1}^2 \sum_{i=1}^2 (\alpha_{j,i} - \alpha)^2 p_j p_i. \end{aligned}$$

Por lo tanto, por el Teorema del Límite Central de Martingalas [31], obtenemos el resultado indicado. \square

Por último, para aplicar los resultados obtenidos en este capítulo, calculamos el valor esperado de varios índices topológicos importantes para una cadena aleatoria de poliominós (ver la Tabla 3.1).

3.3. Cadena de poliominós

En esta sección, el objetivo es obtener expresiones analíticas explícitas para calcular $TI(PC_n)$ donde PC_n es una cadena de poliominós con n cuadrados. Sea $m \geq 1$ y $i \in \{1, 2, 3, 4, \dots, m\}$ nótese que una cadena de poliominós PC_n está formada por una secuencia de segmentos $s_1, s_2, s_3, \dots, s_m$ (ver la Figura 3.4) con longitudes $l(s_i) = l_i$ tales que $\sum_{i=1}^m l_i = n + m - 1$, donde l_i indica el número de cuadrados en s_i .

Teorema 3.3.1. *Sea PC_n una cadena de poliominós que tiene $n \geq 3$ cuadrados y $m \geq 1$ segmento(s) s_i con $i = 1, 2, \dots, m$. Entonces*

TI	A	B	C	D	E
<i>first Zagreb index</i>	0	-2	20	-2	-6
<i>second Zagreb index</i>	-1	-4	32	-4	-24
<i>first hyper-Zagreb index</i>	-2	-26	136	-26	-106
<i>second hyper-Zagreb index</i>	-21	-120	384	-92	-560
<i>modified first Zagreb index</i>	0	-13/144	5/16	-13/144	43/72
<i>Albertson index</i>	-2	-2	4	-6	-2
<i>extended index</i>	1/6	-2/3	7/2	-7/12	5/12
<i>sigma index</i>	2	-10	8	-10	-10
<i>Sombor index</i>	395/3349	-225/113	2599/178	-2102/1065	-2108/441
<i>Randić index</i>	-34/2413	184/3229	1138/1189	224/4583	1338/1279
<i>reciprocal Randić index</i>	-255/2588	-426/763	985/102	-509/870	-665/257
<i>sum-connectivity index</i>	-33/2872	449/6784	461/394	382/6307	1092/1283
<i>reciprocal sum-connectivity index</i>	302/14565	-731/1829	5216/675	-42/107	-59/1243
<i>harmonic index</i>	-11/420	23/210	11/12	2/21	457/420
<i>atom-bond-connectivity (ABC) index</i>	183/6023	-432/7583	991/489	-130/3373	691/796
<i>augmented Zagreb index</i>	-1636/757	2399/1751	944/27	515/269	-1814/137
<i>forgotten index</i>	0	-18	72	-18	-58
<i>geometric-arithmetic index</i>	-307/9318	353/2396	883/306	380/2817	637/565
<i>arithmetic-geometric index</i>	413/10692	-365/2282	3499/1121	-976/6871	969/1126
<i>inverse sum indeg index</i>	-19/210	-8/105	14/3	-2/21	-116/105

Tabla 3.1: La información de interés asociada a cada índice topológico: $\mathbb{E}(TI_n) = (Ap_1^2 + Bp_1 + C)n - 3Ap_1^2 + (D - 3B)p_1 + E$.

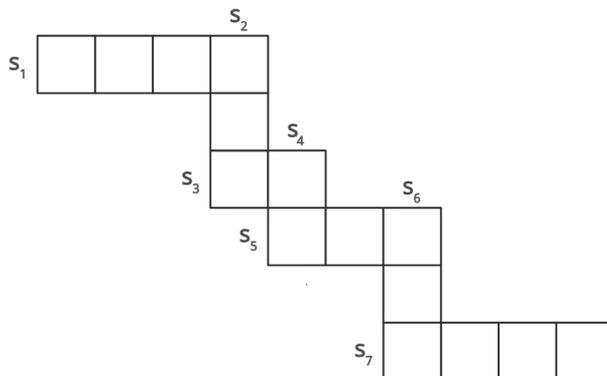


Figura 3.4: Segmentos de una cadena de poliomínos.

$$\begin{aligned}
 TI(PC_n) = & 3f(3,3)n + (4f(3,4) + 2f(2,3) - 6f(3,3))m \\
 & + (f(2,4) - f(2,3) + f(3,3) - f(3,4))(I_1 + I_m) \\
 & + (f(4,4) + 2f(2,4) - 4f(3,4) - 2f(2,3) + 3f(3,3))\gamma \\
 & + 2f(2,2) + 2f(2,3) + f(3,3) - 4f(3,4),
 \end{aligned}$$

$$\text{donde, } I_i = \begin{cases} 1 & \text{si } l_i = 2 \\ 0 & \text{si } l_i \neq 2 \end{cases} \quad \text{y } \gamma = \sum_{i=2}^{m-1} I_i.$$

Demostración. Obsérvese que PC_n es una realización de RPC_n , por lo que conocemos el valor de L_k para $k = 3, \dots, n$. Así, utilizando las ideas presentadas en la sección anterior tenemos,

$$\begin{aligned}
 TI(PC_n) &= TI_2 + \alpha_2 I_1 + \alpha_{1,1}(1 - I_1) + \sum_{k=4}^n \alpha_{L_{k-1}, L_k} \\
 &= TI_2 + (\alpha_2 - \alpha_{1,1})I_1 + \alpha_1 + \sum_{j=1}^2 \sum_{i=1}^2 X_{j,i} \alpha_{j,i},
 \end{aligned}$$

donde $X_{j,i} = |\{k \in \{4, \dots, n\} \mid L_{k-1} = j \text{ y } L_k = i \text{ en } PC_n\}|$ y $I_1 = I_{\{l_1=2\}}$. Ahora, si al tiempo k ($3 \leq k \leq n$), $L_k = 2$ entonces el último segmento en PC_{k-1} está terminado (por lo que se inicia un nuevo segmento en PC_k) y si al tiempo k , $L_k = 1$ entonces se añade un cuadrado al último segmento en PC_{k-1} . Así, $X_2 = |\{k \in \{3, 4, \dots, n\} \mid L_k = 2 \text{ en } PC_n\}| = m - 1$ y $X_1 = |\{k \in \{3, 4, \dots, n\} \mid L_k = 1 \text{ en } PC_n\}| = n - 2 - (m - 1) = n - m - 1$. Además, $X_{1,2} = |\{i \in \{1, 2, \dots, m-1\} \mid l_i \neq 2 \text{ in } PC_n\}|$ y $X_{2,1} = |\{i \in \{2, 3, \dots, m\} \mid l_i \neq 2 \text{ in } PC_n\}|$. Podemos escribir esto como: $X_{1,2} = m - \gamma - 1 - I_1$ y $X_{2,1} = m - \gamma - 1 - I_m$, donde,

$$I_i = \begin{cases} 1 & \text{si } l_i = 2 \\ 0 & \text{si } l_i \neq 2 \end{cases} \quad \text{y } \gamma = \sum_{i=2}^{m-1} I_i.$$

Consecuentemente, $X_{1,1} = n - 2m + \gamma - 1 + I_1 + I_m$ y $X_{2,2} = \gamma$, debido a las siguientes identidades

$$X_{1,1} + X_{2,1} = X_1 - 1 + I_1,$$

$$X_{2,2} + X_{1,2} = X_2 - I_1.$$

Finalmente, llegamos al resultado deseado sustituyendo los valores de $X_{j,i}$, $\alpha_{j,i}$ y α_2 . \square

Observación 3.3.2. 1. Usando que $\sum_{i=1}^m l_i = n + m - 1$, se verifica lo siguiente, $X_{1,1} = \sum_{l_i \neq 2} (l_i - 3)$.

2. Por otro lado, por definición si $f(x, y) = x^a + y^a$ con $a \in \mathbb{R}$ entonces los coeficientes de γ y $I_1 + I_m$ en Teorema 3.3.1 son cero y el coeficiente de m es cero, i.e., la expresión general mostrada en el Teorema 3.3.1 es independiente de m si y sólo si $2 \cdot 3^{a+1} = 4^{a+1} + 2^{a+1}$, equivalentemente $a \in \{0, -1\}$.

3. Por último, en [4] los autores establecieron una expresión general para calcular los BID índices de una cadena de poliominós; lo que se deduce del Teorema 3.3.1. Los BID índices forman una subclase de la clase de todos los índices topológicos basados en grado.

Por definición si $PC_n = Li_n$, deducimos que $m = 1$ y $l_1 = n$ y si $PC_n = Z_n$, entonces $m = n - 1$ y $l_i = 2$ para $i = 1, 2, \dots, m$. Por lo tanto, el siguiente corolario se puede obtener directamente usando el Teorema 3.3.1.

Corolario 3.3.3. Sean Li_n y Z_n cadenas lineales y en zigzag, respectivamente, con $n \geq 3$ cuadrados. Entonces

$$TI(Li_n) = 3f(3, 3)n + 4f(2, 3) + 2f(2, 2) - 5f(3, 3),$$

$$TI(Z_n) = (2f(2, 4) + f(4, 4))n + 4f(2, 3) - 3f(4, 4) + 2f(3, 4) - 4f(2, 4) + 2f(2, 2).$$

Cabe destacar que en 2020, Buragohain et al. [8] introdujeron un índice topológico generalizado para algunas estructuras químicas definido como

$$ISI_{(\alpha, \beta)}(G) = \sum_{uv \in E(G)} (d(u)d(v))^{\alpha} (d(u) + d(v))^{\beta}.$$

En [54] los autores estudiaron el índice generalizado $ISI_{(\alpha,\beta)}$ y el índice (α, β) -Zagreb de una cadena lineal. Utilizando el Corolario 3.3.3 se pueden obtener los resultados mostrados en [54]. Además, tomando $f(x, y) = x^2 + y^2$ en la ecuación (2.1), se obtiene el índice Olvidado. Recientemente, en [9] se mostró el siguiente resultado:

Corolario 3.3.4. *Sean $n \geq 2$ y PC_n una cadena de poliominós con $m \geq 1$ segmento(s). Entonces $F(PC_n) = 54n + 18m - 40$.*

Nótese que, la expresión general obtenida en el Corolario 3.3.4 es independiente de γ, I_1 y I_m ; lo cual tiene sentido debido a la Observación 3.3.2. De manera similar, podemos obtener el resultado anterior a partir del Teorema 3.3.1. Por último, en los siguientes resultados utilizando el Teorema 3.3.1 calcularemos $TI(PC_n)$ de varios tipos de cadenas de poliominós.

Corolario 3.3.5. *Para la cadena de poliominós con $n \geq 3$ cuadrados y 2 segmentos s_1 y s_2 tales que $l_1 = 2$ y $l_2 = n - 1$, PC_n^1 , tenemos lo siguiente:*

$$TI(PC_3^1) = 2f(3, 4) + 4f(2, 3) + 2f(2, 4) + 2f(2, 2),$$

y para $n \geq 4$

$$TI(PC_n^1) = 3f(3, 3)n + 3f(3, 4) + 5f(2, 3) - 10f(3, 3) + f(2, 4) + 2f(2, 2).$$

Corolario 3.3.6. *Para la cadena de poliominós con $n \geq 5$ cuadrados y $m \geq 3$ segmentos s_1, s_2, \dots, s_m tales que $l_1 = l_m = 2$ y $l_2, \dots, l_{m-1} \geq 3$, PC_n^2 , tenemos lo siguiente:*

$$TI(PC_n^2) = 3f(3, 3)n + (4f(3, 4) + 2f(2, 3) - 6f(3, 3))m + 3f(3, 3) - 6f(3, 4) + 2f(2, 4) + 2f(2, 2).$$

Corolario 3.3.7. *Para la cadena de poliominós con $n \geq 6$ cuadrados y $m \geq 3$ segmentos s_1, s_2, \dots, s_m tales que $l_1 = 2$ y $l_2, \dots, l_{m-1}, l_m \geq 3$ o $l_m = 2$ y $l_1, l_2, \dots, l_{m-1} \geq 3$, PC_n^3 , tenemos lo siguiente:*

$$TI(PC_n^3) = 3f(3, 3)n + (4f(3, 4) + 2f(2, 3) - 6f(3, 3))m + 2f(2, 2) + f(2, 3) + 2f(3, 3) - 5f(3, 4) + f(2, 4).$$

Corolario 3.3.8. *Para la cadena de poliominós con $n \geq 7$ cuadrados y $m \geq 3$ segmentos s_1, s_2, \dots, s_m tales que $l_1, l_2, \dots, l_m \geq 3$, PC_n^4 , tenemos lo siguiente:*

$$TI(PC_n^4) = 3f(3, 3)n + (4f(3, 4) + 2f(2, 3) - 6f(3, 3))m + 2f(2, 2) + 2f(2, 3) + f(3, 3) - 4f(3, 4).$$

Los autores en [17, 42, 77] calcularon varios índices topológicos, tales como, el índice de Zagreb redefinido, Armónico y Suma Inversa para L_n , Z_n y PC_n^i con $i = 1, 2$; los cuales se deducen de los Corolarios 3.3.3, 3.3.5 y 3.3.6. Además, en [21, 22, 78] los autores calcularon el índice Olvidado, Randić y Zagreb generalizado para L_n , Z_n y PC_n^i con $i = 1, 2, 3, 4$; por lo tanto, podemos deducir los resultados mencionados anteriormente utilizando los Corolarios 3.3.3, 3.3.5, 3.3.6, 3.3.7 y 3.3.8. De hecho, los resultados mostrados en [21] pueden ser verificados directamente por el Corolario 3.3.4.

Por otro lado, aquí una cadena de poliominós de dimensión $n \geq 1$ con $k = k_1 + k_2 + k_3$ donde k_1 es el número de kinks, k_2 es el número de mediales y k_3 es el número de terminales en una unidad de la cadena de poliominó se denotará por $PC_{n,k}$. En la Figura 3.5, se muestra una representación general de una cadena de poliominós $PC_{n,k}$. Sea $k \geq 3$, por definición de $PC_{n,k}$, tenemos: $m = 2n, \gamma = n - 1$, $I_m = 1$ y $I_1 = I_{\{k=3\}}$. Por tanto, en el siguiente corolario, calcularemos $TI(PC_{n,k})$ para $k \geq 3$ utilizando el Teorema 3.3.1.

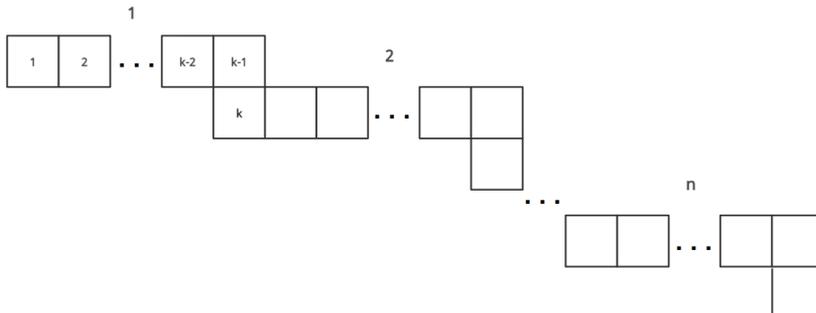


Figura 3.5: Representación general de $PC_{n,k}$.

Observación 3.3.9. *Nota que, por definición $PC_{n,1} = Li_n$ y $PC_{n,2} = Z_{2n}$.*

Corolario 3.3.10. *Sean $k \geq 3$, $n \geq 1$, entonces*

$$\begin{aligned}
 TI(PC_{n,k}) = & (3(k-3)f(3,3) + 4f(3,4) + 2f(2,3) + f(4,4) + \\
 & 2f(2,4))n + (f(2,4) - f(2,3) + f(3,3) - f(3,4))I_{\{k=3\}} + \\
 & 2f(2,2) + 3f(2,3) - f(3,3) - f(3,4) - f(2,4) - f(4,4).
 \end{aligned}$$

De hecho, en [32] Hayat et al. calcularon las expresiones analíticas exactas del índice ABC , GA , ABC_4 y GA_5 para $PC_{n,k}$ con $k = 3, 4, 5$. Estos resultados se pueden obtener como consecuencia del Corolario 3.3.10.

Capítulo 4

Conclusiones

En el Capítulo 1, investigamos una clase de *RST*s donde la variable aleatoria de interés es el número de hojas a medida que avanza el tiempo. En la misma dirección, calculamos la función generadora de momento de las hojas y mostramos que el número de hojas sigue una ley gaussiana asintótica. Luego, investigamos varios índices topológicos útiles para esta clase, incluido el índice de Gini basado en grados, el índice de Hoover basado en grados, el índice de Zagreb generalizado y otros índices asociados con estos. Además, se generalizan los resultados obtenidos en [61]. En particular, concluimos que la clase de *RST* que evoluciona de manera preferencial es relativamente menos regular que la clase de *RST* estudiada en este trabajo para algunos casos y que el índice de Hoover basado en grados es menor que el índice de Gini basado en grados en este modelo.

En el Capítulo 2 y 3, se propuso un enfoque de martingala para el estudio de los índices topológicos en las cadenas aleatorias de espiro y poliomínos. Se determinaron el valor esperado, la varianza y formulamos una martingala para caracterizar el comportamiento asintótico de los índices topológicos. Además, consideramos algunos índices topológicos particulares, tales como, el primer índice de Zagreb, el de Sombor, el Armónico, el Geométrico-Aritmético y el segundo índice de Zagreb. De hecho, a partir de los resultados derivados, se obtuvieron como corolarios varios resultados conocidos en la literatura. Creemos que los resultados obtenidos en este trabajo pueden proporcionar apoyo teórico a la investigación

química, ya que, unifican el estudio de los índices topológicos basados en grados en las cadenas estudiadas evitando razonamientos repetitivos y particulares.

Como trabajo futuro en esta dirección, se estudiarán las cadenas de poliomínos y espiros aleatorias extremas con respecto a varios índices topológicos conocidos basados en grados. Por último, una pregunta natural es si el mismo enfoque puede ser utilizado para otras cadenas aleatorias obteniendo procesos de Markov de otros órdenes. Esto abre una nueva línea de estudio, que esperamos desarrollar en el futuro.

Bibliografía

- [1] AGUILAR-SÁNCHEZ, R., MÉNDEZ-BERMÚDEZ, J. A., RODRÍGUEZ, J. M., AND SIGARRETA, J. M. Normalized Sombor Indices as Complexity Measures of Random Networks. *Entropy* 23, 8 (July 2021), 976.
- [2] ALI, A., DAS, K. C., DIMITROV, D., AND FURTULA, B. Atom–bond connectivity index of graphs: a review over extremal results and bounds. *Discrete Mathematics Letters* 5, 1 (2021), 68–93.
- [3] ALI, A., FURTULA, B., GUTMAN, I., AND VUKICEVIC, D. Augmented Zagreb index: Extremal results and bounds. *MATCH Commun. Math. Comput. Chem* 85 (2021), 211–244.
- [4] ALI, A., RAZA, Z., AND BHATTI, A. A. Bond incident degree (BID) indices of polyomino chains: A unified approach. *Applied Mathematics and Computation* 287 (2016), 28–37.
- [5] AN, M. The first Zagreb index, reciprocal degree distance and Hamiltonian-connectedness of graphs. *Information Processing Letters* 176 (June 2022), 106247.
- [6] BALABAN, A. T. *Chemical applications of graph theory*. Academic Press, 1976.
- [7] BARABÁSI, A.-L., AND ALBERT, R. Emergence of Scaling in Random Networks. *Science* 286, 5439 (oct 1999), 509–512.
- [8] BURAGOHAJ, J., DEKA, B., AND BHARALI, A. A generalized ISI index of some chemical structures. *Journal of Molecular Structure* 1208 (2020), 127843.

-
- [9] CANCAN, M., IMRAN, M., AKHTER, S., SIDDIQUI, M. K., AND HANIF, M. F. Computing Forgotten Topological Index of Extremal Cactus Chains. *Applied Mathematics and Nonlinear Sciences* 6, 1 (2021), 439–446.
- [10] DENG, H. Wiener indices of spiro and polyphenyl hexagonal chains. *Mathematical and Computer Modelling* 55, 3-4 (2012), 634–644.
- [11] DEVILLERS, J., AND BALABAN, A. Topological Indices and Related Descriptors in QSAR and QSPR. *Gordon and Breach* (1999).
- [12] DOMICOLA, C., AND MAHMOUD, H. Degree-based Gini index for graphs. *Probability in the Engineering and Informational Sciences* 34 (2020), 157–171.
- [13] DOŠLIĆ, T., RÉTI, T., AND VUKIČEVIĆ, D. On the vertex degree indices of connected graphs. *Chemical Physics Letters* 512, 4-6 (2011), 283–286.
- [14] DUCOFFE, G., MARINESCU-GHEMECI, R., OBREJA, C., POPA, A., AND TACHE, R. M. Extremal Graphs with respect to the Modified First Zagreb Connection Index. 141–148.
- [15] ESTRADA, E., TORRES, L., RODRIGUEZ, L., AND GUTMAN, I. An atom-bond connectivity index: modelling the enthalpy of formation of alkanes.
- [16] FANG, X., YOU, L., AND LIU, H. The expected values of Sombor indices in random hexagonal chains, phenylene chains and Sombor indices of some chemical graphs. *International Journal of Quantum Chemistry* (2021), e26740.
- [17] FAROOQ, A., HABIB, M., MAHBOOB, A., NAZEER, W., AND KANG, S. M. Zagreb Polynomials and redefined Zagreb indices of Dendrimers and Polyomino Chains. *Open Chemistry* 17, 1 (2019), 1374–1381.
- [18] FENG, Q., AND HU, Z. Asymptotic normality of the Zagreb index of random b-ary recursive trees. *Dal’nevostochnyi Matematicheskii Zhurnal* 15, 1 (2015), 91–101.

BIBLIOGRAFÍA

- [19] FILIPOVSKI, S. New Bounds for the First Zagreb Index. *MATCH Commun. Math. Comput. Chem* 85, 2 (2021), 303–312.
- [20] FURTULA, B., AND GUTMAN, I. A forgotten topological index. *Journal of mathematical chemistry* 53, 4 (2015), 1184–1190.
- [21] GAO, W., WANG, W., JAMIL, M. K., AND FARAHANI, M. R. Electron Energy Studying of Molecular Structures via Forgotten Topological Index Computation. *Journal of Chemistry 2016* (2016), 1–7.
- [22] GAO, W., YAN, L., SHI, L., ET AL. Generalized Zagreb index of polyomino chains and nanotubes. *Optoelectronics and Advanced Materials-Rapid Communications* 11, January-February 2017 (2017), 119–124.
- [23] GOLOMB, S. W. *Polyominoes*. Princeton University Press, Dec. 1994.
- [24] GUT, A. *Probability: A Graduate Course*. Springer, New York, 2005.
- [25] GUTMAN, I. Degree-Based Topological Indices. *Croatica Chemica Acta* 86, 4 (2013), 351–361.
- [26] GUTMAN, I. Geometric approach to degree-based topological indices: Sombor indices. *MATCH Commun. Math. Comput. Chem* (2021).
- [27] GUTMAN, I., AND DAS, K. C. The first Zagreb index 30 years after. *MATCH Commun. Math. Comput. Chem* 50, 1 (2004), 83–92.
- [28] GUTMAN, I., AND KULLI, V. Nirmala energy. *Open Journal of Discrete Applied Mathematics* 4, 2 (2021), 11–16.
- [29] GUTMAN, I., KULLI, V., AND REDZEPOVIC, I. Nirmala index of Kragujevac trees. *International Journal of Mathematics Trends and Technology* 67, 6 (2021), 44–49.

- [30] GUTMAN, I., AND TRINAJSTIĆ, N. Graph theory and molecular orbitals. Total φ -electron energy of alternant hydrocarbons. *Chemical Physics Letters* 17, 4 (Dec. 1972), 535–538.
- [31] HALL, P., AND HEYDE, C. C. *Martingale Limit Theory and Its Application*. Academic press, New York, 2014.
- [32] HAYAT, S., AHMAD, S., UMAİR, H. M., AND SHAOHUI, W. Distance property of chemical graphs. *Hacettepe Journal of Mathematics and Statistics* 47, 5 (2018), 1071–1093.
- [33] HOSOYA, H. Topological Index. A Newly Proposed Quantity Characterizing the Topological Nature of Structural Isomers of Saturated Hydrocarbons. *Bulletin of the Chemical Society of Japan* 44, 9 (1971), 2332–2339.
- [34] IQBAL, N., BHATTI, A. A., ALI, A., AND ALANAZI, A. M. On Bond Incident Connection Indices of Polyomino and Benzenoid Chains. *Polycyclic Aromatic Compounds* (2022), 1–8.
- [35] JAHANBANI, A. The First Zagreb and Randić Indices in Random Spiro Chains. *Polycyclic Aromatic Compounds* (2020), 1–9.
- [36] KAZEMI, R. Gordon-Scantlebury and Platt Indices of Random Plane-oriented Recursive Trees. *Mathematics Interdisciplinary Research* 6, 1 (Mar. 2021).
- [37] KHADIKAR, P. V., KARMARKAR, S., AND AGRAWAL, V. K. A Novel PI Index and Its Applications to QSPR/QSAR Studies. *Journal of Chemical Information and Computer Sciences* 41, 4 (June 2001), 934–949.
- [38] KIER, L. B., AND HALL, L. H. *Molecular connectivity in structure-activity analysis*. Research Studies, Wiley, New York, 1986.
- [39] KIER, L. B., HALL, L. H., MURRAY, W. J., AND RANDI, M. Molecular Connectivity I: Relationship to Nonspecific Local Anesthesia. *Journal of pharmaceutical sciences* 64, 12 (1975), 1971–1974.

- [40] KULLI, V. Nirmala Index. *International Journal of Mathematics Trends and Technology* 67 (03 2021), 8–12.
- [41] KULLI, V., CHALUVARAJU, B., AND ASHA, T. Computation of Nirmala indices of some chemical networks. *Journal of Ultra Scientist of Physical Sciences-A* 33, 4 (2021), 30–41.
- [42] KWUN, Y. C., FAROOQ, A., NAZEER, W., ZAHID, Z., NOREEN, S., AND KANG, S. M. Computations of the M-polynomials and Degree-Based Topological Indices for Dendrimers and Polyomino Chains. *International Journal of Analytical Chemistry* 2018 (2018).
- [43] LI, C.-P., ZHONGLIN, C., MUNIR, M., YASMIN, K., AND LIU, J.-B. M-polynomials and topological indices of linear chains of benzene, naphthalene and anthracene. *Mathematical Biosciences and Engineering* 17, 3 (2020), 2384–2398.
- [44] LI, J., AND WANG, W. The (degree-) Kirchhoff indices in random polygonal chains. *Discrete Applied Mathematics* 304 (2021), 63–75.
- [45] LI, S., SHI, L., AND GAO, W. Topological indices computing on random chain structures. *International Journal of Quantum Chemistry* 121, 8 (Dec. 2020).
- [46] LI, X., AND ZHENG, J. A unified approach to the extremal trees for different indices. *MATCH Commun. Math. Comput. Chem* 54, 1 (2005), 195–208.
- [47] LIN, Y., AND ZHANG, F. A linear algorithm for a perfect matching in polyomino graphs. *Theoretical Computer Science* 675 (2017), 82–88.
- [48] M. I. QURESHI, A. FAHAD, M. K. J. S. A. Zagreb Connection Index of Drugs Related Chemical Structures. *Bio-interface Research in Applied Chemistry* 11, 4 (Dec. 2020), 11920–11930.
- [49] MARTÍNEZ-MARTÍNEZ, C., MÉNDEZ-BERMÚDEZ, J., RODRÍGUEZ, J. M., AND SIGARRETA, J. M. Computational and analytical studies of the Randić index in Erdős–Rényi

- models. *Applied Mathematics and Computation* 377 (July 2020), 125137.
- [50] MEHLER, A., LÜCKING, A., AND WEISS, P. A network model of interpersonal alignment in dialog. *Entropy* 12, 6 (2010), 1440–1483.
- [51] MENÉNDEZ VELÁZQUEZ, A., ET AL. Una breve introducción a la teoría de grafos. *Suma* (1998).
- [52] MILOVANOVIC, I., MATEJIC, M., MILOVANOVIC, E., AND KHOEILAR, R. A note on the first Zagreb index and coindex of graphs. *Communications in Combinatorics and Optimization* 6, 1 (June 2021).
- [53] NIKOLIĆ, S., KOVAČEVIĆ, G., MILIČEVIĆ, A., AND TRINAJSTIĆ, N. The Zagreb indices 30 years after. *Croatica chemica acta* 76, 2 (2003), 113–124.
- [54] PEGU, A., DEKA, B., GOGOI, I. J., AND BHARALI, A. Two generalized topological indices of some graph structures. *J. Math. Comput. Sci.* 11, 5 (2021), 5549–5564.
- [55] PINEDA-PINEDA, J. J., MARTÍNEZ-MARTÍNEZ, C. T., MÉNDEZ-BERMÚDEZ, J. A., MUÑOZ-ROJAS, J., AND SIGARRETA, J. M. Application of Bipartite Networks to the Study of Water Quality. *Sustainability* 12, 12 (June 2020), 5143.
- [56] RAINS, E. M., AND SLOANE, N. J. On cayley’s enumeration of alkanes (or 4-valent trees). *Journal of Integer Sequences* 2 (1999), Art–No.
- [57] RANDIĆ, M. Characterization of Molecular Branching. *Journal of the American Chemical Society* 97, 23 (1975), 6609–6615.
- [58] RAZA, Z. The harmonic and second Zagreb indices in random polyphenyl and spiro chains. *Polycyclic Aromatic Compounds* (2020), 1–10.
- [59] RAZA, Z. The expected values of some indices in random phenylene chains. *The European Physical Journal Plus* 136, 1 (2021), 1–15.

- [60] RAZA, Z., AND IMRAN, M. Expected Values of Some Molecular Descriptors in Random Cyclooctane Chains. *Symmetry* 13, 11 (2021), 2197.
- [61] REN, Y., ZHANG, P., AND DEY, D. K. Investigating Several Fundamental Properties of Random Lobster Trees and Random Spider Trees. *Methodol Comput Appl Probab.* (2021).
- [62] SARDANA, S., AND MADAN, A. Application of graph theory: Relationship of antimycobacterial activity of quinolone derivatives with eccentric connectivity index and Zagreb group parameters. *MATCH Commun. Math. Comput. Chem* 45 (2002), 35–53.
- [63] SEVERINI, T. A., ET AL. *Elements of Distribution Theory.* Cambridge University Press, New York, 2012.
- [64] SEVERO, N. C., AND ZELEN, M. Normal approximation to the chi-square and non-central F probability functions. *Biometrika* 47, 3/4 (1960), 411–416.
- [65] SHAO, Z., JAHANBANI, A., AND SHEIKHOLESAMI, S. M. Multiplicative Topological Indices of Molecular Structure in Anticancer Drugs. *Polycyclic Aromatic Compounds* 42, 2 (Mar. 2020), 475–488.
- [66] SHIU, W. C. Extremal Hosoya index and Merrifield–Simmons index of hexagonal spiders. *Discrete applied mathematics* 156, 15 (2008), 2978–2985.
- [67] SIGARRETA, S., SIGARRETA, S., AND CRUZ-SUÁREZ, H. On bond incident degree indices of random spiro chains. *Polycyclic Aromatic Compounds* (2022), 1–13.
- [68] SIGARRETA, S. C., SIGARRETA, S. M., AND CRUZ-SUÁREZ, H. On degree-based topological indices of random polyomino chains. *Mathematical Biosciences and Engineering* 19, 9 (2022), 8760–8773.
- [69] SIGARRETA, S. C., SIGARRETA, S. M., AND CRUZ-SUÁREZ, H. Probabilistic methods in the study of topological indices on random spider trees. *arXiv preprint arXiv:2204.10840* (2022).

-
- [70] VUKIČEVIĆ, D., AND FURTULA, B. Topological index based on the ratios of geometrical and arithmetical means of end-vertex degrees of edges. *Journal of mathematical chemistry* 46, 4 (2009), 1369–1376.
- [71] WEI, S., KE, X., AND HAO, G. Comparing the expected values of atom-bond connectivity and geometric–arithmetic indices in random spiro chains. *Journal of inequalities and applications* 2018, 1 (2018), 1–11.
- [72] WEI, S., KE, X., AND LIN, F. Perfect matchings in random polyomino chain graphs. *Journal of Mathematical Chemistry* 54, 3 (2016), 690–697.
- [73] WEI, S., AND SHIU, W. C. Enumeration of Wiener indices in random polygonal chains. *Journal of Mathematical Analysis and Applications* 469, 2 (2019), 537–548.
- [74] WIENER, H. Structural Determination of Paraffin Boiling Points. *Journal of the American chemical society* 69, 1 (1947), 17–20.
- [75] WU, T., LÜ, H., AND ZHANG, X. Extremal Matching Energy of Random Polyomino Chains. *Entropy* 19, 12 (2017), 684.
- [76] XIAO, C., AND CHEN, H. Dimer Coverings on Random Polyomino Chains. *Zeitschrift für Naturforschung A* 70, 6 (2015), 465–470.
- [77] YANG, J., XIA, F., AND CHEN, S. On sum-connectivity index of polyomino chains. *Appl. Math. Sci* 5, 6 (2011), 267–271.
- [78] YANG, J., XIA, F., AND CHEN, S. On the Randić index of polyomino chains. *Appl. Math. Sci* 5, 5 (2011), 255–260.
- [79] YARAHMADI, Z. Finding extremal total irregularity of polyomino chain by transformation method. *Journal of New Researches in Mathematics* 7, 30 (2021), 141–150.
- [80] ZHANG, P., AND WANG, X. Several Topological Indices of Random Caterpillars. *Methodol Comput Appl Probab.* (2021).

BIBLIOGRAFÍA

- [81] ZHANG, W., YOU, L., LIU, H., AND HUANG, Y. The expected values and variances for sombor indices in a general random chain. *Applied Mathematics and Computation* 411 (2021), 126521.
- [82] ZHOU, X., AND ZHANG, H. A minimax result for perfect matchings of a polyomino graph. *Discrete Applied Mathematics* 206 (2016), 165–171.