



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

POLÍMEROS ALEATORIOS SUAVES EN
DIMENSIÓN UNO

TESIS

QUE PARA OBTENER EL TÍTULO DE:

MATEMÁTICO

PRESENTA:

DAVID RIVERA BERMÚDEZ

DIRECTORA DE TESIS:

ANA MEDA GUARDIOLA



Ciudad Universitaria, 2016

CDMX



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Agradecimientos:

A LA UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO Y A LA FACULTAD DE CIENCIAS: por permitirme formar parte de este recinto sagrado de conocimiento y cultura, por darme la oportunidad de desarrollarme como una de las mejores mentes del mundo y por ser uno de los pocos lugares donde las personas que nacen con los ojos abiertos y los cerebros dispuestos encuentran su verdadero hogar.

A todos y cada uno de mis profesores, por enseñarme más que lo que esperaban, porque un verdadero profesor no enseña, inspira; y en especial a la Dra. Ana Meda Guardiola, por facilitarme todo su conocimiento, por su infinita paciencia a la hora de desarrollar juntos este trabajo y sobre todo porque con su pasión al momento de enseñar, me inspiró y ayudó a darme cuenta que mi verdadero lugar está en las aulas.

A todos los que han sido mis alumnos, porque la mejor manera de aprender es enseñar.

A mis amigos: Sara, Dana, Mayte, Chícharo, Fer y Ray, mis verdaderos amigos, los que se volvieron mis hermanos... porque ¿quién sabe lo que cuesta llegar aquí si no los amigos?... luchamos juntos, nos ayudamos los unos a los otros en todo lo que pudimos y en lo que no pudimos también, nos vimos caer y levantar una y otra vez, con nuestra directa e insensible honestidad nos apoyamos en los momentos más difíciles, no sólo académicamente; si no en cualquier aspecto de nuestras vidas.

A mi gran familia, porque con su unión, fortaleza y admiración me han enseñado a ser mejor cada día sin dejar de ser yo mismo.

A Reyna, Pinky, Junior, Max, Nina, Kiko, Kitty y Hanna.

A Santiago, para que cuando crezca vea aquí escrito que lo amo mucho de verdad.

A mi hermana, porque es mi hermana aunque a veces ella no lo quiera así... y porque voy a estar con ella siempre, apoyándola e intentando ser un buen ejemplo.

A mi papá, por darme todo lo que necesito y más a costa de su bienestar y su comodidad, porque siempre ha estado a mi lado facilitándome las cosas sin pedirme alguna explicación, y sobre todo; gracias porque con su ejemplo aprendí que no importa cómo te sientas física y mentalmente: todos y cada uno de los días tienes que salir a trabajar.

Jamás encontraré las palabras ni las acciones que mereces, simplemente: Gracias por ser mi mamá y todo lo que eso significa. Gracias a tí por traerme a este mundo y por ser la mejor persona para guiar mi camino, porque ahora entiendo todas y cada una de las razones por las cuales me premiaste y me regañaste, porque ahora conozco el valor del sacrificio y del trabajo y porque siempre has estado, estás y estarás conmigo.

Índice general

1. Los Polímeros y sus Modelos Matemáticos	11
1.1. Polímeros	11
1.1.1. Clasificación	11
1.1.2. Polimerización	13
1.1.3. Física de polímeros	15
1.2. Modelos matemáticos	16
2. Caminatas Aleatorias	19
2.1. Caminata aleatoria simple	19
2.2. Propiedades	20
2.2.1. Probabilidades de transición	21
2.2.2. Regreso a la posición de origen	23
2.3. Caminata aleatoria auto-evasiva	26
2.4. Caminata aleatoria auto-evasiva débil	26
2.4.1. Puentes	27
2.4.2. <i>Desdoblado</i> una trayectoria en un puente	29
3. Grandes Desviaciones	31
3.1. Un <i>volado</i>	32
3.2. Teorema de Cramér	35
3.3. Definiciones generales	41
3.4. Lema de Varadhan	43
4. Polímeros Aleatorios Suaves	47
4.1. Polímeros con auto-repelencia	47
4.2. Grandes desviaciones para puentes	50
4.3. El programa de cinco pasos	54
4.3.1. Paso 1: Añadiendo deriva	54

4.3.2. Paso 2: Tiempos locales y propiedad de Markov	57
4.3.3. Paso 3: El problema variacional	59
4.3.4. Paso 4: La función tasa para puentes.	66
4.3.5. Paso 5: Identificación de la velocidad	72
4.4. Principio de grandes desviaciones sin la condición de puente . .	74
4.5. Demostración del resultado principal	79
5. Comentarios finales	83
Apéndice	87
Bibliografía	95

Introducción

En este trabajo describimos un modelo matemático para un polímero de dimensión uno utilizando principalmente resultados sobre caminatas aleatorias y grandes desviaciones. Estas moléculas, formadas por moléculas más pequeñas, elementales o básicas, son consideradas como caminatas aleatorias en un cierto espacio, donde cada vértice de la trayectoria es uno de estos elementos más pequeños y los enlaces entre cada vértice son enlaces químicos que mantienen unido a todo el polímero. Debido a los factores físicos y químicos implícitos en todos los elementos de la naturaleza, cada modelo que describe un polímero lo hace para un número de moléculas básicas fijo, ya que alterar el sistema agregando un elemento básico podría alterar la composición de toda la molécula y por tanto el modelo que lo describe podría ya no funcionar.

Algunos polímeros tienen un comportamiento denominado *balístico*, están cerca de su punto inicial por un tiempo y después tienden a infinito; mismo que el modelo que trabajamos busca describir.

En el primer capítulo damos un breve resumen sobre lo que son los polímeros, algunas de las propiedades más importantes respecto a su composición, la manera de considerarlos desde el punto de vista de distintas disciplinas y los modelos matemáticos que se utilizan para describirlos.

En el capítulo dos se establecen resultados sobre caminatas aleatorias, las cuales son la piedra angular para el desarrollo de este análisis, se define lo que es una caminata aleatoria simple y algunas de sus propiedades, una caminata aleatoria auto-evasiva y una caminata aleatoria auto-evasiva débil.

En el capítulo tres damos las definiciones y resultados de teoría de grandes

desviaciones que utilizamos en el trabajo.

En el capítulo cuatro establecemos y analizamos el modelo, aquí juntamos las herramientas matemáticas antes mencionadas para describir a los polímeros y probar los resultados pertinentes. Las demostraciones de los teoremas y las gráficas son principalmente de [1], [2] y [4].

Los resultados que se prueban aquí establecen que estos elementos principalmente naturales se puede modelar considerándolos como caminatas aleatorias en ciertos espacios y bajo ciertas condiciones que describen sus distintas y fascinantes propiedades.

Capítulo 1

Los Polímeros y sus Modelos Matemáticos

En este capítulo explicaremos qué es un polímero y algunos de los ingredientes matemáticos requeridos para describirlos. Para la breve descripción de estas macromoléculas nos hemos basado en [12] y en los trabajos de F. den Hollander [1] y [3] para la discusión del modelo matemático.

1.1. Polímeros

Un *polímero* es una macromolécula (molécula con una masa molecular elevada) formada por un gran número de átomos. Generalmente se pueden describir como la repetición de una o unas pocas unidades mínimas llamadas *monómeros*, los cuales están 'atados' por medio de enlaces químicos. Los monómeros pueden ser unidades pequeñas (tales como el CH_2 en el polietileno) o unidades más grandes con una estructura interna (tales como los pares básicos de la doble hélice del ADN, adenina-timina y citosina-guanina). Los polímeros abundan en la naturaleza por la multivalencia de átomos como el carbono, oxígeno, nitrógeno, silicio, azufre y fósforo, los cuales son capaces de formar grandes estructuras concatenadas.

1.1.1. Clasificación

Los polímeros se pueden clasificar en dos tipos:

1. *Homopolímeros*; con todos los monómeros del mismo tipo, como el polietileno (Figura 1.1¹).
2. *Copolímeros*; con dos o más tipos de monómeros, por ejemplo el ADN (Figura 1.2²). La distribución de monómeros diferentes en un copolímero puede ser periódico como en el agar-agar o aleatorio como la carragenina.

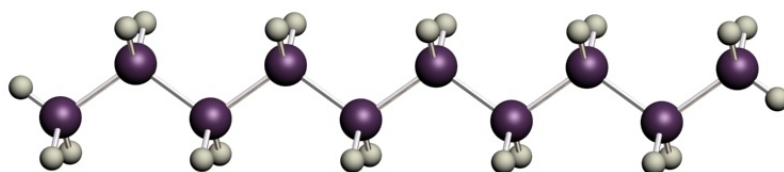


Figura 1.1: Polietileno



Figura 1.2: ADN

Una segunda clasificación, de gran importancia, es en *sintéticos* (como el nylon, el polietileno y el poliestireno) y *naturales* (también llamados biopolímeros) tales como:

¹Fuente: http://www.mim-us.es/estructuras_cristalinas/polietileno.html

²Fuente: <https://www.emaze.com/@ALLIQRR/-Examen-ADN>

1. *proteínas*; cadenas de aminoácidos, las estructuras principales para formar organismos vivos, las cuales realizan una gran cantidad de tareas importantes.
2. *ácidos nucleicos*; ADN y ARN, los bloques en la construcción de los genes, los cuales son el centro mismo del proceso de la vida.
3. *polisacáridos*; agar-agar, amilosa, carragenina, celulosa, los cuales forman parte de la estructura de todos los animales y plantas, y proporcionan una importante fuente de energía.
4. *lignina*; presente en las paredes celulares de organismos del reino *plantae*, el cual llena el espacio entre las fibras de celulosa; se encarga de engrosar el tallo de las plantas.
5. *hule*; presente en el fluido de las células de latex en ciertos tipos de árboles y arbustos.

Además de estos *materiales orgánicos*, arcillas y minerales, existen ejemplos *inorgánicos* de polímeros naturales.

Los polímeros sintéticos generalmente son homopolímeros y los naturales copolímeros (con notables excepciones), los polisacáridos bacterianos tienden a ser periódicos y los polisacáridos de las plantas aleatorios.

Otra clasificación importante de los polímeros es en *lineales*, cuyos monómeros sólo tienen un grupo reactivo (como el CH_2) lo que conduce a una organización lineal como resultado del proceso de polimerización y *ramificados*, donde los monómeros tienen dos o más grupos reactivos (como el ácido hidroxicítrico) formando una compleja red con múltiples conexiones. Muchos polímeros naturales son lineales como el ADN, el ARN, las proteínas y los polisacáridos agar-agar, alginato, amilosa, carragenina y celulosa. Algunos polisacáridos son ramificados como la amilopectina. Muchos polímeros sintéticos son lineales y muchos otros son ramificados, tales como el hule natural y el sintético.

1.1.2. Polimerización

El proceso para construir un polímero de un conjunto de monómeros por medio de distintas reacciones químicas se conoce como *polimerización*, por

ejemplo; la polimerización del etileno nos proporciona polietileno, una muestra típica de este puede contener moléculas de hasta 50,000 átomos de carbono enlazados en una sola cadena. Esta estructura distingue a los polímeros de otros materiales y da lugar a sus propiedades y caracterización.

El *tamaño* de un polímero, es decir; el número de monómeros que lo constituyen (también llamado *grado de polimerización*) puede variar desde 10^3 hasta 10^{10} (moléculas más pequeñas no se consideran polímeros y moléculas más grandes no se conocen). El ADN humano tiene de 10^9 a 10^{10} pares básicos, la lignina entre 10^6 y 10^7 fenil-propanos, mientras que los polisacáridos tienen de 10^3 a 10^4 azúcares.

En ambos tipos de polímeros (sintéticos y naturales) su tamaño puede ser muy extenso, con cantidades variando significativamente de polímero a polímero (como en el nylon y los polisacáridos) o puede ser reducido (como en las proteínas y el ADN); en los polímeros sintéticos el tamaño puede ser reducido a través de ciertos métodos específicos de polimerización. El tamaño de un monómero varía de $1,5 \text{ \AA}$ (con $1 \text{ \AA} = 10^{-10} \text{ m}$) como para el CH_2 en el polietileno, hasta 20 \AA para los pares básicos en el ADN.

Los enlaces químicos en un polímero son flexibles, así que la forma del polímero puede ser cualquiera de muchas *configuraciones espaciales*, entre más larga sea la cadena, más configuraciones son posibles. Es por eso que un polímero puede enrollarse hasta formar pequeños puntos, puede extenderse debido a fuerzas repelentes entre los monómeros dando como resultado el *efecto del volumen excluido* o puede colapsar a una pequeña *bola* debido a las *fuerzas de Van der Waals* entre ellos mismos y un solvente; también pueden interactuar con ciertas superficies en las cuales pueden ser o no ser absorbidos o permanecer entre dos superficies determinadas.

Como mencionamos previamente, un polímero puede ser tanto homogéneo (homopolímeros) como no-homogéneo (copolímeros), un ejemplo de estos es un copolímero con cargas positivas y negativas distribuidas a lo largo de la cadena, o uno constituido por monómeros hidrofóbicos e hidrofílicos, si tal copolímero es puesto cerca de una interfaz que separa agua y aceite, este tratará de moverse alrededor de esta con el objetivo de *acomodar* sus monómeros lo más posible de acuerdo a su configuración y una vez hecho esto, permanecerá cerca de ella.

1.1.3. Física de polímeros

La *física de polímeros* es un área de la física (originalmente física estadística) que estudia las características de los polímeros, su composición, sus fluctuaciones y propiedades mecánicas, así como la cinética de las reacciones que involucran la degradación de los mismos y la polimerización de los monómeros.

La física de polímeros y la química de polímeros están relacionadas también con el área de la ciencia de polímeros, en donde se considera como la parte de las aplicaciones.

Como mencionamos previamente, los polímeros son moléculas muy grandes y debido a esto es muy complicado describirlas utilizando métodos deterministas ([1], capítulo 9). Sin embargo, los enfoques probabilísticos y estadísticos proporcionan resultados que son consistentes con lo visto en la realidad, debido a que son descriptibles de manera eficiente en el límite termodinámico de un número infinito de monómeros (aunque su tamaño es claramente finito).

Las fluctuaciones térmicas continuamente afectan la forma de los polímeros en disoluciones líquidas, y para modelar estos efectos se requiere utilizar resultados de mecánica estadística. La temperatura afecta fuertemente el comportamiento físico de los polímeros, causando transiciones de fase, fusiones, etc.

La aproximación estadística para la física de polímeros está basada en una analogía entre un polímero y ya sea un movimiento browniano u otro tipo de camino aleatorio; la caminata aleatoria auto-evasiva. El modelo más simple posible de un polímero es presentado por la cadena ideal, correspondiente a una caminata aleatoria simple. Las aproximaciones experimentales para su descripción son también muy comunes para determinar las propiedades químicas, físicas, y materiales; usando métodos de caracterización tales como la cromatografía de exclusión molecular, viscosimetría, dispersión dinámica de luz, el monitoreo en línea automático y continuo de las reacciones de polimerización, etcétera. Estos métodos experimentales también ayudan a establecer los modelos matemáticos e incluso a un mejor entendimiento de sus propiedades.

Los modelos que se ajustan de cadenas de polímeros se dividen en dos tipos: *modelos ideales*, y *modelos reales*. Los modelos para una cadena ideal asumen que no hay interacciones entre los monómeros de la cadena. Esta suposición es válida solo para ciertos sistemas poliméricos, donde las interacciones positivas y negativas entre los monómeros se cancelan de manera efectiva. Los modelos de cadena ideal establecen un buen punto de partida para la investigación de sistemas más complejos y son los más adecuados para ecuaciones con más parámetros.

El modelo de la cadena ideal asume que segmentos del polímero pueden trasladarse entre sí como si la cadena fuera una *cadena fantasma* ya que en realidad, dos segmentos no pueden ocupar el mismo espacio al mismo tiempo. Esta interacción entre segmentos es llamada *efecto de volumen excluido*. La formulación más simple del volumen excluido es la caminata aleatoria auto-evasiva, la cual es un modelo para una cadena real; esta es una caminata aleatoria que no puede repetir su ruta previa. Una trayectoria de n pasos de esta caminata en tres dimensiones representa un polímero con interacción de volumen excluido. Debido a la naturaleza auto-evasiva de este modelo, el número de configuraciones posibles es significativamente reducido.

1.2. Modelos matemáticos

En resumidas cuentas, un *polímero* es una larga cadena de *monómeros* (moléculas más pequeñas) conectados entre sí por enlaces químicos; dos características de estos son:

1. *Irregularidad espacial*, porque las cadenas tienden a estar interconectadas en varios y distintos sitios.
2. *Auto-repelencia*, porque los monómeros tienden a tener interacción entre ellos, ya sea atractiva o repulsiva; lo que da lugar a lo que se conoce como el *efecto del volumen excluido*.

Para incorporar la primera de estas características a un modelo matemático, pensaremos al polímero como la trayectoria de una *caminata aleatoria* en \mathbb{Z}^d , $d \geq 1$; donde X_i (el paso al tiempo i) es la orientación del enlace químico entre el $(i - 1)$ -ésimo y el i -ésimo monómero, mientras que S_n (la posición al

tiempo n) corresponde a la ubicación del n -ésimo monómero, el punto final de la cadena.

Para la segunda propiedad, fijamos n y definimos los siguientes elementos

- \mathcal{W}_n = una colección de trayectorias de n pasos en \mathbb{Z}^d .
- H_n = una función que asocia una cierta energía a cada trayectoria en \mathcal{W}_n .

y definimos, para cada $w \in \mathcal{W}_n$ una distribución de probabilidad

$$\mathbb{P}_n(w) = \frac{1}{Z_n} e^{-H_n(w)}$$

donde Z_n es una constante normalizadora. Esta probabilidad es conocida como la *medida de Gibbs* asociada a (\mathcal{W}_n, H_n) y describe al polímero en *equilibrio consigo mismo y con su ambiente*.

La elección de \mathcal{W}_n depende de la aplicación que se tenga en mente y H_n se determinará considerando factores físicos y químicos implícitos en la aplicación y captura la interacción del polímero consigo mismo y/o con su entorno; esta función generalmente depende de dos o tres parámetros incluida la temperatura y normalmente se define de tal modo que sea directamente proporcional al número de auto-intersecciones en la caminata, es decir; es un número positivo o cero.

Una descripción completa de las propiedades espaciales de un polímero resulta ser un trabajo complicado debido a la gran cantidad de factores internos y externos involucrados en la composición de estos elementos, es por eso que se busca simplificar los modelos.

Bajo la medida de Gibbs, trayectorias con poca energía son más probables, mientras que trayectorias con alta energía son menos probables. Generalmente (\mathbb{P}_n) *no* es una familia consistente de distribuciones de probabilidad, es decir; \mathbb{P}_n no es la restricción de \mathbb{P}_{n+1} sobre los primeros n pasos, \mathbb{P}_n no se obtiene simplemente quitando la posición del $(n + 1)$ -ésimo monómero. De hecho, para cada n tenemos una distribución diferente modelando la cadena de un polímero de una longitud fija.

Existen modelos donde la función de energía $H_n(w)$ también depende de un

ambiente aleatorio, es decir; cargas positivas y negativas están distribuidas aleatoriamente a lo largo de toda la cadena o esta está formada por distintos tipos de monómeros, este ambiente aleatorio es denotado por ω y se escribe H_n^ω para enfatizar su dependencia con el ambiente. Tres tipos de medidas de Gibbs consideradas para estos modelos son:

1. La medida de Gibbs para un *sistema templado*

$$\mathbb{P}_n^\omega(w) = \frac{1}{Z_n^\omega} e^{-H_n^\omega(w)}, \quad w \in \mathcal{W}_n$$

2. La medida de Gibbs *para el estado promedio de un sistema templado*

$$\mathbb{E} [\mathbb{P}_n^\omega(w)] = \int \mathbb{P}_n^\omega(w) \mathbb{P}(d\omega), \quad w \in \mathcal{W}_n$$

3. La medida de Gibbs *recocida*

$$\mathbb{P}_n(w) = \frac{1}{Z_n} \int e^{-H_n^\omega(w)} \mathbb{P}(d\omega), \quad w \in \mathcal{W}_n$$

Esta última se utiliza para modelar un polímero cuyo ambiente aleatorio no está congelado pero forma parte del estado de equilibrio.

Al desarrollar estos modelos se está principalmente interesados en el comportamiento del polímero cuando su longitud (n) tiende a infinito.

Se han establecido modelos donde la configuración del polímero cambia con el tiempo; estas situaciones *sin equilibrio* forman un campo bastante interesante y complicado, pero hasta ahora, según las fuentes consultadas [3]; las matemáticas involucradas no están suficientemente desarrolladas.

Capítulo 2

Caminatas Aleatorias

En este capítulo presentamos dos modelos básicos para modelar un polímero, la *caminata aleatoria simple*, que describe a un polímero sin ninguna auto-interacción y la *caminata aleatoria auto-evasiva* que describe a un polímero con auto-interacción, dando lugar al "efecto del volumen excluido", es decir; ningún sitio puede ser ocupado por más de un monómero. En este trabajo consideraremos a la caminata aleatoria simple en \mathbb{Z} como base para desarrollar nuestro análisis para los *polímeros suaves*, nos hemos basado principalmente en [1], [8] y [9] para la descripción de estos procesos.

2.1. Caminata aleatoria simple

Una *caminata aleatoria simple* en \mathbb{Z} es el proceso aleatorio $(S_n)_{n \in \mathbb{N}_0}$ definido por:

$$S_0 = 0, \quad S_n = \sum_{i=0}^n X_i, \quad n \in \mathbb{N}$$

donde $(X_i)_{i \in \mathbb{N}}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas con valores en \mathbb{Z} con

$$\mathbb{P}(X_i = +1) = \frac{1}{2} = \mathbb{P}(X_i = -1)$$

Escribiremos \mathbb{P} y \mathbb{E} para denotar la probabilidad y la esperanza respecto a una caminata aleatoria simple.

Desde el punto de vista de los polímeros, este modelo corresponde a elegir el conjunto de trayectorias y a la función que asocia la energía como:

- $\mathcal{W}_n = \{w = (w_i)_{i=0}^n \in \mathbb{Z}^{n+1} : w_0 = 0, |w_{i+1} - w_i| = 1 \forall 0 \leq i < n\}$.
- $H_n \equiv 0$.

y a la medida \mathbb{P}_n como la distribución uniforme en \mathcal{W}_n .

Esto describe un polímero sin interacción con $(S_i)_{i=0}^n$ tomando valores en \mathcal{W}_n . Claramente $|\mathcal{W}_n| = 2^n$ y \mathbb{P}_n es la restricción de \mathbb{P} en \mathcal{W}_n , por lo que $(\mathbb{P}_n)_{n \in \mathbb{N}_0}$ sí es una familia consistente de medidas de probabilidad.

2.2. Propiedades

Enunciamos algunas definiciones y propiedades útiles para nuestro análisis relativas a una caminata aleatoria en \mathbb{Z} .

Sea $(S_n)_{n \in \mathbb{N}_0}$ una caminata aleatoria en \mathbb{Z} con las siguientes probabilidades de transición definidas para cualquier $n \geq 0$ y para cualesquiera $i, j \in \mathbb{Z}$

$$\mathbb{P}(S_{n+1} = j | S_n = i) = \begin{cases} p & \text{si } j = i + 1 \\ q & \text{si } j = i - 1 \\ 0 & \text{en otro caso} \end{cases}$$

Estas probabilidades no dependen de n , por lo que decimos que son *homogéneas* en el tiempo. Claramente este proceso es una cadena de Markov, es decir; el estado futuro del proceso depende sólo del estado presente y no de los que fueron visitados antes.

Si $(X_i)_{i \in \mathbb{N}}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas con

$$\mathbb{P}(X_i = +1) = p \quad \text{y} \quad \mathbb{P}(X_i = -1) = q$$

podemos definir al proceso $(S_n)_{n \in \mathbb{N}_0}$ como antes

$$S_0 = 0, \quad S_n = \sum_{i=0}^n X_i, \quad n \in \mathbb{N}.$$

PROPOSICIÓN. Para cualquier $n \in \mathbb{N}_0$

1. $\mathbb{E}[S_n] = n(p - q)$.
2. $Var(X_n) = 4npq$.

DEMOSTRACIÓN. Notemos que $\mathbb{E}[X_i] = p - q$ por lo que

$$\mathbb{E}[S_n] = \sum_{i=1}^n \mathbb{E}[X_i] = n\mathbb{E}[X_i] = n(p - q).$$

Por otro lado $\mathbb{E}[X_i^2] = p + q = 1$, entonces $Var(X_i) = 1 - (p - q)^2 = 4pq$ y

$$Var(S_n) = \sum_{i=1}^n Var(X_i) = nVar(X_i) = 4npq.$$

□

De lo anterior, notemos que si la probabilidad de dar un paso a la derecha es mayor que la probabilidad de dar un paso a la izquierda $p > q$, entonces el estado promedio de la caminata después de n pasos es un número positivo, es decir; en promedio la caminata tiende a ir a la derecha y de la misma manera, si $p < q$ en promedio la caminata tiende a ir a la izquierda. Notemos también que conforme n crece, la varianza también lo hace; esto indica que entre más pasos efectúe la caminata, más incierta es la posición final del proceso.

Cuando $p \neq q$ decimos que la caminata es *asimétrica*, si $p = q = \frac{1}{2}$ la caminata es *simétrica* y en promedio el proceso se queda en su posición inicial, ya que $\mathbb{E}[S_n] = 0$, sin embargo, en este caso $Var(S_n) = n$ y es sencillo verificar que este es el valor máximo que se alcanza cuando $p \in (0, 1)$.

2.2.1. Probabilidades de transición

Como hemos establecido que la caminata empieza en 0, podemos deducir que después de que esta dé un número par de pasos, la posición final sólo puede ser un número par o 0, en cambio; si da un número impar de pasos la posición final es un número impar. También es claro que después de efectuar n pasos, la distancia máxima a la que puede llegar es n unidades a la derecha o a la izquierda de 0.

Con estas consideraciones podemos deducir el siguiente resultado.

PROPOSICIÓN. *Para cualesquiera $x, n \in \mathbb{Z}$ tales que $-n \leq x \leq n$ y ambos pares o ambos impares*

$$\mathbb{P}(S_n = x \mid S_0 = 0) = \binom{n}{\frac{n+x}{2}} p^{\frac{n+x}{2}} q^{\frac{n-x}{2}}$$

y esta probabilidad vale 0 si x y n no cumplen con las condiciones previas.

DEMOSTRACIÓN. Elegimos una trayectoria de n pasos y sean R_n y L_n el número de pasos que la caminata da a la derecha y a la izquierda respectivamente. Notemos que

$$S_n = R_n - L_n \quad \text{y} \quad n = R_n + L_n$$

sumando estas dos igualdades tenemos que $S_n + n = 2R_n$ y sabiendo que $S_n = \sum_{i=1}^n X_i$

$$R_n = \frac{1}{2}(S_n + n) = \frac{1}{2} \left(\sum_{i=1}^n X_i + n \right) = \sum_{i=1}^n \frac{1}{2}(X_i + 1)$$

Obsérvese que esta fórmula asocia un valor entero a R_n cuando n y S_n son ambos pares o impares. Como las variables X_i son independientes y toman los valores $+1$ y -1 , las variables $\frac{1}{2}(X_i + 1)$ también son independientes y toman los valores 1 y 0 con probabilidades p y q respectivamente, por esto $R_n \sim \text{Bin}(n, p)$.

Por tanto, para cualquier valor de x que cumpla las condiciones se tiene que

$$\mathbb{P}(S_n = x \mid S_0 = 0) = \mathbb{P} \left(R_n = \frac{1}{2}(n + x) \right) = \binom{n}{\frac{n+x}{2}} p^{\frac{n+x}{2}} q^{\frac{n-x}{2}}.$$

■

Este resultado puede extenderse al caso general de pasar del estado y al estado x en n pasos.

PROPOSICIÓN. Para cualesquiera $y, x, n \in \mathbb{Z}$ tales que n y $x - y$ son ambos pares o ambos impares y $-n \leq x - y \leq n$

$$\mathbb{P}(S_n = x \mid S_0 = y) = \binom{n}{\frac{n+x-y}{2}} p^{\frac{n+x-y}{2}} q^{\frac{n-x+y}{2}}$$

y esta probabilidad vale 0 si x, y y n no cumplen con las condiciones previas.

DEMOSTRACIÓN. Como $S_0 = y$, el proceso $(T_n)_{n \in \mathbb{N}_0}$ con $T_n = S_n - y$, es una caminata aleatoria que inicia en 0 y como

$$\mathbb{P}(S_n = x \mid S_0 = y) = \mathbb{P}(T_n = x - y \mid T_0 = 0)$$

por lo anterior anterior se obtiene el resultado. ■

2.2.2. Regreso a la posición de origen

Nos interesa conocer la probabilidad de que una caminata aleatoria que inicia en el estado 0 regrese eventualmente a su punto de partida.

TEOREMA. Para una caminata aleatoria en \mathbb{Z} , la probabilidad de regresar al punto de origen es

$$\mathbb{P}(S_n = 0 \text{ p.a. } n > 0 \mid S_0 = 0) = 1 - |p - q|$$

DEMOSTRACIÓN. Primero definimos los siguientes elementos

1. Para cada $n \geq 0$ definimos $p_n = \mathbb{P}(S_n = 0 \mid S_0 = 0)$, es decir; p_n es la probabilidad de que la caminata esté en el estado 0 al tiempo n , la cual es distinta de cero sólomente cuando n es un número par, y además $p_0 = 1$.
2. Definimos a f_k como la probabilidad de que la caminata visite por primera vez el estado 0 en el paso $k \geq 0$, con $f_0 = 0$. Notemos que la probabilidad de que la caminata regrese en algún momento a 0 es

$$\sum_{k=0}^{\infty} f_k.$$

Esta serie es convergente ya que las f_k 's son probabilidades de eventos disjuntos y su valor es a lo más 1. Recordemos que los valores que toma cada f_k son estrictamente positivos sólo cuando k es un número par.

OBSERVACIÓN.
$$p_n = \sum_{k=0}^n f_k p_{n-k}.$$

Utilizaremos esta expresión para encontrar la función generadora de probabilidad de la sucesión f_0, f_1, f_2, \dots , es decir;

$$G(t) = \sum_{k=0}^{\infty} f_k t^k.$$

Multiplicando por t^n ambos lados de la expresión para p_n y sumando sobre n tenemos que

$$\begin{aligned} \sum_{n=1}^{\infty} p_n t^n &= \sum_{n=1}^{\infty} \left(\sum_{k=0}^n f_k p_{n-k} \right) t^n = \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} f_k p_{n-k} t^n \\ &= \sum_{k=0}^{\infty} f_k t^k \sum_{n=k}^{\infty} p_{n-k} t^{n-k} = G(t) \sum_{n=0}^{\infty} p_n t^n \end{aligned}$$

por lo que

$$\begin{aligned} \sum_{n=1}^{\infty} p_n t^n &= G(t) \sum_{n=0}^{\infty} p_n t^n \\ \left(\sum_{n=0}^{\infty} p_n t^n \right) - 1 &= G(t) \sum_{n=0}^{\infty} p_n t^n \end{aligned}$$

3. Necesitamos una expresión para $\sum_{n=0}^{\infty} p_n t^n$ la cual es la función generadora de la sucesión p_0, p_1, p_2, \dots

Recordemos que para cualquier $a \in \mathbb{R}$ y cualquier $n \in \mathbb{N}$ se tiene el coeficiente binomial

$$\binom{a}{n} = \frac{a(a-1)\cdots(a-(n-1))}{n!}$$

y que la expansión

$$(1+t)^a = \sum_{n=0}^{\infty} \binom{a}{n} t^n$$

es válida para $|t| < 1$.

En particular

$$\binom{2n}{n} = \frac{(2n)!}{n!n!} = \frac{2n(2n-1)(2n-2)\cdots 3\cdot 2\cdot 1}{n!n!}$$

agrupando los números pares e impares en el numerador y factorizando

$$\begin{aligned} &= \frac{2^n n! (2n-1)(2n-3)\cdots 5\cdot 3\cdot 1}{n!n!} = \frac{2^n 2^n}{n!} \left(\frac{2n-1}{2}\right) \left(\frac{2n-3}{2}\right) \cdots \left(\frac{5}{2}\right) \left(\frac{3}{2}\right) \left(\frac{1}{2}\right) \\ &= \frac{4^n}{n!} (-1)^n \left(-n + \frac{1}{2}\right) \left(-n + \frac{3}{2}\right) \cdots \left(-\frac{5}{2}\right) \left(-\frac{3}{2}\right) \left(-\frac{1}{2}\right) = (-4)^n \binom{-\frac{1}{2}}{n}. \end{aligned}$$

Entonces, utilizando las probabilidades de transición en n pasos

$$\begin{aligned} \sum_{n=0}^{\infty} p_n t^n &= \sum_{k=0}^{\infty} p_{2k} t^{2k} = \sum_{k=0}^{\infty} \binom{2k}{k} p^k q^k t^{2k} \\ &= \sum_{k=0}^{\infty} (-4)^k \binom{-\frac{1}{2}}{k} p^k q^k t^{2k} = \sum_{k=0}^{\infty} \binom{-\frac{1}{2}}{k} (-4pqt^2)^k = (1 - 4pqt^2)^{-\frac{1}{2}} \end{aligned}$$

por lo que

$$(1 - 4pqt^2)^{-\frac{1}{2}} - 1 = G(t)(1 - 4pqt^2)^{-\frac{1}{2}}$$

por tanto

$$G(t) = 1 - (1 - 4pqt^2)^{\frac{1}{2}}.$$

Utilizando esta última expresión podemos calcular la probabilidad de un eventual regreso al estado 0, por el lema de Abel

$$\begin{aligned} \sum_{n=0}^{\infty} f_n &= \lim_{t \rightarrow 1^-} \sum_{n=0}^{\infty} f_n t^n = \lim_{t \rightarrow 1^-} G(t) = \lim_{t \rightarrow 1^-} 1 - (1 - 4pqt^2)^{\frac{1}{2}} = 1 - (1 - 4pq)^{\frac{1}{2}} \\ &= 1 - (1 - 4p + 4p^2)^{\frac{1}{2}} = 1 - [(1 - 2p)^2]^{\frac{1}{2}} = 1 - |1 - 2p| = 1 - |p - q|. \end{aligned}$$

■

2.3. Caminata aleatoria auto-evasiva

Una *caminata aleatoria auto-evasiva* en \mathbb{Z}^d es una caminata aleatoria donde las auto-intersecciones en las trayectorias están prohibidas, es decir; ningún sitio puede ser ocupado más de una vez.

Para los polímeros, esto corresponde a elegir el conjunto de trayectorias y a la función que asocia la energía como:

- $\mathcal{W}_n = \{w = (w_i)_{i=0}^n \in (\mathbb{Z}^d)^{n+1} : w_0 = 0, \|w_{i+1} - w_i\| = 1 \forall 0 \leq i < n, w_i \neq w_j \forall i \neq j\}$.
- $H_n \equiv 0$.

y a la medida \mathbb{P}_n como la distribución uniforme en \mathcal{W}_n .

Esto describe un polímero auto-evasivo con $(S_i)_{i=0}^n$ tomando valores en \mathcal{W}_n . Notemos que en este caso $(\mathbb{P}_n)_{n \in \mathbb{N}_0}$ no es una familia consistente de medidas de probabilidad.

2.4. Caminata aleatoria auto-evasiva débil

Una *caminata aleatoria auto-evasiva débil* es una variación de la caminata aleatoria auto-evasiva en la cual las auto-intersecciones en las trayectorias no están prohibidas pero son penalizadas, este es el modelo que usaremos para describir a los polímeros aleatorios de los cuales se habla en este trabajo.

Una caminata aleatoria auto-evasiva débil (*polímero suave*) en \mathbb{Z} se define como el proceso aleatorio $(S_i)_{i=0}^n$ con valores en el conjunto de trayectorias y función de energía:

- $\mathcal{W}_n = \{w = (w_i)_{i=0}^n \in \mathbb{Z}^{n+1} : w_0 = 0, |w_{i+1} - w_i| = 1 \forall 0 \leq i < n\}$,
- $H_n(w) = \beta I_n(w)$,

donde $\beta \in [0, \infty)$ y

$$I_n(w) = \sum_{i,j=0:i < j}^n 1_{\{w_i=w_j\}}$$

es el número de auto-intersecciones en la trayectoria.

Escribiremos \mathbb{P}_n^β para denotar la distribución de probabilidad de este proceso de longitud n , con parámetro β , la cual se define como:

$$\mathbb{P}_n^\beta(w) = \frac{1}{Z_n^\beta} e^{-\beta I_n(w)} \mathbb{P}_n(w) \quad , \quad w \in \mathcal{W}_n$$

donde \mathbb{P}_n es la restricción de \mathbb{P} en los primeros n pasos y Z_n^β es una constante normalizadora; de hecho

$$\begin{aligned} \sum_{w \in \mathcal{W}_n} \mathbb{P}_n^\beta(w) = 1 &\quad \Rightarrow \quad \sum_{w \in \mathcal{W}_n} \frac{1}{Z_n^\beta} e^{-\beta I_n(w)} \mathbb{P}_n(w) = 1 \\ \Rightarrow \quad Z_n^\beta = \sum_{w \in \mathcal{W}_n} e^{-\beta I_n(w)} \mathbb{P}_n(w) &\quad \Rightarrow \quad Z_n^\beta = \mathbb{E}_n [e^{-\beta I_n}] \end{aligned}$$

con \mathbb{E}_n denotando la esperanza respecto a \mathbb{P}_n .

Este modelo es llamado *Caminata Aleatoria Auto-evasiva Débil o Modelo de Domb-Joyce*, cada auto-intersección contribuye con un nivel de energía β a H_n y por tanto es penalizada por un factor $e^{-\beta}$; pensaremos a β como la *fuerza de auto-repelencia*.

Entonces, \mathbb{P}_n^β es la distribución del proceso aleatorio $(S_i)_{i=0}^n$ con auto-repelencia débil y valores en \mathcal{W}_n , donde

$$H_n [(S_i)_{i=0}^n] = \beta I [(S_i)_{i=0}^n] \quad \text{y} \quad I_n [(S_i)_{i=0}^n] = \sum_{i,j=0: i < j}^n 1_{\{S_i=S_j\}}.$$

Notemos que $(\mathbb{P}_n^\beta)_{n \in \mathbb{N}_0}$ no es una familia consistente de distribuciones de probabilidad cuando $\beta \in (0, \infty)$; si $\beta = 0$ corresponde a una caminata aleatoria simple y el caso $\beta = \infty$ a una caminata aleatoria auto-evasiva.

2.4.1. Puentes

Bajo esta última distribución de probabilidad y para nuestro análisis obtendremos algunos resultados que nos serán útiles más adelante.

Primero consideraremos que tenemos una trayectoria de n pasos contenida completamente entre sus puntos extremos, es decir; no hace *lazos* a la izquierda de w_0 ni a la derecha de w_n , esto es lo que se conoce como una *trayectoria puente* o simplemente un *puente*.

Ahora, reescribiremos la definición de \mathbb{P}_n^β de una manera más conveniente para nuestro análisis. Sea

$$\widehat{I}_n(w) = \sum_{i,j=0}^n 1_{\{w_i=w_j\}}.$$

Notemos que

$$1_{\{w_i=w_i\}} = 1 \quad \text{para todo } i = 0, 1, \dots, n \quad \text{y} \quad 1_{\{w_i=w_j\}} = 1_{\{w_j=w_i\}},$$

así, se tiene que $\widehat{I}_n(w) = 2I(w) + (n + 1)$, por lo que podemos poner $\widehat{I}_n(w)$ en el factor exponencial que sólo cambia β por 2β .

La siguiente definición es de vital importancia. Sea

$$l_n(x) = \sum_{i=0}^n 1_{\{w_i=x\}}, \quad x \in \mathbb{Z}, \quad n \in \mathbb{N}_0$$

el *número de visitas* o el *tiempo local* en x al tiempo n .

Nótemos que $\sum_{x \in \mathbb{Z}} l_n(x) = n + 1$ para todo $n \in \mathbb{N}$ y que $\widehat{I}_n(w) = \sum_{x \in \mathbb{Z}} l_n^2(x)$

$$\begin{aligned} \widehat{I}_n(w) &= \sum_{i,j=0}^n 1_{\{w_i=w_j\}} = \sum_{x \in \mathbb{Z}} \sum_{i,j=0}^n 1_{\{w_i=w_j=x\}} = \sum_{x \in \mathbb{Z}} \sum_{i,j=0}^n 1_{\{w_i=w_j=x\}} \\ &= \sum_{x \in \mathbb{Z}} \sum_{i,j=0}^n 1_{\{w_i=x\}} 1_{\{w_j=x\}} = \sum_{x \in \mathbb{Z}} \left(\sum_{i=0}^n 1_{\{w_i=x\}} \sum_{j=0}^n 1_{\{w_j=x\}} \right) \\ &= \sum_{x \in \mathbb{Z}} l_n(x) l_n(x) = \sum_{x \in \mathbb{Z}} l_n^2(x). \end{aligned}$$

2.4.2. *Desdoblado una trayectoria en un puente*

Este resultado muestra que la condición de puente no altera la constante normalizadora.

LEMA. Para $n \rightarrow \infty$,

$$\mathbb{E} \left[e^{-\beta \widehat{I}_n} 1_{\{S_n \geq 0\}} \right] = e^{o(n)} \mathbb{E} \left[e^{-\beta \widehat{I}_n} 1_{A_n} \right]$$

donde $\widehat{I}_n = \widehat{I}_n((S_i)_{i=0}^n)$ y $A_n = \{S_0 \leq S_i \leq S_n, \forall 0 \leq i \leq n\}$.

DEMOSTRACIÓN. Sea n fija y supongamos que nuestra trayectoria es un *medio-puente* a la derecha, es decir; $S_i > S_0$ para todo $0 < i \leq n$. Aplicaremos un proceso de reflexión empezando desde el punto final izquierdo de la trayectoria como sigue:

Sea $i_0 = 0$ y para $j = 1, 2, \dots$, definimos (R_j, i_j) recursivamente como

$$R_j = \max_{i_{j-1} < i < n} (-1)^j (S_{i_{j-1}} - S_i),$$

i_j = el índice i donde se alcanza el máximo.

La recursión se detiene en el menor entero k tal que $i_k = n$. Lo que esta definición establece es que R_j es el *espacio* que ocupa la caminata $(S_{i_{j-1}}, \dots, S_n)$.

Cada subcaminata $(S_{i_{j-1}}, \dots, S_{i_j})$ se encuentra estrictamente de un lado del punto $S_{i_{j-1}}$ y

$$R_1 + \dots + R_k \leq n \text{ y } R_1 > R_2 > \dots > R_k \geq 1.$$

Si para $j = 1, 2, \dots, k-1$ reflejamos (S_{i_j}, \dots, S_n) al rededor de S_{i_j} , la trayectoria se convierte en un puente, es decir; en una trayectoria que cumple que $S_0 \leq S_i \leq S_n, \forall 0 < i \leq n$, más aún; este puente es menos penalizado que la trayectoria original porque tiene menos auto-intersecciones.

Ahora quitaremos la condición de medio puente y sólo supondremos que $S_n \geq 0$.

Sean

$$i_- \min \left\{ 0 \leq i \leq n \mid S_i = \min_{0 \leq j \leq n} S_j \right\}$$

$$i_+ \text{ máx} \left\{ 0 \leq i \leq n \mid S_i = \text{máx}_{0 \leq j \leq n} S_j \right\}$$

Cuando $i_- > 0$ y $i_+ < n$, ambas subcaminatas (S_0, \dots, S_{i_-}) y (S_{i_+}, \dots, S_n) son medios puentes y se puede aplicar el proceso anterior. Si desdoblamos ambas piezas hacia afuera de la trayectoria, al final obtenemos un puente (los casos $i_- = 0$ e $i_+ = n$ no necesitan reflexión) y como la trayectoria *convertida* en puente (w_p) tiene menos auto-intersecciones que la trayectoria original (w) tenemos que

$$e^{-\beta \widehat{I}_n(w)} \leq e^{-\beta \widehat{I}_n(w_p)}$$

y podemos concluir que

$$\mathbb{E} \left[e^{-\beta \widehat{I}_n} 1_{\{S_n \geq 0\}} \right] \leq N_n^2 \mathbb{E} \left[e^{-\beta \widehat{I}_n} 1_{A_n} \right]$$

donde N_n es el número de soluciones de

$$R_1 + \dots + R_k \leq n \text{ y } R_1 > R_2 > \dots > R_k \geq 1.$$

sumadas sobre k (el número de particiones ordenadas del conjunto $\{1, 2, \dots, n\}$) y $N_n = e^{O(\sqrt{n})}$ ([7] *Madras and Slade - The Self-Avoiding Walk*, teorema 3.1.4), el cual es un factor inofensivo y se sigue el resultado. ■

Capítulo 3

Grandes Desviaciones

En este capítulo presentamos los resultados sobre grandes desviaciones que utilizaremos más adelante, para esto nos basamos en [1] y [2].

Sea $(X_n)_{n \in \mathbb{N}}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas en un espacio de probabilidad $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$, con

$$\mathbb{E}[X_i] = \mu \in \mathbb{R} \quad \text{y} \quad \text{Var}(X_i) = \sigma^2 \in (0, \infty)$$

donde $\mathcal{B}(\mathbb{R})$ es la σ -álgebra de Borel sobre \mathbb{R} y \mathbb{E} denota la esperanza respecto a \mathbb{P} .

Denotemos por $S_n = \sum_{i=1}^n X_i$ a la n -ésima suma parcial y recordemos dos de los resultados más importantes en teoría de probabilidad:

- *Ley de los Grandes Números*

$$\mathbb{P} \left(\lim_{n \rightarrow \infty} \frac{1}{n} S_n = \mu \right) = 1$$

- *Teorema del Límite Central*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq a \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-\frac{x^2}{2}} dx$$

Mientras que la ley de los grandes números afirma que el promedio empírico $\frac{1}{n}S_n$ converge a μ cuando $n \rightarrow \infty$, el teorema de límite central aproxima la probabilidad de que S_n difiera de $n\mu$ por un factor de orden \sqrt{n} ; desviaciones de este tipo son llamadas *normales*.

En teoría de grandes desviaciones se trabaja con eventos donde S_n difiere de $n\mu$ por un factor de orden n ; un ejemplo es el evento:

$$\{S_n \geq (\mu + a)n\}, \quad a > 0$$

cuya probabilidad tiende a cero cuando $n \rightarrow \infty$. El objetivo de esta Teoría es inferir la *velocidad* a la que esto ocurre.

Resulta que bajo cierta condición en la cola de la distribución de las X_i 's, el decaimiento de la probabilidad es exponencial cuando n crece, es decir;

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq (\mu + a)n) = -I(a) < 0, \quad a > 0.$$

Donde $I(a)$ es conocida como la *función tasa*, esta función es fundamental para un correcto y completo análisis del comportamiento antes mencionado.

Este comportamiento es porque generalmente algunos de los componentes X_1, X_2, \dots, X_n en la suma S_n deben *desviarse* de μ para crear la *gran desviación* an 'lejos' de $n\mu$.

3.1. Un *volado*

TEOREMA. Sea $(X_n)_{n \in \mathbb{N}}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con función de densidad de probabilidad:

$$\mathbb{P}(X_i = 0) = \frac{1}{2} = \mathbb{P}(X_i = 1)$$

es decir, X_i tiene distribución Bernoulli con parámetro $p = \frac{1}{2}$.

Sea $S_n = \sum_{i=1}^n X_i$, entonces; $\forall a > \frac{1}{2}$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq an) = -I(a)$$

donde

$$I(z) = \begin{cases} \log 2 + z \log z + (1 - z) \log(1 - z) & \text{si } z \in [0, 1] \\ \infty & \text{en otro caso} \end{cases}$$

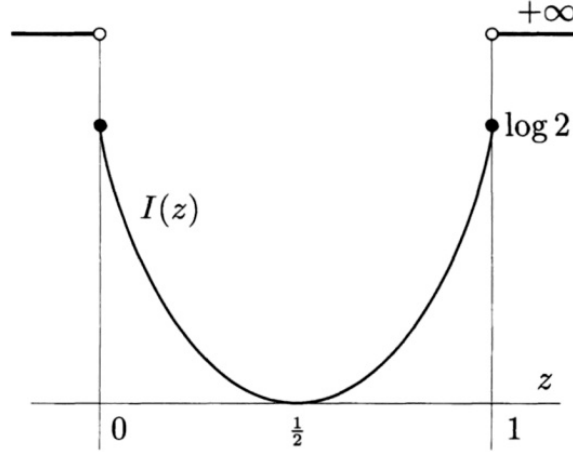


Figura 3.1: Función tasa para un *volado* [1].

DEMOSTRACIÓN. Como las X_i 's son independientes, S_n tiene una distribución *Binomial* de parámetros n y $p = \frac{1}{2}$; es decir

$$\mathbb{P}(S_n = x) = \binom{n}{x} \left(\frac{1}{2}\right)^x \left(1 - \frac{1}{2}\right)^{n-x} = \binom{n}{x} 2^{-n}$$

entonces

$$\mathbb{P}(S_n \geq an) = \sum_{x=an}^n \binom{n}{x} 2^{-n}$$

Si $a > 1$, $\mathbb{P}(S_n \geq an) = 0$; por lo que: $\log \mathbb{P}(S_n \geq an) = \log 0 = -\infty$ y

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq an) = -\infty = -I(a)$$

Sea $a \in (\frac{1}{2}, 1]$, note que:

$$2^{-n} Q_n(a) \leq \sum_{x=an}^n \binom{n}{x} 2^{-n} \leq (n+1) 2^{-n} Q_n(a)$$

es decir

$$2^{-n}Q_n(a) \leq \mathbb{P}(S_n \geq an) \leq (n+1)2^{-n}Q_n(a)$$

donde $Q_n(a) = \max_{k \geq an} \binom{n}{k}$.

OBSERVACIÓN.

$$Q_n(a) = \binom{n}{k}, \text{ donde } k = \lceil an \rceil \text{ y } \lceil an \rceil = \min\{z \in \mathbb{Z}: z \geq an\}.$$

La *Fórmula de Stirling*

$$\forall n \in \mathbb{N}, \quad n! = n^n e^{-n} \sqrt{2\pi n} \left(1 + O\left(\frac{1}{n}\right)\right),$$

nos permite deducir que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log Q_n(a) = -a \log a - (1-a) \log(1-a)$$

entonces

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log 2^{-n} Q_n(a) &= -(\log 2 + a \log a + (1-a) \log(1-a)) \\ \lim_{n \rightarrow \infty} \frac{1}{n} \log (n+1) 2^{-n} Q_n(a) &= -(\log 2 + a \log a + (1-a) \log(1-a)) \end{aligned}$$

por lo que

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq an) &= -(\log 2 + a \log a + (1-a) \log(1-a)) \\ \therefore \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq an) &= -I(a) \end{aligned}$$

■

Como $\mathbb{E}[X_i] = \frac{1}{2}$ y $a > \frac{1}{2}$, el teorema anterior trata con grandes desviaciones *por arriba*, y es claro que por la relación de simetría $I(1-z) = I(z)$ se sigue un resultado análogo para $\mathbb{P}(S_n \leq an)$ con $a < \frac{1}{2}$.

Note que la función tasa I es finita y estrictamente convexa en $[0, 1]$ e infinita fuera de $[0, 1]$; tiene un único cero en $z = \frac{1}{2}$ el cual corresponde con la ley de los grandes números.

3.2. Teorema de Cramér

El siguiente teorema identifica el comportamiento desde el punto de vista de las grandes desviaciones del promedio empírico $\frac{1}{n}S_n$ bajo cierta condición en la cola de la distribución de X_i .

TEOREMA. Sea $(X_n)_{n \in \mathbb{N}}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas en \mathbb{R} con

$$\varphi(t) = \mathbb{E} [e^{tX_i}] < \infty, \quad \forall t \in \mathbb{R}$$

Sea $S_n = \sum_{i=1}^n X_i$, entonces; $\forall a > \mathbb{E} [X_i]$:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} (S_n \geq an) = -I(a)$$

donde

$$I(z) = \sup_{t \in \mathbb{R}} [zt - \log \varphi(t)].$$

DEMOSTRACIÓN. Note que $\forall a \in \mathbb{R}$, la transformación $T_a(X) = X + a$ implica que

$$\varphi_a(t) = \mathbb{E} [e^{t(X_i+a)}] = \mathbb{E} [e^{tX_i} e^{ta}] = e^{ta} \mathbb{E} [e^{tX_i}] = e^{ta} \varphi(t)$$

y en consecuencia

$$\begin{aligned} I_a(a) &= \sup_{t \in \mathbb{R}} [at - \log \varphi_a(t)] = \sup_{t \in \mathbb{R}} [at - \log e^{ta} \varphi(t)] = \sup_{t \in \mathbb{R}} [at - at - \log \varphi(t)] \\ &= \sup_{t \in \mathbb{R}} [0 - \log \varphi(t)] = I(0) \end{aligned}$$

Por lo que, sin pérdida de generalidad; podemos suponer que $a = 0$ y que $\mathbb{E} [X_i] < 0$, además de que X_i es no-degenerada, es decir; no toma sólo un valor.

Sea $\rho = \inf_{t \in \mathbb{R}} \varphi(t)$ y note que

$$I(0) = \sup_{t \in \mathbb{R}} [-\log \varphi(t)] = -\inf_{t \in \mathbb{R}} [\log \varphi(t)] = -\log \inf_{t \in \mathbb{R}} \varphi(t) = -\log \rho$$

con $I(0) = \infty$ si $\rho = 0$.

Basta demostrar entonces que:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq 0) = \log \rho$$

Sea $F_{X_i}(x) = \mathbb{P}(X_i \leq x)$ la función de distribución de X_i , se sigue del hecho de que $\varphi(t) = \mathbb{E}[e^{tX_i}] < \infty$ que $\varphi \in C^\infty(\mathbb{R})$ con

$$\varphi'(t) = \int_{\mathbb{R}} x e^{tx} dF(x) \quad \text{y} \quad \varphi''(t) = \int_{\mathbb{R}} x^2 e^{tx} dF(x) > 0$$

por lo que φ es estrictamente convexa y $\varphi'(0) = \mathbb{E}[X_i] < 0$.

De acuerdo a donde se puede situar la densidad de \mathbb{P} tenemos tres casos:

1. $\mathbb{P}(X_i < 0) = 1$

Sea $x < 0$ y $(t_n)_{n \in \mathbb{N}}$ una sucesión de números positivos en \mathbb{R} tal que $t_n < t_{n+1}$ y $\lim_{n \rightarrow \infty} t_n = \infty$, entonces:

$$e^{t_{n+1}x} < e^{t_n x} \quad \text{lo cual implica que } \varphi \text{ es estrictamente decreciente.}$$

Además

$$0 < e^{t_n x} \leq 1 \quad \text{y} \quad \lim_{n \rightarrow \infty} e^{t_n x} = 0$$

por el *teorema de la convergencia dominada de Lebesgue*:

$$\lim_{n \rightarrow \infty} \int_{-\infty}^0 e^{t_n x} dF(x) = \int_{-\infty}^0 \lim_{n \rightarrow \infty} e^{t_n x} dF(x) = 0$$

y como esto sucede para toda $x < 0$ y para toda $(t_n)_{n \in \mathbb{N}}$ sucesión creciente en \mathbb{R}

$$\lim_{t \rightarrow \infty} \int_{-\infty}^0 e^{tx} dF(x) = 0 \quad \text{por lo que} \quad \rho = \inf_{t \in \mathbb{R}} \varphi(t) = \lim_{t \rightarrow \infty} \varphi(t) = 0.$$

Por otro lado $\mathbb{P}(S_n \geq 0) = 0$, por tanto $\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq 0) = -\infty$

$$\therefore \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq 0) = \log \rho.$$

2. $\mathbb{P}(X_i \leq 0) = 1$ y $\mathbb{P}(X_i = 0) > 0$

De la misma manera φ es estrictamente decreciente y

$$\rho = \lim_{t \rightarrow \infty} \varphi(t) = \lim_{t \rightarrow \infty} \int_{-\infty}^{0^-} e^{tx} dF(x) + e^{0t} \mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 0) > 0.$$

Además

$$\mathbb{P}(S_n \geq 0) = \mathbb{P}\left(\bigcap_{i=1}^n \{X_i = 0\}\right) = \prod_{i=1}^n \mathbb{P}(X_i = 0) = \rho^n$$

por lo que

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq 0) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \rho^n \\ \therefore \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq 0) &= \log \rho. \end{aligned}$$

3. $\mathbb{P}(X_i < 0) > 0$ y $\mathbb{P}(X_i > 0) > 0$

En este caso

$$\lim_{t \rightarrow \infty} \varphi(t) = \infty = \lim_{t \rightarrow -\infty} \varphi(t)$$

y como φ es estrictamente convexa, existe un único $\tau \in \mathbb{R}$ tal que:

$$\tau > 0^1, \quad \varphi(\tau) = \rho \text{ y } \varphi'(\tau) = 0$$

por la *Desigualdad de Markov*²:

$$\begin{aligned} \mathbb{P}(S_n \geq 0) &= \mathbb{P}(e^{\tau S_n} \geq e^{\tau 0}) = \mathbb{P}(e^{\tau S_n} \geq 1) \leq \mathbb{E}[e^{\tau S_n}] \\ &= \mathbb{E}\left[\prod_{i=1}^n e^{\tau X_i}\right] = \prod_{i=1}^n \mathbb{E}[e^{\tau X_i}] = \prod_{i=1}^n \varphi(\tau) = \varphi^n(\tau) = \rho^n \end{aligned}$$

por lo que

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq 0) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \rho^n = \log \rho.$$

¹ $\varphi'(0) = \mathbb{E}[X_i] < 0$ implica que $\tau > 0$, es decir; φ 'sigue decreciendo' en 0.

²También llamada *Desigualdad de Chebyshev*.

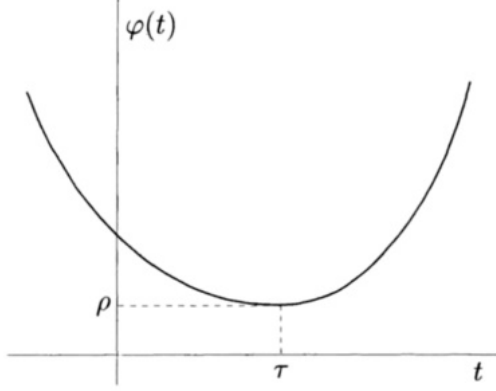


Figura 3.2: Relación entre φ , τ y ρ [1].

Por otro lado, sea $(\widehat{X}_n)_{n \in \mathbb{N}}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con función de distribución

$$\widehat{F}(x) = \frac{1}{\rho} \int_{-\infty}^x e^{\tau y} dF(y)$$

\widehat{F} es llamada la *Transformación de Cramér de F*.

Note que $d\widehat{F}(x) = \frac{1}{\rho} e^{\tau x} dF(x)$ entonces, $dF(x) = \rho e^{-\tau x} d\widehat{F}(x)$.

La cota inferior se obtiene a partir de tres lemas:

LEMA 1. $\mathbb{E}[\widehat{X}_i] = \widehat{\mu} = 0$ y $Var(\widehat{X}_i) = \widehat{\sigma}^2 \in (0, \infty)$.

DEMOSTRACIÓN. Sea $\widehat{\varphi}(t) = \mathbb{E} [e^{t\widehat{X}}]$, entonces

$$\begin{aligned} \widehat{\varphi}(t) &= \int_{\mathbb{R}} e^{tx} d\widehat{F}(x) = \frac{1}{\rho} \int_{\mathbb{R}} e^{tx} e^{\tau x} dF(x) \\ &= \frac{1}{\rho} \int_{\mathbb{R}} e^{x(t+\tau)} dF(x) = \frac{1}{\rho} \varphi(t + \tau) < \infty \end{aligned}$$

entonces $\widehat{\varphi} \in C^\infty(\mathbb{R})$, por tanto

$$\mathbb{E}[\widehat{X}_i] = \widehat{\varphi}'(0) = \frac{1}{\rho}\varphi'(0 + \tau) = \frac{1}{\rho}\varphi'(\tau) = 0$$

$$\text{Var}(\widehat{X}_i) = \mathbb{E}[\widehat{X}_i^2] = \widehat{\varphi}''(0) = \frac{1}{\rho}\varphi''(0 + \tau) = \frac{1}{\rho}\varphi''(\tau) \in (0, \infty)$$

□

LEMA 2. Sea $\widehat{S}_n = \sum_{i=1}^n \widehat{X}_i$, entonces $\mathbb{P}(S_n \geq 0) = \rho^n \mathbb{E} \left[e^{-\tau \widehat{S}_n} 1_{\{\widehat{S}_n \geq 0\}} \right]$.

DEMOSTRACIÓN.

$$\begin{aligned} \mathbb{P}(S_n \geq 0) &= \int_{\{x_1+x_2+\dots+x_n \geq 0\}} dF(x_1)dF(x_2)\cdots dF(x_n) \\ &= \int_{\{x_1+x_2+\dots+x_n \geq 0\}} \left(\rho e^{-\tau x_1} d\widehat{F}(x_1) \right) \left(\rho e^{-\tau x_2} d\widehat{F}(x_2) \right) \cdots \left(\rho e^{-\tau x_n} d\widehat{F}(x_n) \right) \\ &= \rho^n \int_{\{x_1+x_2+\dots+x_n \geq 0\}} e^{-\tau(x_1+x_2+\dots+x_n)} d\widehat{F}(x_1)d\widehat{F}(x_2)\cdots d\widehat{F}(x_n) \\ &\Rightarrow \mathbb{P}(S_n \geq 0) = \rho^n \mathbb{E} \left[e^{-\tau \widehat{S}_n} 1_{\{\widehat{S}_n \geq 0\}} \right]. \end{aligned}$$

□

LEMA 3. $\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} \left[e^{-\tau \widehat{S}_n} 1_{\{\widehat{S}_n \geq 0\}} \right] \geq 0$.

DEMOSTRACIÓN. Sea $n \in \{k \in \mathbb{N} : \widehat{S}_k \geq 0\}$, por el LEMA 1 podemos aplicar el teorema del límite central a \widehat{S}_n , es decir;

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\widehat{S}_n}{\widehat{\sigma}\sqrt{n}} \leq \alpha \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-\frac{x^2}{2}} dx$$

elegimos α tal que³

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-\frac{x^2}{2}} dx \geq \frac{3}{4}$$

³ $\alpha > 0,675$

y aplicando la *desigualdad de Markov*

$$\mathbb{P}\left(\frac{\widehat{S}_n}{\widehat{\sigma}\sqrt{n}} \leq \alpha\right) = \mathbb{P}\left(e^{-\tau\widehat{S}_n} \geq e^{-\tau\widehat{\sigma}\sqrt{n}\alpha}\right) \leq \frac{\mathbb{E}\left[e^{-\tau\widehat{S}_n}\right]}{e^{-\tau\widehat{\sigma}\sqrt{n}\alpha}}$$

entonces

$$e^{-\tau\widehat{\sigma}\sqrt{n}\alpha} \mathbb{P}\left(\frac{\widehat{S}_n}{\widehat{\sigma}\sqrt{n}} \leq \alpha\right) \leq \mathbb{E}\left[e^{-\tau\widehat{S}_n}\right],$$

por lo que

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \left[e^{-\tau\widehat{\sigma}\sqrt{n}\alpha} \mathbb{P}\left(\frac{\widehat{S}_n}{\widehat{\sigma}\sqrt{n}} \leq \alpha\right) \right] \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}\left[e^{-\tau\widehat{S}_n}\right]$$

y además

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \left[e^{-\tau\widehat{\sigma}\sqrt{n}\alpha} \mathbb{P}\left(\frac{\widehat{S}_n}{\widehat{\sigma}\sqrt{n}} \leq \alpha\right) \right] = 0$$

por tanto

$$0 \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}\left[e^{-\tau\widehat{S}_n} 1_{\{\widehat{S}_n \geq 0\}}\right].$$

□

De esta manera, por los LEMAS 2 y 3 se tiene que

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq 0) = \liminf_{n \rightarrow \infty} \frac{1}{n} \log \rho^n \mathbb{E}\left[e^{-\tau\widehat{S}_n} 1_{\{\widehat{S}_n \geq 0\}}\right]$$

$$= \log \rho + \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}\left[e^{-\tau\widehat{S}_n} 1_{\{\widehat{S}_n \geq 0\}}\right] \geq \log \rho$$

$$\therefore \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq 0) = \log \rho.$$

■

3.3. Definiciones generales

En esta sección presentamos las definiciones básicas en Teoría de grandes desviaciones.

Sea \mathcal{X} un espacio métrico completo y separable con distancia

$$d : \mathcal{X} \times \mathcal{X} \rightarrow [0, \infty)$$

DEFINICIÓN. Una función $f : \mathcal{X} \rightarrow [-\infty, \infty]$ es *semicontinua inferiormente* si satisface alguna de las siguientes propiedades:

- (i) $\liminf_{n \rightarrow \infty} f(x_n) \geq f(x)$, $\forall (x_n)$, x tal que $x_n \rightarrow x$ en \mathcal{X} cuando $n \rightarrow \infty$.
- (ii) $\lim_{\epsilon \rightarrow 0^+} \inf_{y \in B_\epsilon(x)} f(y) = f(x)$ con $B_\epsilon(x) = \{y \in \mathcal{X} : d(x, y) < \epsilon\}$.
- (iii) f tiene *conjuntos de nivel* cerrados, i.e.,
 $f^{-1}([-\infty, c]) = \{x \in \mathcal{X} : f(x) \leq c\}$ es cerrado para toda $c \in \mathbb{R}$.

Las definiciones anteriores son equivalentes.

DEFINICIÓN. La función $I : \mathcal{X} \rightarrow [0, \infty]$ es llamada *función tasa* si

T1 $I \not\equiv \infty$.

T2 I es semicontinua inferiormente.

T3 I tiene conjuntos de nivel compactos.

DEFINICIÓN. Decimos que una sucesión de medidas de probabilidad (\mathbb{P}_n) en \mathcal{X} satisface el *principio de grandes desviaciones* con tasa n y función tasa I si

P1 I es una función tasa en el sentido de la definición anterior.

P2 $\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n(C) \leq -I(C)$, $\forall C \subset \mathcal{X}$ cerrado.

P3 $\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n(O) \geq -I(O)$, $\forall O \subset \mathcal{X}$ abierto.

Las cotas están en términos de la función conjuntista definida por

$$I(S) = \inf_{x \in S} I(x), \quad S \subset \mathcal{X}$$

El objetivo de esta teoría es construir una gran cantidad de teoremas basados en estas dos definiciones.

TEOREMA. *Sea (\mathbb{P}_n) una sucesión de medidas de probabilidad que satisface el principio de grandes desviaciones, entonces; la función tasa I asociada, es única.*

DEMOSTRACIÓN. Sean I y J funciones tasa para (\mathbb{P}_n) , debemos demostrar que $I(x) = J(x)$ para todo $x \in \mathcal{X}$.

Sea $x \in \mathcal{X}$ y sea $B_N = B_{\frac{1}{N}}(x)$ la bola con centro en x y radio $\frac{1}{N}$, con $N \in \mathbb{N}$, entonces; por la definición de función tasa

$$\begin{aligned} -I(x) &\leq -I(B_{N+1}) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n(B_{N+1}) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n(\text{cl}(B_{N+1})) \leq -J(\text{cl}(B_{N+1})) \leq -J(B_N) \end{aligned}$$

donde el orden de las desigualdades se da por los siguientes hechos

1. $I(B_{N+1}) = \inf_{y \in B_{N+1}} I(y)$, $I(B_{N+1}) \leq I(y) \quad \forall y \in B_{N+1}$ y entonces $-I(y) \leq -I(B_{N+1})$.

2. **P3.**

- 3.

$$B_{N+1} \subseteq \text{cl}(B_{N+1}) \quad \text{por lo que} \quad \mathbb{P}_n(B_{N+1}) \leq \mathbb{P}_n(\text{cl}(B_{N+1}))$$

$$\frac{1}{n} \log \mathbb{P}_n(B_{N+1}) \leq \frac{1}{n} \log \mathbb{P}_n(\text{cl}(B_{N+1}))$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n(B_{N+1}) \leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n(\text{cl}(B_{N+1}))$$

y por tanto

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n(B_{N+1}) &\leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n(B_{N+1}) \\ &\leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n(\text{cl}(B_{N+1})) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n(\text{cl}(B_{N+1})). \end{aligned}$$

4. P2.

5. Como $B_{N+1} \subset B_N$ entonces $\text{cl}(B_{N+1}) \subseteq B_N$.

Por otro lado

$$J(B_N) = \inf_{y \in B_N} J(y) \text{ y } \lim_{N \rightarrow \infty} J(B_N) = \lim_{N \rightarrow \infty} \inf_{y \in B_N} J(y) = J(x)$$

ya que J es semicontinua inferiormente y $\lim_{N \rightarrow \infty} B_N(x) = x$, entonces $-I(x) \leq -J(x)$, es decir $I(x) \geq J(x)$ y de la misma manera $J(x) \geq I(x)$ por lo que $I(x) = J(x)$ para todo $x \in \mathcal{X}$ y por tanto I es única. ■

3.4. Lema de Varadhan

En esta sección enunciamos uno de los resultados importantes de la teoría de grandes desviaciones.

Decimos que dos sucesiones de números positivos (α_n) y (β_n) son *logarítmicamente equivalentes* $\alpha_n \simeq \beta_n$ si y sólo si

$$\lim_{n \rightarrow \infty} \frac{1}{n} (\log \alpha_n - \log \beta_n) = 0.$$

OBSERVACIÓN. $\alpha_n + \beta_n \simeq \alpha_n \vee \beta_n$.

TEOREMA. (Lema de Varadhan) Sea (\mathbb{P}_n) una sucesión de medidas de probabilidad que satisface el principio de grandes desviaciones en \mathcal{X} con tasa n y función tasa I . Sea $F : \mathcal{X} \rightarrow \mathbb{R}$ una función continua acotada por arriba; entonces

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{\mathcal{X}} e^{nF(x)} \mathbb{P}_n(dx) = \sup_{x \in \mathcal{X}} [F(x) - I(x)]$$

DEMOSTRACIÓN. Sea

$$J_n(S) = \int_S e^{nF(x)} \mathbb{P}_n(dx), \quad S \subset \mathcal{X} \text{ boreliano,}$$

y hacemos

$$a = \sup_{x \in \mathcal{X}} F(x), \quad b = \sup_{x \in \mathcal{X}} [F(x) - I(x)].$$

Notemos que $-\infty < b \leq a < \infty$, porque $I \geq 0$ y F es continua y acotada por arriba. Probaremos que existen cotas superior e inferior.

· Cota superior:

Particionamos al espacio \mathcal{X} de acuerdo a los valores que toma F ; sea $C = F^{-1}([b, a])$, y para $N \in \mathbb{N}$ definimos los conjuntos

$$C_j^N = F^{-1}([c_{j-1}^N, c_j^N]), \quad j = 1, \dots, N$$

donde $c_j^N = b + \frac{j}{N}(a - b)$ para $j = 0, 1, \dots, N$, es claro que

$$C = \bigcup_{j=1}^N C_j^N.$$

Ahora, como F es continua, todos los C_j^N 's son cerrados en \mathcal{X} , entonces; por el principio de grandes desviaciones

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n(C_j^N) \leq -I(C_j^N), \quad \forall j.$$

Si tomamos en cuenta que $F(x) \leq c_j^N$ en C_j^N , obtenemos la desigualdad

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log J_n(C) \leq \max_{1 \leq j \leq N} [c_j^N - I(C_j^N)]$$

lo cual puede desarrollarse más utilizando la desigualdad

$$c_j^N \leq \inf_{x \in C_j^N} F(x) + \frac{1}{N}(a - b)$$

para obtener

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log J_n(C) &\leq \max_{1 \leq j \leq N} \left\{ \inf_{x \in C_j^N} F(x) - \inf_{x \in C_j^N} I(x) \right\} + \frac{1}{N}(a - b) \\ &\leq \max_{1 \leq j \leq N} \sup_{x \in C_j^N} [F(x) - I(x)] + \frac{1}{N}(a - b) = \sup_{x \in C} [F(x) - I(x)] + \frac{1}{N}(a - b) \\ &\leq b + \frac{1}{N}(a - b) \end{aligned}$$

Haciendo $N \rightarrow \infty$ tenemos que

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log J_n(C) \leq b = \sup_{x \in \mathcal{X}} [F(x) - I(x)]$$

· Cota inferior:

Sea $x \in \mathcal{X}$ y $\epsilon > 0$, entonces por la continuidad de F , el conjunto

$$O_{x,\epsilon} = \{y \in \mathcal{X} : F(y) > F(x) - \epsilon\}$$

es una vecindad abierta de x . Se sigue por el principio de grandes desviaciones que

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n(O_{x,\epsilon}) \geq -I(O_{x,\epsilon}),$$

como $I(O_{x,\epsilon}) \leq I(x)$, tenemos que

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n(O_{x,\epsilon}) \geq F(x) - \epsilon - I(x),$$

usando el hecho de que $J_n(\mathcal{X}) \geq J_n(O_{x,\epsilon})$, haciendo $\epsilon \rightarrow 0$ y tomando el supremo sobre todas las $x \in \mathcal{X}$ obtenemos

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log J_n(\mathcal{X}) \geq b = \sup_{x \in \mathcal{X}} [F(x) - I(x)],$$

por lo que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log J_n(\mathcal{X}) = \sup_{x \in \mathcal{X}} [F(x) - I(x)].$$

■

Existe una versión del principio de grandes desviaciones en el que algunas de las condiciones en la definición anterior son debilitadas.

DEFINICIÓN. La función $I : \mathcal{X} \rightarrow [0, \infty]$ es llamada *función tasa débil* si

- $I \not\equiv \infty$.
- I es semicontinua inferiormente.
- I tiene conjuntos de nivel cerrados.

DEFINICIÓN. Decimos que una sucesión de medidas de probabilidad (\mathbb{P}_n) en \mathcal{X} satisface el *principio de grandes desviaciones débil* con tasa n y función tasa I si

- I es una función tasa débil en el sentido de la definición anterior.
- $\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n(K) \leq -I(K), \forall K \subset \mathcal{X}$ compacto.
- $\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n(O) \geq -I(O), \forall O \subset \mathcal{X}$ abierto.

Capítulo 4

Polímeros Aleatorios Suaves

A continuación consideraremos a la caminata aleatoria auto-evasiva débil cuya definición se dió en el capítulo 2, donde las auto-intersecciones no están prohibidas pero son penalizadas, nos referimos a esto como un *polímero suave*.

Probaremos que el polímero suave tiene un comportamiento *balístico* en $d = 1$, es decir; bajo la distribución \mathbb{P}_n^β las trayectorias están cerca del origen (su punto inicial) por un tiempo y después tienden a infinito a una velocidad estrictamente positiva debido a la auto-repelencia.

La prueba utiliza una representación *Markoviana* del número de visitas a cada estado (los tiempos locales totales) de una caminata aleatoria simple, en combinación con teoría de grandes desviaciones; una técnica poderosa que se utiliza para muchos otros modelos en dimensiones más altas. La estructura del capítulo es la siguiente: se define el modelo, se establece el resultado principal, un principio de grandes desviaciones para la localización del punto final derecho, se enuncia un programa de cinco pasos para probar el principio de grandes desviaciones para *polímeros-puente* y por último se remueve esta condición.

4.1. Polímeros con auto-repelencia

De la misma manera que para la caminata aleatoria auto-evasiva débil, un *polímero suave* en \mathbb{Z} de longitud n se define como el proceso aleatorio $(S_i)_{i=0}^n$ con valores en el conjunto de trayectorias y función de energía:

- $\mathcal{W}_n = \{w = (w_i)_{i=0}^n \in \mathbb{Z}^{n+1} : w_0 = 0, |w_{i+1} - w_i| = 1 \ \forall 0 \leq i < n\}$,
- $H_n(w) = \beta I_n(w)$,

donde $\beta \in [0, \infty)$,

$$I_n(w) = \sum_{i,j=0:i<j}^n 1_{\{w_i=w_j\}}$$

es el número de auto-intersecciones en la trayectoria y

$$\mathbb{P}_n^\beta(w) = \frac{1}{Z_n^\beta} e^{-\beta I_n(w)} \mathbb{P}_n(w) \ , \ w \in \mathcal{W}_n$$

Bajo este modelo, existen muchas preguntas que se pueden establecer respecto al comportamiento de estos procesos [1], pero en este trabajo sólo se exponen resultados relacionados con el comportamiento asintótico de

$|S_n|$ = la distancia entre los puntos finales de un polímero de longitud n

bajo la distribución \mathbb{P}_n^β en el límite cuando $n \rightarrow \infty$ y para un parámetro β fijo.

Como mencionamos previamente, podemos esperar que las trayectorias bajo la distribución \mathbb{P}_n^β estén cerca del origen por un tiempo y después tiendan a infinito a una velocidad estrictamente positiva debido a la auto-repelencia, esto es llamado *comportamiento balístico* el cual fue probado por primera vez por E. Bolthausen (*One self-repellent one-dimensional random walks*).

TEOREMA 1. *Para todo $\beta \in (0, \infty)$ existe $\theta^*(\beta) \in (0, 1)$ tal que*

$$\lim_{n \rightarrow \infty} \mathbb{P}_n^\beta \left(\left| \frac{1}{n} S_n - \theta^*(\beta) \right| < \epsilon \mid S_n \geq 0 \right) = 1 \ , \ \forall \epsilon > 0.$$

TEOREMA 2. *La función $\beta \rightarrow \theta^*(\beta)$ se puede obtener en términos de un problema variacional y se sigue de la solución al mismo que:*

- $\beta \rightarrow \theta^*(\beta)$ es analítica en $(0, \infty)$,
- $\lim_{\beta \rightarrow 0^+} \theta^*(\beta) = 0$,
- $\lim_{\beta \rightarrow \infty} \theta^*(\beta) = 1$.

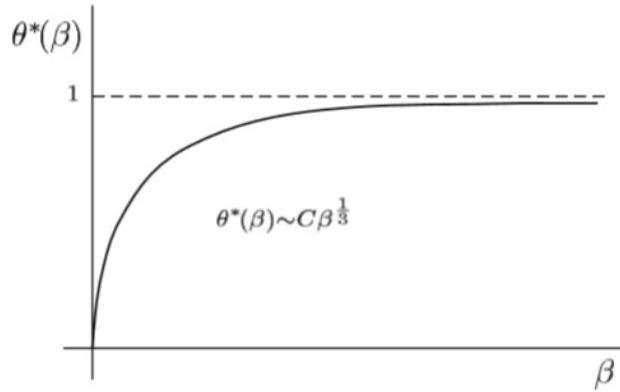


Figura 4.1: Velocidad lineal de un polímero suave [3].

La cantidad $\theta^*(\beta)$ es la *velocidad del polímero con fuerza de auto-repelencia* β (Fig. 4.1).

El teorema 1 se sigue del siguiente principio de grandes desviaciones:

TEOREMA 3. Para toda $\beta \in (0, \infty)$ la familia $(\mathbb{P}_n^{+, \beta})_{n \in \mathbb{N}_0}$ definida por

$$\mathbb{P}_n^{+, \beta}(\cdot) = \mathbb{P}_n^\beta \left(\frac{1}{n} S_n \in \cdot \mid S_n \geq 0 \right)$$

satisface un Principio de Grandes Desviaciones en $[0, 1]$ con tasa n y cuya función tasa I_β tiene a $\theta^*(\beta)$ como único cero (Figura 4.4).

Este resultado explica lo que está detrás del teorema 1, en esencia lo que establece este resultado es que

$$\lim_{\delta \rightarrow 0^+} \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^{+, \beta}([\theta - \delta, \theta + \delta]) = -I_\beta(\theta).$$

Antes de continuar recordemos algunos elementos, la reescritura de \mathbb{P}_n^β , donde

$$\widehat{I}_n(w) = \sum_{i,j=0}^n 1_{\{w_i=w_j\}},$$

el número de visitas o el tiempo local en x al tiempo n

$$l_n(x) = \sum_{i=0}^n 1_{\{w_i=x\}}, \quad x \in \mathbb{Z}, \quad n \in \mathbb{N}_0,$$

y que $\sum_{x \in \mathbb{Z}} l_n(x) = n + 1$ para todo $n \in \mathbb{N}$ y $\widehat{I}_n(w) = \sum_{x \in \mathbb{Z}} l_n^2(x)$.

De este modo vemos que la prueba del teorema anterior realmente consiste en analizar las propiedades desde el punto de vista de la teoría de grandes desviaciones de la sucesión aleatoria $\{l_n(x)\}_{x \in \mathbb{Z}}$ bajo la distribución \mathbb{P}_n de una caminata aleatoria simple.

Veremos esta sucesión tiene una estructura similar a una cadena de Markov.

Para lo siguiente, recordemos que \mathbb{P} y \mathbb{E} denotan la probabilidad y la esperanza respecto a una caminata aleatoria simple y que \mathbb{P}_n es la restricción de \mathbb{P} sobre \mathcal{W}_n .

4.2. Grandes desviaciones para puentes

Con el objetivo de obtener el Principio de Grandes Desviaciones para $\mathbb{P}_n^{+, \beta}$, empezaremos deduciendo un principio de grandes desviaciones bajo la restricción de que la trayectoria es un *puente*, es decir; se encuentra entre sus puntos extremos, esta condición es crucial para la prueba de nuestro resultado principal y posteriormente será debilitada.

TEOREMA 4. Para toda $\beta \in (0, \infty)$ la familia $(\mathbb{P}_n^{\beta, \text{puente}})_{n \in \mathbb{N}_0}$, definida por

$$\mathbb{P}_n^{\beta, \text{puente}}(\cdot) = \mathbb{P}_n^\beta \left(\frac{1}{n} S_n \in \cdot \mid A_n \right)$$

satisface un Principio de Grandes Desviaciones en $(0, 1]$ con tasa n y función tasa J_β y cuyo único cero es $\theta^*(\beta)$ (Figura 4.2).

Para la prueba de este teorema seguiremos el siguiente programa:

(i) Elegimos $\theta \in (0, 1]$ y consideramos la siguiente cantidad

$$\mathbb{P}_n^\beta(S_n = \lceil \theta n \rceil \mid A_n) = \frac{\widehat{K}_n(\theta)}{\int_{\theta \in (0,1]} \widehat{K}_n(\theta) d(\theta n)}$$

donde

$$\widehat{K}_n(\theta) = \mathbb{E}_n \left[e^{-\beta \widehat{I}_n} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} \right],$$

$\lceil \theta n \rceil$ y n deben tener la misma paridad. El caso $\theta = 0$ no es relevante para los puentes.

Notemos que

$$\mathbb{P}_n^\beta(S_n = \lceil \theta n \rceil \mid A_n) = \frac{\mathbb{P}_n^\beta(\{S_n = \lceil \theta n \rceil\} \cap A_n)}{\mathbb{P}_n^\beta(A_n)}$$

además

$$\begin{aligned} \mathbb{P}_n^\beta(\{S_n = \lceil \theta n \rceil\} \cap A_n) &= \sum_{w \in \mathcal{W}_n} \frac{1}{Z_n^\beta} e^{-\beta \widehat{I}_n(w)} 1_{\{S_n = \lceil \theta n \rceil\}}^{(w)} 1_{A_n}^{(w)} \mathbb{P}_n(w) \\ &= \frac{1}{Z_n^\beta} \mathbb{E}_n \left[e^{-\beta \widehat{I}_n} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} \right] = \frac{1}{Z_n^\beta} \widehat{K}_n(\theta) \end{aligned}$$

y

$$\begin{aligned} \int_{\theta \in (0,1]} \widehat{K}_n(\theta) d(\theta n) &= \int_{\theta \in (0,1]} \mathbb{E}_n \left[e^{-\beta \widehat{I}_n} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} \right] d(\theta n) \\ &= \int_{\theta \in (0,1]} \sum_{w \in \mathcal{W}_n} e^{-\beta \widehat{I}_n(w)} 1_{\{S_n = \lceil \theta n \rceil\}}^{(w)} 1_{A_n}^{(w)} \mathbb{P}_n(w) d(\theta n) \\ &= \sum_{w \in \mathcal{W}_n} e^{-\beta \widehat{I}_n(w)} 1_{A_n}^{(w)} \mathbb{P}_n(w) \int_{\theta \in (0,1]} 1_{\{S_n = \lceil \theta n \rceil\}}^{(w)} d(\theta n) \\ &= Z_n^\beta \sum_{w \in \mathcal{W}_n : w \in A_n} \frac{1}{Z_n^\beta} e^{-\beta \widehat{I}_n(w)} \mathbb{P}_n(w) = Z_n^\beta \mathbb{P}_n^\beta(A_n). \end{aligned}$$

Por tanto

$$\mathbb{P}_n^\beta(S_n = \lceil \theta n \rceil \mid A_n) = \frac{\widehat{K}_n(\theta)}{\int_{\theta \in (0,1]} \widehat{K}_n(\theta) d(\theta n)} = \frac{\widehat{K}_n(\theta)}{\mathbb{P}_n^\beta(A_n)}.$$

(ii) Veremos que existe una función $\widehat{J}_\beta : (0, 1] \rightarrow (0, \infty)$ tal que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \widehat{K}_n(\theta) = -\widehat{J}_\beta(\theta)$$

con la propiedad de que $\theta \rightarrow \widehat{J}_\beta(\theta)$ es continua, estrictamente convexa y mínima en $\theta^*(\beta)$. Identificaremos a \widehat{J}_β en términos de un problema variacional.

(iii) Combinando (i) y (ii) obtenemos que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^\beta(S_n = \lceil \theta n \rceil | A_n) = -J_\beta(\theta)$$

donde

$$J_\beta(\theta) = \widehat{J}_\beta(\theta) - \inf_{\theta \in (0, 1]} \widehat{J}_\beta(\theta).$$

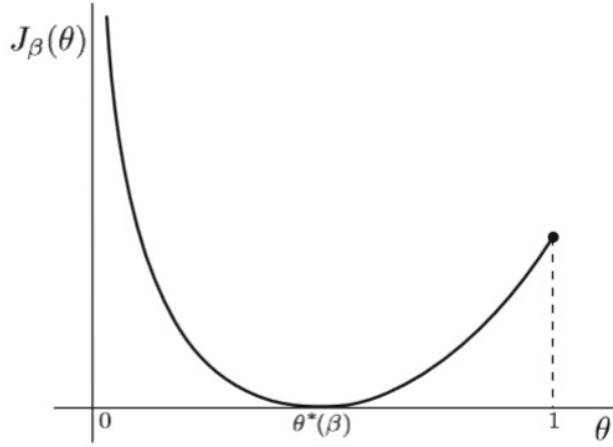


Figura 4.2: Función tasa J_β para los polímeros puente [3].

ya que

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^\beta(S_n = \lceil \theta n \rceil | A_n) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{\widehat{K}_n(\theta)}{\mathbb{P}_n^\beta(A_n)} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \widehat{K}_n(\theta) - \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^\beta(A_n) = -\widehat{J}_\beta(\theta) + \inf_{\theta \in (0, 1]} \widehat{J}_\beta(\theta) = -J_\beta(\theta). \end{aligned}$$

Notemos que

$$\begin{aligned}
\inf_{\theta \in (0,1]} \widehat{J}_\beta(\theta) &= - \sup_{\theta \in (0,1]} \{-\widehat{J}_\beta(\theta)\} = - \sup_{\theta \in (0,1]} \lim_{n \rightarrow \infty} \frac{1}{n} \log \widehat{K}_n(\theta) \\
&= - \sup_{\theta \in (0,1]} \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}_n \left[e^{-\beta \widehat{I}_n} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} \right] \\
&= - \sup_{\theta \in (0,1]} \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_{w \in \mathcal{W}_n} e^{-\beta \widehat{I}_n(w)} 1_{\{S_n = \lceil \theta n \rceil\}}^{(w)} 1_{A_n}^{(w)} \mathbb{P}_n(w) \\
&= - \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_{w \in \mathcal{W}_n} e^{-\beta \widehat{I}_n(w)} 1_{A_n}^{(w)} \mathbb{P}_n(w) \sup_{\theta \in (0,1]} 1_{\{S_n = \lceil \theta n \rceil\}}^{(w)} \\
&= - \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}_n \left[e^{-\beta \widehat{I}_n} 1_{A_n} \right] = - \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^\beta(A_n).
\end{aligned}$$

Evidentemente, $\theta \rightarrow J_\beta(\theta)$ también es continua, estrictamente convexa y mínima en $\theta^*(\beta)$, el cual es su único cero.

Estos resultados se siguen cumpliendo cuando θ se reemplaza por $\theta_n \rightarrow \theta$ cuando $n \rightarrow \infty$, que es por lo que tenemos el teorema 4.

Notemos que si queremos que la *velocidad* de polímero $\frac{1}{n} S_n$ esté cerca de un valor $\theta \in (0, 1]$, ($\frac{1}{n} S_n \in [\theta - \epsilon, \theta + \epsilon]$ con $\epsilon \rightarrow 0$) esto es equivalente a pedir que la caminata esté en el estado $\lceil \theta n \rceil$ al tiempo n ($S_n = \lceil \theta n \rceil$), es por eso que si obtenemos la forma y las propiedades de la función J_β obtenemos el principio de grandes desviaciones para los *polímeros suaves puente*.

El programa anterior se llevará a cabo en la siguiente sección por medio de cinco pasos. Los dos primeros son una preparación necesaria para obtener las cantidades necesarias en la forma adecuada para aplicar teoría de grandes desviaciones.

1. Anadiendo deriva,
2. Tiempos locales y propiedad de Markov.

La aplicación de grandes desviaciones y el análisis del problema variacional, se llevará a cabo en los últimos tres pasos.

3. El problema variacional,
4. La función tasa para puentes. Solución al problema variacional.
5. Identificación de la velocidad.

La prueba es técnica pero poderosa y puede ser utilizada para otros modelos de dimensión 1 y en modelos de dimensiones más altas.

Una vez completada la prueba de este teorema, se debilita la condición de puente y el principio de grandes desviaciones que queremos se sigue inmediatamente. Resulta que la función tasa asociada es diferente de J_β pero sigue teniendo a $\theta^*(\beta)$ como único cero.

4.3. El programa de cinco pasos

4.3.1. Paso 1: Añadiendo deriva

El objetivo es probar que si condicionamos a que la trayectoria esté en el estado $[\theta n]$ al tiempo n , entonces la probabilidad condicional puede ser aproximada, en una escala exponencial; por medio de la distribución de probabilidad de una caminata aleatoria con deriva θ .

Empezaremos por obtener diferentes equivalencias para $\widehat{K}_n(\theta)$.

Sea $\theta \in (0, 1)$ fijo y \mathbb{P}_θ y \mathbb{E}_θ la probabilidad y la esperanza para una *caminata aleatoria con deriva* θ , es decir;

$$\mathbb{P}_\theta(x, x+1) = \frac{1}{2}(1+\theta), \quad \text{la probabilidad de dar un paso a la derecha,}$$

$$\mathbb{P}_\theta(x, x-1) = \frac{1}{2}(1-\theta), \quad \text{la probabilidad de dar un paso a la izquierda;}$$

entonces, podemos reescribir

$$\widehat{K}_n(\theta) = (1-\theta)^{-\frac{n-[\theta n]}{2}} (1+\theta)^{-\frac{n+[\theta n]}{2}} \widetilde{K}_n(\theta),$$

donde

$$\widetilde{K}_n(\theta) = \mathbb{E}_\theta \left[e^{-\beta \widehat{I}_n} 1_{\{S_n = [\theta n]\}} 1_{A_n} \right]$$

Notemos que todas las trayectorias que van de 0 a $\lceil \theta n \rceil$ dan el mismo número de pasos a la derecha y el mismo número de pasos a la izquierda.

Es por eso que debemos analizar el comportamiento asintótico de $\tilde{K}_n(\theta)$, es decir; nuestra tarea es relacionar el polímero suave con deriva θ con la caminata aleatoria con deriva θ .

La ventaja de la reformulación anterior es que la trayectoria no se preocupa por regresar a $[0, S_n]$ después del tiempo n .

LEMA 5. Para toda $\theta \in (0, 1)$ y $n \in \mathbb{N}_0$,

$$\mathbb{E}_\theta \left[e^{-\beta \hat{I}_n} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} \right] = \frac{1}{\theta} \mathbb{E}_\theta \left[e^{-\beta \hat{I}_n} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} 1_{B_n} \right]$$

donde

$$B_n = \{S_i > S_n, \forall i > n\}.$$

DEMOSTRACIÓN. Nótese que \hat{I}_n no depende de S_i para cualquier $i > n$, por lo que

$$\begin{aligned} \mathbb{E}_\theta \left[e^{-\beta \hat{I}_n} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} 1_{B_n} \right] &= \mathbb{E}_\theta \left[\mathbb{E}_\theta \left[e^{-\beta \hat{I}_n} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} 1_{B_n} \mid 1_{B_n} \right] \right] \\ &= \mathbb{E}_\theta \left[1_{B_n} \mathbb{E}_\theta \left[e^{-\beta \hat{I}_n} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} \mid 1_{B_n} \right] \right] = \mathbb{E}_\theta \left[1_{B_n} \mathbb{E}_\theta \left[e^{-\beta \hat{I}_n} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} \right] \right] \\ &= \mathbb{E}_\theta \left[e^{-\beta \hat{I}_n} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} \right] \mathbb{E}_\theta [1_{B_n}] = \mathbb{E}_\theta \left[e^{-\beta \hat{I}_n} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} \right] \mathbb{P}_\theta(B_n) \end{aligned}$$

ahora

$$\mathbb{P}_\theta(B_n) = \mathbb{P}_\theta(\{S_i > S_n, \forall i > n\}) = 1 - \mathbb{P}_\theta(\{S_i = S_n, \text{p.a. } i > n\})$$

como $(S_n)_{n \in \mathbb{N}_0}$ es una cadena homogénea y $S_0 = 0$

$$\mathbb{P}_\theta(\{S_i = S_n, \text{p.a. } i > n\}) = \mathbb{P}_\theta(\{S_i = 0, \text{p.a. } i > 0\}) = \mathbb{P}_\theta(\{\text{de regresar a } 0\}).$$

Sabemos que la probabilidad de regresar a 0, en una caminata aleatoria simple es

$$1 - |p - q|$$

donde p es la probabilidad de dar un paso a la derecha y q la de dar un paso a la izquierda, entonces

$$\mathbb{P}_\theta(\{\text{de regresar a } 0\}) = 1 - \left| \frac{1+\theta}{2} - \frac{1-\theta}{2} \right| = 1 - \theta$$

por tanto

$$\begin{aligned} \mathbb{P}_\theta(B_n) &= 1 - \mathbb{P}_\theta(\{S_i = 0, \text{ p.a. } i\}) = 1 - (1 - \theta) = \theta \\ \therefore \mathbb{E}_\theta \left[e^{-\beta \widehat{I}_n} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} \right] &= \frac{1}{\theta} \mathbb{E}_\theta \left[e^{-\beta \widehat{I}_n} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} 1_{B_n} \right] \end{aligned}$$

■

Una consecuencia importante de este lema es que en el evento

$$\{S_n = \lceil \theta n \rceil\} \cap A_n \cap B_n$$

podemos escribir

$$\widehat{I}_n = \sum_{x=0}^{\lceil \theta n \rceil} l^2(x)$$

donde

$$l(x) = \sum_{i \in \mathbb{N}_0} 1_{\{S_i = x\}}, \quad x \in \mathbb{Z}$$

es el *tiempo local total en el sitio* x (número de visitas al estado x), de hecho, esto se sigue de que en el evento $\{S_n = \lceil \theta n \rceil\} \cap A_n \cap B_n$ tenemos

$$l_n(x) = \begin{cases} l(x) & \text{si } 0 \leq x \leq \lceil \theta n \rceil \\ 0 & \text{en otro caso} \end{cases}$$

por lo que podemos reescribir

$$\widetilde{K}_n(\theta) = \frac{1}{\theta} \mathbb{E}_\theta \left[e^{-\beta \sum_{x=0}^{\lceil \theta n \rceil} l^2(x)} 1_{\{S_n = \lceil \theta n \rceil\}} 1_{A_n} 1_{B_n} \right]$$

Nótese que el tiempo n ha sido reemplazado por el estado $\lceil \theta n \rceil$.

El tiempo local total tiene una agradable estructura como veremos a continuación.

4.3.2. Paso 2: Tiempos locales y propiedad de Markov

En esta sección mostraremos que $\{l(x)\}_{x \in \mathbb{N}_0}$ tiene una estructura de Markov. Esto es lo que nos permitirá analizar el comportamiento asintótico de $\widehat{K}_n(\theta)$ por medio de un principio de grandes desviaciones para estas cadenas.

Sea

$$m(x) = \sum_{i \in \mathbb{N}_0} 1_{\{S_i=x, S_{i+1}=x+1\}}, \quad x \in \mathbb{Z}$$

el número total de saltos de x a $x+1$.

Entonces, en el evento $\{S_n = \lceil \theta n \rceil\} \cap A_n \cap B_n$ el número total de saltos de $x-1$ a x es igual a

$$\sum_{i \in \mathbb{N}_0} 1_{\{S_i=x, S_{i+1}=x-1\}} = m(x) - 1, \quad 0 \leq x \leq \lceil \theta n \rceil$$

porque el número de saltos a lo largo del enlace entre $x-1$ y x debe ser al menos 1.

Como $l(x)$ es la suma del número de saltos a x que vienen de la izquierda y de la derecha, tenemos que

$$l(x) = m(x-1) + m(x) - 1_{\{x>0\}}, \quad 0 \leq x \leq \lceil \theta n \rceil.$$

Más aún, en el evento $\{S_n = \lceil \theta n \rceil\} \cap A_n \cap B_n$, el tiempo que la trayectoria está entre 0 y $\lceil \theta n \rceil$ es $n+1$, por lo que

$$\{S_n = \lceil \theta n \rceil\} \cap A_n \cap B_n = \left\{ \sum_{x=0}^{\lceil \theta n \rceil} [m(x-1) + m(x) - 1_{\{x>0\}}] = n+1, m(-1) = 0, m(\lceil \theta n \rceil) = 1 \right\}.$$

$m(-1) = 0$ porque la trayectoria es un puente que inicia en cero y $m(\lceil \theta n \rceil) = 1$ porque no debe regresar a $\lceil \theta n \rceil$ después del tiempo n .

Por tanto, podemos volver a reescribir

$$\tilde{K}_n(\theta) = \frac{1}{\theta} \mathbb{E}_\theta \left[e^{-\beta \sum_{x=0}^{\lceil \theta n \rceil} [m(x-1) + m(x) - 1_{\{x>0\}}]^2} 1_{\left\{ \sum_{x=0}^{\lceil \theta n \rceil} [m(x-1) + m(x) - 1_{\{x>0\}}] = n+1 \right\}} \right]$$

$$\times 1_{\{m(-1)=0\}} 1_{\{m(\lceil \theta n \rceil)=1\}}$$

La función indicadora $1_{\{x>0\}}$ y las restricciones $m(-1) = 0$ y $m(\lceil \theta n \rceil)$ deben ser pensadas como inofensivas condiciones de frontera.

La razón principal para esta reformulación es el siguiente resultado.

LEMA 6. *Para todo $\theta \in (0, 1)$, bajo la distribución \mathbb{P}_θ , $\{m(x)\}_{x \in \mathbb{N}_0}$ es una cadena de Markov con espacio de estados \mathbb{N} y probabilidades de transición*

$$\begin{aligned} \mathbb{P}_\theta(i, j) &= \mathbb{P}_\theta(m(x+1) = j \mid m(x) = i) \\ &= \binom{i+j-2}{i-1} \left(\frac{1+\theta}{2}\right)^i \left(\frac{1-\theta}{2}\right)^{j-1}, \quad i, j \in \mathbb{N}. \end{aligned}$$

DEMOSTRACIÓN. Sea x fija. Si $m(x) = i$, entonces el enlace $(x, x+1)$ recibe i cruces ascendentes e $i-1$ cruces descendentes.

Para $s = 1, \dots, i-1$, definimos Z_s como el número de cruces ascendentes del enlace $(x+1, x+2)$ entre el s -ésimo cruce ascendente y el s -ésimo cruce descendente de $(x, x+1)$ y Z como el número de cruces ascendentes de $(x+1, x+2)$ después del i -ésimo cruce ascendente de $(x, x+1)$, la cual es diferente de las otras porque ya no se permiten más cruces descendentes de $(x, x+1)$. El número $m(x)$ me indica cuantas Z_s 's van a existir.

Consideremos el evento $\{Z_s = k \mid m(x) = i\}$ el cual se puede ver como el evento:

E : la s -ésima vez que la caminata está en $x+1$ viniendo de x , esta da k saltos a la derecha de $x+1$ antes de dar uno a la izquierda.

Sea Y el número de saltos a la derecha de $x+1$ que la caminata da antes del primero a la izquierda, entonces $Y \sim \text{Geo}\left(\frac{1-\theta}{2}\right)$, por lo que

$$\mathbb{P}_\theta(E) = \mathbb{P}_\theta(Y = k) = \left(\frac{1-\theta}{2}\right) \left(\frac{1+\theta}{2}\right)^k$$

por lo que

$$\mathbb{P}_\theta(Z_s = k \mid m(x) = i) = \left(\frac{1+\theta}{1-\theta}\right) \left(\frac{1-\theta}{2}\right)^{k+1}, \quad k \in \mathbb{N}_0, s = 1, \dots, i-1$$

Para Z sabemos que ya no hay más cruces descendentes de $(x, x+1)$, es decir; la caminata ya no regresa a x , entonces para que esta recorra de manera ascendente el enlace $(x+1, x+2)$ debe recorrerlo de manera descendente $k-1$ veces y después dar un salto a la derecha de $x+2$, por lo que

$$\begin{aligned} \mathbb{P}_\theta(Z = k | m(x) = i) &= \left(\frac{1+\theta}{2}\right) \left(\frac{1-\theta}{2}\right)^{k-1} \\ &= \left(\frac{1+\theta}{1-\theta}\right) \left(\frac{1-\theta}{2}\right)^k, \text{ con } k \in \mathbb{N} \end{aligned}$$

Si $m(x+1) = j$ eso significa que $Z_1 + \dots + Z_{i-1} + Z = j$, así podemos ver que nuestro proceso es un proceso de Markov: para la cantidad $m(x+1)$ es irrelevante lo que la caminata hace a la izquierda de x , sólo el valor de $m(x)$ es el que le afecta. Más aún, para cada Z_s , Z_s+1 tiene la misma distribución que Z ; como $(Z_1+1) + \dots + (Z_{i-1}+1) + Z = i+j-1$ y el número de maneras de dividir $i+j-1$ en pedazos de longitud mayor o igual a 1 es $\binom{i+j-1}{i-1}$; obtenemos que

$$\mathbb{P}_\theta(i, j) = \binom{i+j-2}{i-1} \left(\frac{1+\theta}{2}\right)^i \left(\frac{1-\theta}{2}\right)^{j-1}, \quad i, j \in \mathbb{N}.$$

■

4.3.3. Paso 3: El problema variacional

Hemos completado la reescritura de $\tilde{K}_n(\theta)$ y estamos listos para aplicar la teoría de grandes desviaciones. En esta sección construimos el problema variacional implícito en el principio de grandes desviaciones para puentes.

La ventaja de la última reescritura de $\hat{K}_n(\theta)$ es que puede ser expresada en términos de la *medida empírica por pares* asociada a $\{m(x)\}_{x \in \mathbb{N}_0}$, esto es porque esa cantidad es la que mejor se adapta a la *dependencia por pares* de una cadena de Markov. Para esto definimos

$$L_N^2 = \frac{1}{N} \sum_{x=0}^{N-1} \delta_{(m(x-1), m(x))}, \quad N \in \mathbb{N}$$

con *condiciones de frontera* $(m(-1) = m(N))$.

Sea $\widetilde{\mathcal{M}}_1(\mathbb{N} \times \mathbb{N})$ el conjunto de todas las medidas de probabilidad en $\mathbb{N} \times \mathbb{N}$ con funciones marginales idénticas y

$$A_\theta = \left\{ \nu \in \widetilde{\mathcal{M}}_1(\mathbb{N} \times \mathbb{N}) \left| \sum_{i,j \in \mathbb{N}} (i+j-1)\nu(i,j) = \frac{1}{\theta} \right. \right\}.$$

Definimos

$$F_\beta(\nu) = -\beta \sum_{i,j \in \mathbb{N}} (i+j-1)^2 \nu(i,j), \quad \text{con } \nu \in \widetilde{\mathcal{M}}_1(\mathbb{N} \times \mathbb{N})$$

entonces tenemos que

$$\widetilde{K}_n(\theta) = e^{o(n)} \mathbb{E}_\theta \left[e^{NF_\beta(L_N^2)} 1_{\{L_N^2 \in A_\theta\}} \right] \quad \text{con } N = \lceil \theta n \rceil + 1.$$

Notemos que

1. El factor exponencial en la última expresión para $\widetilde{K}_n(\theta)$ es el mismo en la expresión anterior con un error despreciable que surge al forzar las condiciones de frontera.

Si $m(x-1) = i$ y $m(x) = j$ tenemos que $\delta_{(m(x-1), m(x))}(i, j) = 1$ y

$$[m(x-1) + m(x) - 1]^2 = (i+j-1)^2 = \sum_{i,j \in \mathbb{N}} (i+j-1)^2 \delta_{(m(x-1), m(x))}(i, j),$$

si además $N = \lceil \theta n \rceil + 1$

$$\begin{aligned} -\beta \sum_{x=0}^{\lceil \theta n \rceil} [m(x-1) + m(x) - 1]^2 &= -\beta \sum_{x=0}^{N-1} \sum_{i,j \in \mathbb{N}} (i+j-1)^2 \delta_{(m(x-1), m(x))}(i, j) \\ &= -\beta N \sum_{i,j \in \mathbb{N}} (i+j-1)^2 \frac{1}{N} \sum_{x=0}^{N-1} \delta_{(m(x-1), m(x))}(i, j) \\ &= -\beta N \sum_{i,j \in \mathbb{N}} (i+j-1)^2 L_N^2(i, j) = NF_\beta(L_N^2) \end{aligned}$$

2. La primer restricción en $\tilde{K}_n(\theta)$, es asintóticamente la misma en la expresión anterior.

Si

$$\sum_{x=0}^{\lceil \theta n \rceil} [m(x-1) + m(x) - 1] = n + 1$$

de la misma forma que en la observación anterior

$$\sum_{i,j \in \mathbb{N}} (i+j-1)L_N^2(i,j) = \frac{n+1}{N} = \frac{n+1}{\lceil \theta n \rceil + 1}$$

y además tenemos que

$$\theta n \leq \lceil \theta n \rceil \leq \theta n + 1 \quad \text{y} \quad \frac{n+1}{\theta n + 2} \leq \frac{n+1}{\lceil \theta n \rceil + 1} \leq \frac{n+1}{\theta n + 1}$$

por lo que

$$\lim_{n \rightarrow \infty} \frac{n+1}{\lceil \theta n \rceil + 1} = \frac{1}{\theta}$$

y

$$\sum_{i,j \in \mathbb{N}} (i+j-1)L_N^2(i,j) = \frac{1}{\theta} \quad \text{si } n \rightarrow \infty.$$

3. La segunda restricción es despreciable cuando $n \rightarrow \infty$.

La razón de esta nueva representación de $\tilde{K}_n(\theta)$ es que nos permite utilizar el principio de grandes desviaciones para la medida empírica por pares L_N^2 , basada en la propiedad de Markov de la cadena $\{m(x)\}_{x \in \mathbb{N}_0}$ establecida previamente.

LEMA 7. Para todo $\theta \in (0, 1)$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \tilde{K}_n(\theta) = -\tilde{J}_\beta(\theta)$$

con

$$\tilde{J}_\beta(\theta) = \theta \inf_{\nu \in A_\theta} [-F_\beta(\nu) + I_{\mathbb{P}_\theta}^2(\nu)]$$

donde

$$I_{\mathbb{P}_\theta}^2(\nu) = \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \left(\frac{\nu(i,j)}{\bar{\nu}(i)\mathbb{P}_\theta(i,j)} \right) \quad \text{y} \quad \bar{\nu}(i) = \sum_{j \in \mathbb{N}} \nu(i,j)$$

DEMOSTRACIÓN. La fórmula anterior es la función tasa del principio de grandes desviaciones para $(L_N^2)_{N \in \mathbb{N}}$ bajo la distribución de la cadena de Markov $\{m(x)\}_{x \in \mathbb{N}_0}$ con probabilidades de transición \mathbb{P}_θ la cual es continua y estrictamente convexa en $\widetilde{\mathcal{M}}_1(\mathbb{N} \times \mathbb{N})$ ([1] sección IV, teorema IV.3, teorema IV.5).

Con el objetivo de aplicar el lema de Varadhan, necesitamos superar la dificultad técnica de que el espacio de estados \mathbb{N} es infinito. Esto se puede hacer por un argumento de *truncación* ([4] sección 2, problema 2) porque como se vió previamente en el lema 6; el proceso $\{m(x)\}_{x \in \mathbb{N}_0}$ es en efecto una cadena de Markov.

Ahora, aplicamos el lema de Varadhan a la última representación de $\widetilde{K}_n(\theta)$, la cual es una integral exponencial restringida a A_θ ; aquí surge otra dificultad, el principio de grandes desviaciones debe ser 'transferido' de $\widetilde{\mathcal{M}}_1(\mathbb{N} \times \mathbb{N})$ a A_θ ([2] capítulo 4, lema 4.1.5).

Ahora, sea

$$\widetilde{R}_N(\theta) = \mathbb{E}_\theta[e^{NF_\beta(L_N^2)} 1_{\{L_N^2 \in A_\theta\}}]$$

entonces

$$\frac{1}{N} \log \widetilde{R}_N(\theta) = \frac{1}{N} \log \mathbb{E}_\theta[e^{NF_\beta(L_N^2)} 1_{\{L_N^2 \in A_\theta\}}]$$

por el lema de Varadhan

$$\frac{1}{N} \log \widetilde{R}_N(\theta) = \sup_{\nu \in A_\theta} [F_\beta(\nu) - I_{\mathbb{P}_\theta^2}(\nu)], \quad \text{si } N \rightarrow \infty$$

es decir;

$$\mathbb{E}_\theta[e^{NF_\beta(L_N^2)} 1_{\{L_N^2 \in A_\theta\}}] = e^{N \sup_{\nu \in A_\theta} [F_\beta(\nu) - I_{\mathbb{P}_\theta^2}(\nu)]}, \quad \text{si } N \rightarrow \infty$$

por lo que

$$\begin{aligned} \frac{1}{n} \log \widetilde{K}_n(\theta) &= \frac{1}{n} \log \mathbb{E}_\theta \left[e^{NF_\beta(L_N^2)} 1_{\{L_N^2 \in A_\theta\}} \right] \\ &= \frac{1}{n} \log \left\{ e^{N \sup_{\nu \in A_\theta} [F_\beta(\nu) - I_{\mathbb{P}_\theta^2}(\nu)]} \right\} \\ &= \frac{N}{n} \sup_{\nu \in A_\theta} [F_\beta(\nu) - I_{\mathbb{P}_\theta^2}(\nu)], \quad \text{si } N \rightarrow \infty. \end{aligned}$$

Si $n \rightarrow \infty$ entonces $N \rightarrow \infty$, es decir; todo lo anterior se sigue cumpliendo y

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \tilde{K}_n(\theta) = \lim_{n \rightarrow \infty} \frac{N}{n} \sup_{\nu \in A_\theta} [F_\beta(\nu) - I_{\mathbb{P}_\theta}^2(\nu)]$$

Como $N = [\theta n] + 1$ y $\theta n < [\theta n] < \theta n + 1$,

$$\theta < \frac{[\theta n]}{n} < \theta + \frac{1}{n} \quad \text{y} \quad \lim_{n \rightarrow \infty} \frac{[\theta n] + 1}{n} = \lim_{n \rightarrow \infty} \frac{[\theta n]}{n} = \theta.$$

Entonces

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \tilde{K}_n(\theta) = \theta \sup_{\nu \in A_\theta} [F_\beta(\nu) - I_{\mathbb{P}_\theta}^2(\nu)] = -\theta \inf_{\nu \in A_\theta} [-F_\beta(\nu) + I_{\mathbb{P}_\theta}^2(\nu)] = -\tilde{J}_\beta(\theta)$$

y como $I_{\mathbb{P}_\theta}^2$ es continua en $\tilde{\mathcal{M}}(\mathbb{N} \times \mathbb{N})$ y $\nu \rightarrow F_\beta(\nu)$ es lineal, $\tilde{J}_\beta(\theta)$ es continua. ■

Recordemos que

$$\hat{K}_n(\theta) = (1 - \theta)^{-\frac{n - [\theta n]}{2}} (1 + \theta)^{-\frac{n + [\theta n]}{n}} \tilde{K}_n(\theta)$$

donde

$$\tilde{K}_n(\theta) = \mathbb{E}_\theta \left[e^{-\beta \hat{I}_n} 1_{\{S_n = [\theta n]\}} 1_{A_n} \right]$$

y reescribimos el lema anterior de la siguiente manera

LEMA 8. *Para todo $\theta \in (0, 1)$,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \hat{K}_n(\theta) = -\hat{J}_\beta(\theta),$$

con

$$\hat{J}_\beta(\theta) = \theta \inf_{\nu \in A_\theta} [-F_\beta(\nu) + I_{\mathbb{P}_0}^2(\nu)]$$

es decir; la misma fórmula variacional que la del lema anterior pero con la distribución \mathbb{P}_0 en vez de \mathbb{P}_θ

$$\mathbb{P}_0(i, j) = \binom{i + j - 2}{i - 1} \left(\frac{1}{2}\right)^{i+j-1}, \quad i, j \in \mathbb{N}.$$

DEMOSTRACIÓN. Primero notemos que si $\nu \in A_\theta$, entonces

$$\sum_{i,j \in \mathbb{N}} (i+j-1)\nu(i,j) = \sum_{i,j \in \mathbb{N}} i\nu(i,j) + j\nu(i,j) - \nu(i,j) = \frac{1}{\theta}$$

$$\sum_{i \in \mathbb{N}} i \sum_{j \in \mathbb{N}} \nu(i,j) + \sum_{j \in \mathbb{N}} j \sum_{i \in \mathbb{N}} \nu(i,j) - \sum_{i,j \in \mathbb{N}} \nu(i,j) = \frac{1}{\theta}$$

$$\sum_{i \in \mathbb{N}} i\bar{\nu}(i) + \sum_{j \in \mathbb{N}} j\bar{\nu}(j) - 1 = \frac{1}{\theta}$$

$$2 \sum_{i \in \mathbb{N}} i\bar{\nu}(i) = \frac{1}{\theta} + 1 \quad \text{por lo que} \quad \sum_{i \in \mathbb{N}} i\bar{\nu}(i) = \frac{1+\theta}{2\theta}$$

por otro lado

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n} \log \left[(1-\theta)^{-\frac{n-[\theta n]}{2}} (1+\theta)^{-\frac{n+[\theta n]}{2}} \right] \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \left(-\frac{n-[\theta n]}{2} \right) \log(1-\theta) + \frac{1}{n} \left(-\frac{n+[\theta n]}{2} \right) \log(1+\theta) \\ &= \frac{1}{2} \log(1-\theta) \lim_{n \rightarrow \infty} \left(-\frac{n-[\theta n]}{n} \right) - \frac{1}{2} \log(1+\theta) \lim_{n \rightarrow \infty} \left(\frac{n+[\theta n]}{n} \right) \\ &= -\frac{1}{2} \log(1-\theta) \lim_{n \rightarrow \infty} \left(1 - \frac{[\theta n]}{n} \right) - \frac{1}{2} \log(1+\theta) \lim_{n \rightarrow \infty} \left(1 + \frac{[\theta n]}{n} \right) \\ &= -\frac{1-\theta}{2} \log(1-\theta) - \frac{1+\theta}{2} \log(1+\theta) \end{aligned}$$

además

$$\begin{aligned} I_{\mathbb{P}_0}^2(\nu) - I_{\mathbb{P}_\theta}^2(\nu) &= \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \left(\frac{\nu(i,j)}{\bar{\nu}(i)\mathbb{P}_0(i,j)} \right) - \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \left(\frac{\nu(i,j)}{\bar{\nu}(i)\mathbb{P}_\theta(i,j)} \right) \\ &= \sum_{i,j \in \mathbb{N}} \nu(i,j) [\log \nu(i,j) - \log \bar{\nu}(i) - \log \mathbb{P}_0(i,j) - \log \nu(i,j) + \log \bar{\nu}(i) + \log \mathbb{P}_\theta(i,j)] \\ &= \sum_{i,j \in \mathbb{N}} \nu(i,j) [\log \mathbb{P}_\theta(i,j) - \log \mathbb{P}_0(i,j)] = \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \left[\frac{\mathbb{P}_\theta(i,j)}{\mathbb{P}_0(i,j)} \right] \end{aligned}$$

$$\begin{aligned}
&= \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \left[\frac{\binom{i+j-2}{i-1} \left(\frac{1+\theta}{2}\right)^i \left(\frac{1-\theta}{2}\right)^{j-1}}{\binom{i+j-2}{i-1} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{j-1}} \right] = \sum_{i,j \in \mathbb{N}} \nu(i,j) \log [(1+\theta)^i (1-\theta)^{j-1}] \\
&= \sum_{i,j \in \mathbb{N}} \nu(i,j) [i \log(1+\theta) + (j-1) \log(1-\theta)] \\
&= \log(1+\theta) \sum_{i,j \in \mathbb{N}} i \nu(i,j) + \log(1-\theta) \sum_{i,j \in \mathbb{N}} (j-1) \nu(i,j) \\
&= \log(1+\theta) \sum_{i \in \mathbb{N}} i \bar{\nu}(i) + \log(1-\theta) \left(\sum_{j \in \mathbb{N}} j \bar{\nu}(j) - \sum_{j \in \mathbb{N}} \bar{\nu}(j) \right) \\
&= \frac{1+\theta}{2\theta} \log(1+\theta) + \frac{1-\theta}{2\theta} \log(1-\theta).
\end{aligned}$$

Por lo establecido en el lema anterior

$$\begin{aligned}
&\lim_{n \rightarrow \infty} \frac{1}{n} \log \tilde{K}_n(\theta) = -\tilde{J}_\beta(\theta) = -\theta \inf_{\nu \in A_\theta} [-F_\beta(\nu) + I_{\mathbb{P}_\theta}^2(\nu)] \\
&= \theta \sup_{\nu \in A_\theta} [F_\beta(\nu) - I_{\mathbb{P}_\theta}^2(\nu)] = \theta \sup_{\nu \in A_\theta} [F_\beta(\nu) + I_{\mathbb{P}_0}^2(\nu) - I_{\mathbb{P}_\theta}^2(\nu) - I_{\mathbb{P}_0}^2(\nu)] \\
&= \theta \left(\sup_{\nu \in A_\theta} [F_\beta(\nu) - I_{\mathbb{P}_0}^2(\nu)] + \frac{1+\theta}{2\theta} \log(1+\theta) + \frac{1-\theta}{2\theta} \log(1-\theta) \right) \\
&= -\theta \inf_{\nu \in A_\theta} [-F_\beta(\nu) + I_{\mathbb{P}_0}^2(\nu)] + \frac{1+\theta}{2} \log(1+\theta) + \frac{1-\theta}{2} \log(1-\theta)
\end{aligned}$$

por lo que

$$\begin{aligned}
&\lim_{n \rightarrow \infty} \frac{1}{n} \log \hat{K}_n(\theta) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left[(1-\theta)^{-\frac{n-[\theta n]}{2}} (1+\theta)^{-\frac{n+[\theta n]}{n}} \tilde{K}_n(\theta) \right] \\
&= \lim_{n \rightarrow \infty} \frac{1}{n} \log \left[(1-\theta)^{-\frac{n-[\theta n]}{2}} (1+\theta)^{-\frac{n+[\theta n]}{2}} \right] + \lim_{n \rightarrow \infty} \frac{1}{n} \log \tilde{K}_n(\theta) \\
&= -\theta \inf_{\nu \in A_\theta} [-F_\beta(\nu) + I_{\mathbb{P}_0}^2(\nu)]
\end{aligned}$$

por lo tanto

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \hat{K}_n(\theta) = -\theta \inf_{\nu \in A_\theta} [-F_\beta(\nu) + I_{\mathbb{P}_0}^2(\nu)] = -\hat{J}_\beta(\theta).$$

■

Hemos identificado $\widehat{J}_\beta(\theta)$ para $\theta \in (0, 1)$, la función tasa para los polímeros puente, la cual es válida para $\theta = 1$.

Si $\theta = 1$ la caminata sólo da pasos a la derecha, es decir; no tiene auto-intersecciones, por lo que para esta trayectoria w , $\widehat{I}_n(w) = 0$ y $1_{A_n}(w) = 1$, entonces

$$\widehat{K}_n(1) = \mathbb{E}_n[e^{-\beta \widehat{I}_n} 1_{\{S_n=n\}} 1_{A_n}] = \mathbb{E}_n[1_{\{S_n=n\}}] = \mathbb{P}_n(S_n = n) = \frac{1}{2^n}$$

por tanto

$$\widehat{J}_\beta(1) = - \lim_{n \rightarrow \infty} \frac{1}{n} \widehat{K}_n(1) = - \lim_{n \rightarrow \infty} \frac{1}{n} \log 2^{-n} = \log 2.$$

4.3.4. Paso 4: La función tasa para puentes. Solución al problema variacional.

Continuamos dando la solución al problema variacional del lema 8, el cual explica el comportamiento cualitativo de la función $\theta \rightarrow J_\beta(\theta)$, (Figura 4.2). El problema requiere minimizar una funcional no-lineal bajo cierta restricción lineal. Es posible encontrar la solución en términos de un problema de eigenvalores.

Sea $\beta \in (0, \infty)$ fija, $r \in \mathbb{R}$ y $A_{r,\beta}$ la matriz de $\mathbb{N} \times \mathbb{N}$ con entradas

$$A_{r,\beta}(i, j) = e^{r(i+j-1) - \beta(i+j-1)^2} \mathbb{P}_0(i, j), \quad i, j \in \mathbb{N}.$$

LEMA 9. *Sea $\beta \in (0, \infty)$ fijo. Para todo $r \in \mathbb{R}$, $A_{r,\beta}$ es un operador auto-adjunto en $\ell^2(\mathbb{N})$ con un único y más grade eigenvalor $\lambda_{r,\beta}$ y correspondiente eigenvector $\tau_{r,\beta}$ (normalizado $\|\tau_{r,\beta}\|_2 = 1$).*

DEMOSTRACIÓN. [4], sección 1. Para cualquier $r \in \mathbb{R}$ y $i, j \in \mathbb{N}$, tenemos que $A_{r,\beta}(i, j) = A_{r,\beta}(j, i) > 0$ y $\sum_{i,j \in \mathbb{N}} A_{r,\beta}(i, j)^2 < \infty$, por lo que el operador $A_{r,\beta} : \ell^2(\mathbb{N}) \rightarrow \ell^2(\mathbb{N})$ es un operador auto-adjunto positivo y compacto y por tanto existe un único eigenvalor $\lambda_{r,\beta}$ con eigenvector asociado $\tau_{r,\beta}$.

■

El siguiente lema nos da una breve descripción de las propiedades del eigenvalor $\lambda_{r,\beta}$ asociado a $A_{r,\beta}$, ([3] capítulo 3, sección 3.4.5., [4] sección 1).

LEMA 10.

1. $(r, \beta) \rightarrow \lambda_{r,\beta}$ es analítica en $\mathbb{R} \times (0, \infty)$.
2. $\lim_{r \rightarrow -\infty} \frac{\partial}{\partial r} \log \lambda_{r,\beta} = 1$ y $\lim_{r \rightarrow \infty} \frac{\partial}{\partial r} \log \lambda_{r,\beta} = \infty$ para toda $\beta \in (0, \infty)$.
3. $r \rightarrow \log \lambda_{r,\beta}$ es estrictamente convexa para toda $\beta \in (0, \infty)$.

DEMOSTRACIÓN.

1. La analiticidad se sigue porque $\lambda_{r,\beta}$ tiene multiplicidad 1 y todos los elementos de $A_{r,\beta}$ son analíticos.
2. Esto se sigue por una estimación directa en el eigenvector $\tau_{r,\beta}$ para $r \rightarrow -\infty$ y $r \rightarrow \infty$ respectivamente.
3. La convexidad se sigue de las siguientes observaciones.
 - a) $\lambda_{r,\beta} = \sup_{x \in \ell^2(\mathbb{N}) : x > 0, \|x\|_2 = 1} \sum_{i,j \in \mathbb{N}} x(i) A_{r,\beta}(i,j) x(j)$;
 - b) $r \rightarrow \log A_{r,\beta}(i,j)$ es lineal para toda $i, j \in \mathbb{N}$ y $\beta \in (0, \infty)$;
 - c) la log-convexidad se preserva bajo sumas y al tomar supremos.

La convexidad estricta se sigue de la convexidad en combinación con 1 y 2. ■

Con ayuda del lema 10 podemos expresar a $\widehat{J}_\beta(\theta)$ en términos del eigenvalor $\lambda_{r,\beta}$, para algún valor r que depende de θ .

LEMA 11. (Variación sobre ν). Sea $\beta \in (0, \infty)$ fija, entonces; para todo $\theta \in (0, 1)$

$$\widehat{J}_\beta(\theta) = r - \theta \log \lambda_{r,\beta} \Big|_{r=r_\beta(\theta)},$$

donde $r_\beta(\theta) \in \mathbb{R}$ es la única solución a la ecuación

$$\frac{1}{\theta} = \frac{\partial}{\partial r} \log \lambda_{r,\beta}.$$

DEMOSTRACIÓN. El hecho de que la ecuación anterior tenga una solución para todo $\theta \in (0, 1)$ y que esta sea única se sigue del lema 10, ya que $r \rightarrow \frac{\partial}{\partial r} \log \lambda_{r,\beta}$ es estrictamente creciente o bien, $r \rightarrow \lambda_{r,\beta}$ es estrictamente log-convexa ([4], sección 3, proposición 4).

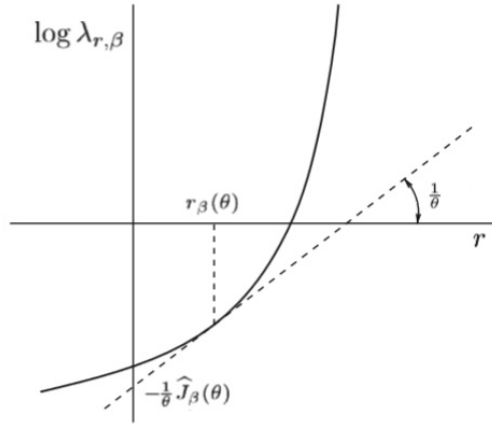


Figura 4.3: Identificación de \widehat{J}_β [3].

En la figura 4.3 podemos ver que $\theta \rightarrow r_\beta(\theta)$ es decreciente.

Consideramos la siguiente familia de medidas de probabilidad

$$\nu_{r,\beta}(i, j) = \frac{1}{\lambda_{r,\beta}} \tau_{r,\beta}(i) A_{r,\beta}(i, j) \tau_{r,\beta}(j), \quad i, j \in \mathbb{N}.$$

OBSERVACIÓN. $\nu_{r,\beta} \in \widetilde{\mathcal{M}}(\mathbb{N} \times \mathbb{N})$.

$$\begin{aligned} \bar{\nu}_{r,\beta}(i) &= \sum_{j \in \mathbb{N}} \nu_{r,\beta}(i, j) = \sum_{j \in \mathbb{N}} \frac{1}{\lambda_{r,\beta}} \tau_{r,\beta}(i) A_{r,\beta}(i, j) \tau_{r,\beta}(j) \\ &= \frac{1}{\lambda_{r,\beta}} \tau_{r,\beta}(i) \sum_{j \in \mathbb{N}} A_{r,\beta}(i, j) \tau_{r,\beta}(j) \end{aligned}$$

como $\tau_{r,\beta}$ es el eigenvector asociado al eigenvalor $\lambda_{r,\beta}$ de $A_{r,\beta}$ se tiene que

$$A_{r,\beta} \tau_{r,\beta} = \lambda_{r,\beta} \tau_{r,\beta} \quad \text{y} \quad \sum_{j \in \mathbb{N}} A_{r,\beta}(i, j) \tau_{r,\beta}(j) = \lambda_{r,\beta} \tau_{r,\beta}(i)$$

por lo que $\bar{\nu}_{r,\beta}(i) = \tau_{r,\beta}^2(i)$ y de la misma manera $\bar{\nu}_{r,\beta}(j) = \tau_{r,\beta}^2(j)$ entonces $\bar{\nu}_{r,\beta}(i) = \bar{\nu}_{r,\beta}(j)$, por tanto

$$\nu_{r,\beta} \in \widetilde{\mathcal{M}}(\mathbb{N} \times \mathbb{N}).$$

□

Luego

$$\begin{aligned} I_{\mathbb{P}_0}^2(\nu_{r,\beta}) &= \sum_{i,j \in \mathbb{N}} \nu_{r,\beta}(i,j) \log \left(\frac{\nu_{r,\beta}(i,j)}{\bar{\nu}_{r,\beta}(i)\mathbb{P}_0(i,j)} \right) \\ &= \sum_{i,j \in \mathbb{N}} \nu_{r,\beta}(i,j) [\log \nu_{r,\beta}(i,j) - \log \bar{\nu}_{r,\beta}(i) - \log \mathbb{P}_0(i,j)] \\ &= \sum_{i,j \in \mathbb{N}} \nu_{r,\beta}(i,j) \left[\log \left(\frac{1}{\lambda_{r,\beta}} \tau_{r,\beta}(i) A_{r,\beta}(i,j) \tau_{r,\beta}(j) \right) - \log \tau_{r,\beta}^2(i) - \log \mathbb{P}_0(i,j) \right] \\ &= \sum_{i,j \in \mathbb{N}} \nu_{r,\beta}(i,j) [-\log \lambda_{r,\beta} - \log \tau_{r,\beta}(i) + \log A_{r,\beta}(i,j) + \log \tau_{r,\beta}(j) - \log \mathbb{P}_0(i,j)] \\ &= \sum_{i,j \in \mathbb{N}} \nu_{r,\beta}(i,j) \left[-\log \lambda_{r,\beta} + \log \frac{\tau_{r,\beta}(j)}{\tau_{r,\beta}(i)} + \log \left(e^{r(i+j-1) - \beta(i+j-1)^2} \mathbb{P}_0(i,j) \right) - \log \mathbb{P}_0(i,j) \right] \\ &= \sum_{i,j \in \mathbb{N}} \nu_{r,\beta}(i,j) \left[r(i+j-1) - \beta(i+j-1)^2 - \log \lambda_{r,\beta} + \log \frac{\tau_{r,\beta}(j)}{\tau_{r,\beta}(i)} \right]. \end{aligned}$$

Notemos que

1. Si $r = r_\beta(\theta)$ entonces $\sum_{i,j \in \mathbb{N}} (i+j-1) \nu_{r,\beta}(i,j) = \frac{1}{\theta}$, es decir; $\nu_{r,\beta} \in A_\theta$

Notemos que

$$\begin{aligned} \frac{\partial}{\partial r} A_{r,\beta}(i,j) &= \frac{\partial}{\partial r} e^{r(i+j-1) - \beta(i+j-1)^2} \mathbb{P}_0(i,j) \\ &= (i+j-1) e^{r(i+j-1) - \beta(i+j-1)^2} \mathbb{P}_0(i,j) = (i+j-1) A_{r,\beta}(i,j), \end{aligned}$$

por este hecho y nuevamente porque $A_{r,\beta} \tau_{r,\beta} = \lambda_{r,\beta} \tau_{r,\beta}$ y $\|\tau_{r,\beta}\|_2 = 1$ se tiene que

$$\sum_{i,j \in \mathbb{N}} (i+j-1) \nu_{r,\beta}(i,j) = \frac{1}{\lambda_{r,\beta}} \sum_{i,j \in \mathbb{N}} \tau_{r,\beta}(i) (i+j-1) A_{r,\beta}(i,j) \tau_{r,\beta}(j)$$

$$\begin{aligned}
&= \frac{1}{\lambda_{r,\beta}} \sum_{i,j \in \mathbb{N}} \tau_{r,\beta}(i) \left[\frac{\partial}{\partial r} A_{r,\beta}(i,j) \right] \tau_{r,\beta}(j) = \frac{1}{\lambda_{r,\beta}} \frac{\partial}{\partial r} \sum_{i,j \in \mathbb{N}} \tau_{r,\beta}(i) A_{r,\beta}(i,j) \tau_{r,\beta}(j) \\
&= \frac{1}{\lambda_{r,\beta}} \frac{\partial}{\partial r} \lambda_{r,\beta} = \frac{\partial}{\partial r} \log \lambda_{r,\beta} = \frac{1}{\theta},
\end{aligned}$$

por lo que $\nu_{r,\beta} \in A_\theta$ y por tanto

$$r \sum_{i,j \in \mathbb{N}} (i+j-1) \nu_{r,\beta}(i,j) = \frac{r}{\theta}.$$

$$2. \quad -\beta \sum_{i,j \in \mathbb{N}} \nu_{r,\beta}(i,j) (i+j-1)^2 = F_\beta(\nu_{r,\beta}).$$

$$3. \quad \sum_{i,j \in \mathbb{N}} \nu_{r,\beta}(i,j) \log \lambda_{r,\beta} = \log \lambda_{r,\beta}.$$

4. Como $\nu_{r,\beta}$ tiene marginales idénticas

$$\begin{aligned}
&\sum_{i,j \in \mathbb{N}} \nu_{r,\beta}(i,j) \log \frac{\tau_{r,\beta}(j)}{\tau_{r,\beta}(i)} \\
&= \sum_{j \in \mathbb{N}} \log \tau_{r,\beta}(j) \sum_{i \in \mathbb{N}} \nu_{r,\beta}(i,j) - \sum_{i \in \mathbb{N}} \log \tau_{r,\beta}(i) \sum_{j \in \mathbb{N}} \nu_{r,\beta}(i,j) \\
&= \sum_{j \in \mathbb{N}} \log \tau_{r,\beta}(j) \bar{\nu}_{r,\beta}(j) - \sum_{i \in \mathbb{N}} \log \tau_{r,\beta}(i) \bar{\nu}_{r,\beta}(i) = 0.
\end{aligned}$$

De esta manera si $r = r_\beta(\theta)$

$$I_{\mathbb{P}_0}^2(\nu_{r,\beta}) = \frac{r}{\theta} + F_\beta(\nu_{r,\beta}) - \log \lambda_{r,\beta} \quad \text{y} \quad -\theta[F_\beta(\nu_{r,\beta}) - I_{\mathbb{P}_0}^2(\nu_{r,\beta})] = r - \theta \log \lambda_{r,\beta}.$$

Verifiquemos que $\nu_{r_\beta(\theta),\beta}$ es un minimizador del problema variacional

$$\widehat{J}_\beta(\theta) = \theta \inf_{\nu \in A_\theta} [-F_\beta(\nu) + I_{\mathbb{P}_0}^2(\nu)]$$

Si $\nu \in A_\theta$ podemos escribir

$$-\theta[F_\beta(\nu) - I_{\mathbb{P}_0}^2(\nu)] = -\theta \left[F_\beta(\nu) - \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \left(\frac{\nu(i,j)}{\bar{\nu}(i) \mathbb{P}_0(i,j)} \right) \right]$$

$$\begin{aligned}
&= -\theta \left[F_\beta(\nu) - \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \nu(i,j) + \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \bar{\nu}(i) + \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \mathbb{P}_0(i,j) \right] \\
&\quad \text{como } A_{r,\beta}(i,j) = e^{r(i+j-1) - \beta(i+j-1)^2} \mathbb{P}_0(i,j), \\
&\quad \mathbb{P}_0(i,j) = e^{-r(i+j-1) + \beta(i+j-1)^2} A_{r,\beta}(i,j) \quad \text{y} \\
&\quad \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \mathbb{P}_0(i,j) \\
&= -r \sum_{i,j \in \mathbb{N}} (i+j-1) \nu(i,j) + \beta \sum_{i,j \in \mathbb{N}} (i+j-1)^2 \nu(i,j) + \sum_{i,j \in \mathbb{N}} \nu(i,j) \log A_{r,\beta}(i,j) \\
&= -\frac{r}{\theta} - F_\beta(\nu) + \sum_{i,j \in \mathbb{N}} \nu(i,j) \log A_{r,\beta}(i,j), \quad \text{ya que } \nu \in A_\theta.
\end{aligned}$$

con esta última expresión y sumando y restando $\log \lambda_{r,\beta}$ de la forma adecuada obtenemos que

$$-\theta[F_\beta(\nu) - I_{\mathbb{P}_0}^2(\nu)] = [r - \theta \log \lambda_{r,\beta}] + \theta \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \left(\frac{\nu(i,j)}{\bar{\nu}(i) \frac{1}{\lambda_{r,\beta}} A_{r,\beta}(i,j)} \right)$$

como $\bar{\nu}(i) = \tau_{r,\beta}^2(i)$, entonces

$$\begin{aligned}
&= [r - \theta \log \lambda_{r,\beta}] + \theta \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \left(\frac{\nu(i,j)}{\bar{\nu}(i)} \frac{\tau_{r,\beta}(i) \tau_{r,\beta}(j)}{\frac{1}{\lambda_{r,\beta}} \tau_{r,\beta}(i) A_{r,\beta}(i,j) \tau_{r,\beta}(j)} \right) \\
&= [r - \theta \log \lambda_{r,\beta}] + \theta \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \left(\frac{\nu(i,j)}{\bar{\nu}(i)} \frac{\sqrt{\bar{\nu}_{r,\beta}(i) \bar{\nu}_{r,\beta}(j)}}{\nu_{r,\beta}(i,j)} \right)
\end{aligned}$$

y como ν tiene marginales idénticas

$$= [r - \theta \log \lambda_{r,\beta}] + \theta \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \left(\frac{\nu(i,j)}{\bar{\nu}(i)} \frac{\bar{\nu}_{r,\beta}(i)}{\nu_{r,\beta}(i,j)} \right).$$

El primer término es precisamente el valor que obtuvimos cuando $\nu = \nu_{r,\beta}$.

Sobre el segundo término,

$$\theta \sum_{i,j \in \mathbb{N}} \nu(i,j) \log \left(\frac{\nu(i,j)}{\bar{\nu}(i)} \frac{\bar{\nu}_{r,\beta}(i)}{\nu_{r,\beta}(i,j)} \right) = \theta \sum_{i \in \mathbb{N}} \bar{\nu}(i) H([\nu(i)] \mid [\nu_{r,\beta}(i)]),$$

donde $[\nu(i)], [\nu_{r,\beta}(i)] \in \mathcal{M}(\mathbb{N})$ (el conjunto de todas las medidas de probabilidad en \mathbb{N}) se definen como

$$[\nu(i)](j) = \frac{\nu(i,j)}{\nu(i)} \quad \text{y} \quad [\nu_{r,\beta}(i)](j) = \frac{\nu_{r,\beta}(i,j)}{\nu_{r,\beta}(i)}$$

mientras que si $\mu_1, \mu_2 \in \mathcal{M}(\mathbb{N})$

$$H(\mu_1 | \mu_2) = \sum_{i \in \mathbb{N}} \mu_1(i) \log \left(\frac{\mu_1(i)}{\mu_2(i)} \right)$$

denota la entropía relativa de μ_1 respecto a μ_2 ([1], capítulo 2, [10] sección 5.1).

Claramente este término es mayor o igual a cero cumpliéndose la igualdad cuando $\nu = \nu_{r,\beta}$, por lo tanto

$$\widehat{J}_\beta(\theta) = \theta \inf_{\nu \in A_\theta} [-F_\beta(\nu) + I_{\mathbb{P}_0}^2(\nu)] = -\theta[F_\beta(\nu_{r,\beta}) - I_{\mathbb{P}_0}^2(\nu_{r,\beta})] = r - \theta \log \lambda_{r,\beta} \Big|_{r=r_\beta(\theta)}.$$

■

4.3.5. Paso 5: Identificación de la velocidad

El lema anterior nos proporciona una buena representación de $\widehat{J}_\beta(\theta)$, en términos de una familia de eigenvalores $\lambda_{r,\beta}$, $r \in \mathbb{R}$. Como vimos en la Figura 4.2, $\theta^*(\beta)$ debe ser identificado como el único punto mínimo de $\theta \rightarrow \widehat{J}_\beta(\theta)$.

LEMA 12. (Variación sobre θ). Sea $\beta \in (0, \infty)$ fija, entonces

$$\frac{1}{\theta^*(\beta)} = \frac{\partial}{\partial r} \log \lambda_{r,\beta} \Big|_{r=r^*(\beta)}$$

donde $r^*(\beta)$ es la única solución a la ecuación

$$\lambda_{r,\beta} = 1.$$

DEMOSTRACIÓN. El hecho de que la ecuación anterior tenga solución para toda $\beta \in (0, \infty)$ y esta sea única se sigue del lema 10 y la Figura 4.3.

Derivando $\widehat{J}_\beta(\theta)$ respecto a θ

$$\begin{aligned} \frac{\partial}{\partial \theta} \widehat{J}_\beta(\theta) &= \frac{\partial}{\partial \theta} (r_\beta(\theta) - \theta \log \lambda_{r_\beta(\theta), \beta}) \\ &= \frac{\partial}{\partial \theta} r_\beta(\theta) - \log \lambda_{r_\beta(\theta), \beta} - \theta \left[\frac{\partial}{\partial r} \log \lambda_{r, \beta} \Big|_{r=r_\beta(\theta)} \right] \left[\frac{\partial}{\partial \theta} r_\beta(\theta) \right] \\ &= \frac{\partial}{\partial \theta} r_\beta(\theta) - \log \lambda_{r_\beta(\theta), \beta} - \theta \left[\frac{1}{\theta} \right] \frac{\partial}{\partial \theta} r_\beta(\theta) = -\log \lambda_{r_\beta(\theta), \beta}. \end{aligned}$$

Entonces

$$\frac{\partial}{\partial \theta} \widehat{J}_\beta(\theta) = -\log \lambda_{r_\beta(\theta), \beta}$$

y esta derivada es cero si y sólo si $\lambda_{r_\beta(\theta), \beta} = 1$, es decir; la solución a la ecuación $\lambda_{r, \beta} = 1$ nos da el valor mínimo de $\theta \rightarrow \widehat{J}_\beta(\theta)$. Sabemos que $r \rightarrow r_\beta(\theta)$ es decreciente y como cada elemento de $A_{r, \beta}$ es estrictamente creciente respecto a r , se sigue del lema 10.3a que $r \rightarrow \lambda_{r, \beta}$ es estrictamente creciente, por lo que la solución de $\lambda_{r, \beta} = 1$ existe y es única ([4] proposición 5).

Sea $\theta^*(\beta)$ este valor mínimo y definimos $r^*(\beta) = r_\beta(\theta^*(\beta))$; por el lema anterior

$$\frac{1}{\theta^*(\beta)} = \frac{\partial}{\partial r} \log \lambda_{r, \beta} \Big|_{r=r^*(\beta)}.$$

Notemos que

$$\begin{aligned} \frac{\partial^2}{\partial \theta^2} \widehat{J}_\beta(\theta) &= -\frac{\partial}{\partial \theta} \log \lambda_{r_\beta(\theta), \beta} = -\frac{\partial}{\partial r} \log \lambda_{r, \beta} \Big|_{r=r_\beta(\theta)} \frac{\partial}{\partial \theta} r_\beta(\theta) \\ &= -\frac{1}{\theta} \frac{\partial}{\partial \theta} r_\beta(\theta) > 0, \end{aligned}$$

y en la Figura 4.3 se puede ver que $\theta \rightarrow r_\beta(\theta)$ tiene una pendiente negativa por ser decreciente, así que $r_\beta(\theta^*(\beta))$ es el único minimizador de $\theta \rightarrow \widehat{J}_\beta(\theta)$ y esta función es por tanto estrictamente convexa. ■

Por esto $\inf_{\theta \in (0,1]} \widehat{J}_\beta(\theta) = r^*(\beta)$ y recordando que

$$J_\beta(\theta) = \widehat{J}_\beta(\theta) - \inf_{\theta \in (0,1]} \widehat{J}_\beta(\theta)$$

tenemos que si $\theta^*(\beta)$ es el único mínimo de \widehat{J}_β , también es el único mínimo y cero de J_β (Figura 4.2).

$$\begin{aligned} J_\beta(\theta^*(\beta)) &= \widehat{J}_\beta(\theta^*(\beta)) - r^*(\beta) \\ &= r^*(\beta) - \theta^*(\beta) \log \lambda_{r^*(\beta),\beta} - r^*(\beta) = r^*(\beta) - r^*(\beta) = 0. \end{aligned}$$

Aquí termina nuestro análisis de la función tasa J_β para los polímeros suaves puente y la prueba del teorema 4.

4.4. Principio de grandes desviaciones sin la condición de puente

En las secciones anteriores probamos el principio de grandes desviaciones para un polímero suave puente, en esta sección damos un breve esquema sobre como obtener el principio de grandes desviaciones sin esta condición ([1] capítulo 9), es decir; el teorema 3 que implica el teorema 1. Resulta que la función tasa asociada a este principio de grandes desviaciones tiene una parte lineal entre 0 y una *velocidad crítica* $\theta^{**}(\beta)$ que es menor a $\theta^*(\beta)$ (Figura 4.4).

TEOREMA 3. Para todo $\beta \in (0, \infty)$ la familia $(\mathbb{P}_n^{+,\beta})_{n \in \mathbb{N}}$ definida como

$$\mathbb{P}_n^{+,\beta}(\cdot) = \mathbb{P}_n^\beta \left(\frac{1}{n} S_n \in \cdot \mid S_n \geq 0 \right)$$

satisface el principio de grandes desviaciones en $[0, 1]$ con tasa n y con función tasa I_β dada por

$$I_\beta(\theta) = \begin{cases} J_\beta(\theta) & \text{si } \theta > \theta^{**}(\beta), \\ I_\beta(0) + \frac{\theta}{\theta^{**}(\beta)} [J_\beta(\theta^{**}(\beta)) - I_\beta(0)] & \text{si } \theta \leq \theta^{**}(\beta), \end{cases}$$

donde $\theta^{**}(\beta)$ es la única solución a la ecuación

$$J_\beta(\theta) - I_\beta(0) = \theta \frac{\partial}{\partial \theta} J_\beta(\theta),$$

e $I_\beta(0)$ se define en los siguientes lemas, más aún; $\theta^{**}(\beta) \in (0, \theta^*(\beta))$.

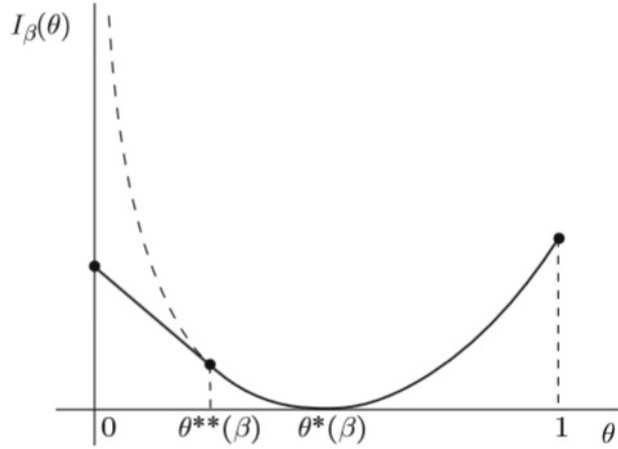


Figura 4.4: Función tasa I_β en comparación con J_β de los polímeros puente [3].

La parte lineal se puede ver por el siguiente hecho:

Si se requiere que el polímero suave se mueva a una velocidad $\theta < \theta^{**}(\beta)$, entonces *prefiere* violar la condición de puente moviéndose a velocidad $\theta^{**}(\beta)$ entre 0 y $[\theta n]$ y hacer dos *lazos*, uno a la izquierda de 0 y otro a la derecha de $[\theta n]$. La penalización por hacer estas vueltas es menor que la penalización por permanecer como un puente y la longitud de estos lazos es proporcional a la cantidad $\theta^{**}(\beta) - \theta$, es decir; la penalización por no comportarse como un puente crece linealmente con $\theta^{**}(\beta) - \theta$.

Veamos como es que se obtiene el resultado anterior.

Para $\theta \in (0, 1]$ definimos

$$-J_\beta^\circ(\theta) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^\beta \left(\max_{0 \leq i \leq n} S_i = [\theta n], S_i \geq 0 \forall 0 \leq i \leq n \mid S_n = 0 \right)$$

con el límite a lo largo de los enteros pares. Esta es la función tasa para cuando el polímero realiza un lazo que crece a velocidad θ .

Sea $j_\beta^\circ = \inf_{\theta \in (0,1]} J_\beta^\circ(\theta)$.

LEMA 13. Para todo $\theta \in (0, 1]$

$$I_\beta(\theta) = \inf_{t_1, t_2 \geq 0, t_1 + t_2 \leq 1} \left[(1 - t_1 - t_2) J_\beta \left(\frac{\theta}{1 - t_1 - t_2} \right) + (t_1 + t_2) j_\beta^\circ \right]$$

DEMOSTRACIÓN. Sean t_1 y t_2 fijos. La trayectoria puede decidir primero pasar un tiempo nt_1 a la izquierda de 0, después un tiempo $(1 - t_1 - t_2)n$ entre 0 y $[\theta n]$ y finalmente un tiempo a la derecha de $[\theta n]$. El costo por este comportamiento es

$$t_1 j_\beta^\circ \quad , \quad (1 - t_1 - t_2) J_\beta \left(\frac{\theta}{1 - t_1 - t_2} \right) \quad \text{y} \quad t_2 j_\beta^\circ$$

respectivamente, y una vez que se optimiza sobre t_1 y t_2 obtenemos el resultado. Debemos descartar la posibilidad de que la trayectoria prefiera una estrategia más complicada por las penalizaciones que podría tener, sin embargo, en tal caso podemos superar esta dificultad por medio de un argumento similar al de la demostración del último lema en el capítulo 2.

□

Para simplificar este nuevo problema variacional hacemos la siguiente observación.

LEMA 14. $I_\beta(0) = j_\beta^\circ$.

DEMOSTRACIÓN. Suponiendo que el polímero forma un lazo, podemos ver que este puede ser a la izquierda o a la derecha de 0, o bien, podría estar formado por lazos más pequeños, como esta última alternativa *en más costosa* para la trayectoria se sigue el resultado. De nuevo, esto puede verificarse por medio de un argumento similar al de la demostración del último lema en el capítulo 2.

□

Considerando los dos lemas anteriores obtenemos el siguiente problema variacional

$$I_\beta(\theta) = \inf_{0 \leq t \leq 1} \left[(1-t)J_\beta\left(\frac{\theta}{1-t}\right) + tI_\beta(0) \right],$$

el cual se puede reescribir como

$$I_\beta(\theta) - I_\beta(0) = \inf_{0 \leq s \leq 1} \left[s \left\{ J_\beta\left(\frac{\theta}{s}\right) - I_\beta(0) \right\} \right].$$

Como $\theta \rightarrow J_\beta(\theta)$ es estrictamente convexa tenemos el supremo se alcanza en

$$s = \begin{cases} 1 & \text{si } \theta \geq \theta^{**}(\beta), \\ \frac{\theta}{\theta^{**}(\beta)} & \text{si } \theta \leq \theta^{**}(\beta) \end{cases}$$

donde $\theta^{**}(\beta)$ es la única solución de la ecuación

$$J_\beta(\theta) - I_\beta(0) = \theta \frac{\partial}{\partial \theta} J_\beta(\theta).$$

Esto completa la prueba del teorema 3. Notemos que esta última expresión nos dice que la parte lineal de I_β es tangente a J_β en $\theta = \theta^{**}(\beta)$, por lo que tenemos que $\theta^{**}(\beta) \in (0, \theta^*(\beta))$ siempre que $I_\beta(0) > 0$.

Para completar el cálculo de I_β nos hace falta identificar $I_\beta(0)$ y probar que $I_\beta(0) > 0$, esto puede hacerse con ayuda de un análisis similar al de los pasos 4 y 5 de la sección anterior.

Sea

$$-\widehat{J}_\beta^\circ(\theta) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} \left[e^{-\beta \widehat{I}_n} 1_{\{\max_{0 \leq i \leq n} S_i = \lceil \theta n \rceil, S_i \geq 0, \forall 0 \leq i \leq n\}} \mid S_n = 0 \right],$$

la misma que en la definición de $J_\beta^\circ(\theta)$ pero sin la constante normalizadora Z_n^β , (recuerde la definición de \mathbb{P}_n^β).

Por el cálculo hecho en (iii), los lemas 11 y 12 y la Figura 4.2 tenemos

$$\lim_{n \rightarrow \infty} \frac{1}{n} Z_n^\beta = -r^*(\beta) = - \inf_{\theta \in (0,1]} \widehat{J}_\beta^\circ(\theta)$$

por lo que

$$J_\beta^\circ(\theta) = \widehat{J}_\beta^\circ(\theta) - r^*(\beta).$$

LEMA 15. Para todo $\theta \in (0, 1)$

$$\widehat{J}_\beta^\circ(\theta) = \theta \inf_{\nu \in A_\theta^\circ} [-F_\beta^\circ(\nu) + I_{\mathbb{P}_0^\circ}^2(\nu)],$$

donde

$$F_\beta^\circ(\nu) = -\beta \sum_{i,j \in \mathbb{N}_0} (i+j)^2 \nu(i,j),$$

$$A_\theta^\circ = \left\{ \nu \in \widetilde{\mathcal{M}}_1(\mathbb{N}_0 \times \mathbb{N}_0) \mid \sum_{i,j \in \mathbb{N}_0} (i+j) \nu(i,j) = \frac{1}{\theta}, \nu(\mathbb{N} \times \mathbb{N}) = 1 \right\},$$

y $I_{\mathbb{P}_0^\circ}^2(\nu)$ es la misma que $I_{\mathbb{P}_0}^2$ pero reemplazando a \mathbb{P}_0 por

$$\mathbb{P}_0^\circ(i,j) = 1_{\{i \neq 0\}} \mathbb{P}_0(i,j+1) + 1_{\{i=j=0\}}, \quad i,j \in \mathbb{N}_0.$$

DEMOSTRACIÓN. Estas son las mismas fórmulas que las que utilizamos para $\widehat{J}_\beta(\theta)$, excepto que (i,j) ha sido reemplazado por $(i,j-1)$, esto es porque ahora estamos considerando trayectorias que hacen lazos en vez de trayectorias que sólo se mueven a la derecha.

El resultado análogo a los lemas 11 y 12 bajo esta condición puede establecerse como sigue

LEMA 16. Sea $\beta \in (0, \infty)$ fija, para $r \in \mathbb{R}$ sea $A_{r,\beta}^\circ$ la matriz de $\mathbb{N} \times \mathbb{N}$ con entradas

$$A_{r,\beta}^\circ(i,j) = e^{r(i+j) - \beta(i+j)^2} \mathbb{P}_0^\circ(i,j), \quad i,j \in \mathbb{N},$$

con

$$\mathbb{P}_0^\circ(i,j) = 1_{\{i \neq 0\}} \mathbb{P}_0(i,j+1) + 1_{\{i=j=0\}}, \quad i,j \in \mathbb{N}_0.$$

Sea $\lambda_{r,\beta}^\circ$ el único y más grande eigenvalor de $A_{r,\beta}^\circ$ visto como un operador en $\ell^2(\mathbb{N})$, entonces

$$I_\beta(0) = r^{**}(\beta) - r^*(\beta)$$

donde $r^{**}(\beta) \in (0, \infty)$ es la única solución de la ecuación

$$\lambda_{r,\beta}^\circ = 1$$

DEMOSTRACIÓN. Por los lemas 14 y 15 tenemos que

$$I_\beta(0) = \inf_{\theta \in (0,1]} J_\beta^\circ(\theta) = \inf_{\theta \in (0,1]} \widehat{J}_\beta^\circ(\theta) - r^*(\beta).$$

Este último ínfimo es $r^{**}(\beta)$. La prueba de este resultado sigue la misma línea argumentativa que el de la sección anterior. ■

En conclusión, se ha identificado a la función tasa I_β en términos de los eigenvalores $\lambda_{r,\beta}$ y $\lambda_{r,\beta}^\circ$ asociados a los operadores $A_{r,\beta}$ y $A_{r,\beta}^\circ$.

OBSERVACIÓN. $r^{**}(\beta) > r^*(\beta)$, para todo $\beta \in (0, \infty)$.

DEMOSTRACIÓN. Notemos que $A_{r,\beta}^\circ = A_{r,\beta}B$, donde B es la matriz de $\mathbb{N} \times \mathbb{N}$ con entradas $B(i, j) = 1_{\{i=j+1\}}$, $i, j \in \mathbb{N}$.

Entonces

$$\lambda_{r,\beta}^\circ = \|A_{r,\beta}^\circ\| < \|A_{r,\beta}\| \|B\| = \|A_{r,\beta}\| = \lambda_{r,\beta},$$

donde $\|\cdot\|$ es la norma en $\ell^2(\mathbb{N})$. Como ambos eigenvalores son estrictamente crecientes en r y como $r^{**}(\beta)$, $r^*(\beta)$ son los valores cuando estos valen 1, se tiene que $r^{**}(\beta) > r^*(\beta)$. □

De los lemas 11 y 12 tenemos que

$$\widehat{J}_\beta(\theta) - \theta \frac{\partial}{\partial \theta} \widehat{J}_\beta(\theta) = r_\beta(\theta) \quad \text{y} \quad \inf_{\theta \in (0,1]} \widehat{J}_\beta(\theta) = r^*(\beta),$$

Como $J_\beta(\theta) = \widehat{J}_\beta(\theta) - r^*(\beta)$, tenemos que $J_\beta(\theta) - \theta \frac{\partial}{\partial \theta} J_\beta(\theta) = r_\beta(\theta) - r^*(\beta)$ y como $I_\beta(0) = r^{**}(\beta) - r^*(\beta)$, se tiene que $\theta^{**}(\beta)$ es la única solución de la ecuación $r_\beta(\theta) = r^{**}(\beta)$ y como $\theta \rightarrow r_\beta(\theta)$ es decreciente para todo $\beta \in (0, \infty)$ se sigue que $\theta^{**}(\beta) < \theta^*(\beta)$.

4.5. Demostración del resultado principal

El principio de grandes desviaciones para $\mathbb{P}_n^{+,\beta}$ implica los teoremas 1 y 2.

TEOREMA 1. Para todo $\beta \in (0, \infty)$ existe $\theta^*(\beta) \in (0, 1)$ tal que

$$\lim_{n \rightarrow \infty} \mathbb{P}_n^\beta \left(\left| \frac{1}{n} S_n - \theta^*(\beta) \right| < \epsilon \mid S_n \geq 0 \right) = 1, \quad \forall \epsilon > 0.$$

DEMOSTRACIÓN. El lema 12 establece la existencia de $\theta^*(\beta)$, por otro lado; notemos que

$$\begin{aligned} \mathbb{P}_n^\beta \left(\left| \frac{1}{n} S_n - \theta^*(\beta) \right| < \epsilon \mid S_n \geq 0 \right) &= \mathbb{P}_n^\beta \left(\theta^*(\beta) - \epsilon < \frac{1}{n} S_n < \theta^*(\beta) + \epsilon \mid S_n \geq 0 \right) \\ &= \mathbb{P}_n^\beta \left(\frac{1}{n} S_n \in (\theta^*(\beta) - \epsilon, \theta^*(\beta) + \epsilon) \mid S_n \geq 0 \right) = \mathbb{P}_n^{+, \beta}((\theta^*(\beta) - \epsilon, \theta^*(\beta) + \epsilon)) \end{aligned}$$

Como $(\mathbb{P}_n^{+, \beta})_{n \in \mathbb{N}}$ satisface un principio de grandes desviaciones con función tasa I_β tenemos que

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^{+, \beta}((\theta^*(\beta) - \epsilon, \theta^*(\beta) + \epsilon)) &\geq -I_\beta((\theta^*(\beta) - \epsilon, \theta^*(\beta) + \epsilon)) \\ &= - \inf_{\theta \in (\theta^*(\beta) - \epsilon, \theta^*(\beta) + \epsilon)} I_\beta(\theta) = -I_\beta(\theta^*(\beta)) = 0 \end{aligned}$$

y

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^{+, \beta}([\theta^*(\beta) - \epsilon, \theta^*(\beta) + \epsilon]) &\leq -I_\beta([\theta^*(\beta) - \epsilon, \theta^*(\beta) + \epsilon]) \\ &= - \inf_{\theta \in [\theta^*(\beta) - \epsilon, \theta^*(\beta) + \epsilon]} I_\beta(\theta) = -I_\beta(\theta^*(\beta)) = 0 \end{aligned}$$

como $(\theta^*(\beta) - \epsilon, \theta^*(\beta) + \epsilon) \subset [\theta^*(\beta) - \epsilon, \theta^*(\beta) + \epsilon]$

$$\begin{aligned} 0 &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^{+, \beta}((\theta^*(\beta) - \epsilon, \theta^*(\beta) + \epsilon)) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^{+, \beta}([\theta^*(\beta) - \epsilon, \theta^*(\beta) + \epsilon]) \leq 0. \end{aligned}$$

Por tanto

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}_n^{+, \beta}((\theta^*(\beta) - \epsilon, \theta^*(\beta) + \epsilon)) = 0$$

y

$$\lim_{n \rightarrow \infty} \mathbb{P}_n^{+, \beta}((\theta^*(\beta) - \epsilon, \theta^*(\beta) + \epsilon)) = 1$$

■

TEOREMA 2. La función $\beta \rightarrow \theta^*(\beta)$ se puede obtener en términos de un problema variacional y se sigue de la solución al mismo que:

- $\beta \longrightarrow \theta^*(\beta)$ es analítica en $(0, \infty)$,
- $\lim_{\beta \rightarrow 0^+} \theta^*(\beta) = 0$,
- $\lim_{\beta \rightarrow \infty} \theta^*(\beta) = 1$.

DEMOSTRACIÓN. [4] sección 3.5 teorema 2B. Para la analiticidad nos referimos a [1] capítulo 9 y [5] sección 3.

Para el primer límite. Sea x_k el vector $x_k(i) = k^{i-1}$ con $i \in \mathbb{N}$, por la siguiente identidad

$$\sum_{j \in \mathbb{N}} \binom{i+j-2}{i-1} \alpha^{j-1} = \frac{1}{(1-\alpha)^i}, \quad i \in \mathbb{N}, \alpha \in [0, 1)$$

y la definición de $\mathbb{P}_0(i, j)$ tenemos que si $\beta = 0$

$$A_{r,0}x_k = T(r, k)x_{T(r,k)}$$

con $T(r, k) = (2e^{-r} - k)^{-1}$. Se sigue que $\lambda_{r,0}$ es la solución de la ecuación $k = T(r, k)$, esto es

$$\lambda_{r,0} = e^{-r} - (e^{-2r} - 1)^{\frac{1}{2}}, \quad r \leq 0.$$

Ahora elegimos $\epsilon > 0$, como $r \rightarrow \log \lambda_{r,\beta}$ es convexa para todo $\beta > 0$ tenemos que

$$\frac{1}{\theta^*(\beta)} = \frac{1}{\lambda_{r,\beta}} \frac{\partial}{\partial r} \lambda_{r,\beta} \Big|_{r=r^*(\beta)} \geq \frac{1}{\lambda_{r,\beta}} \frac{\partial}{\partial r} \lambda_{r,\beta} \Big|_{r=-\epsilon}$$

donde utilizamos que $r^*(\beta) > 0$ para todo $\beta > 0$, mientras $\beta \rightarrow 0$ el lado derecho de esta última expresión tiende a

$$\frac{1}{\lambda_{r,0}} \frac{\partial}{\partial r} \lambda_{r,0} \Big|_{r=-\epsilon}$$

porque $\beta \rightarrow \lambda_{r,\beta}$ es analítica para todo $r < 0$ ([5], lema 11). Haciendo $\epsilon \rightarrow 0$ y usando el hecho de que $\lambda_{-\epsilon,0}$ y $\frac{\partial}{\partial r} \lambda_{r,0} \Big|_{r=-\epsilon} \rightarrow \infty$ obtenemos que

$$\lim_{\beta \rightarrow 0^+} \theta^*(\beta) = 0$$

Para el segundo límite. Definimos

$$\widehat{\lambda}_{r,\beta} = \lambda_{r+\beta^2,\beta}$$

y tenemos que $A_{r+\beta,\beta} \rightarrow \frac{1}{2}e^r \delta_{(1,1)}$ puntualmente cuando $\beta \rightarrow \infty$, (note que $\mathbb{P}_0(1,1) = \frac{1}{2}$) por tanto

$$\lim_{\beta \rightarrow \infty} \widehat{\lambda}_{r,\beta} = \frac{1}{2}e^r.$$

Ahora, definimos $\widehat{r}^*(\beta)$ tal que

$$\widehat{\lambda}_{\widehat{r}^*(\beta),\beta} = 1,$$

por lo anterior tenemos que $\widehat{r}^*(\beta) \rightarrow \log 2$ cuando $\beta \rightarrow \infty$ y como $\widehat{r}^*(\beta) = r^*(\beta) - \beta^2$ tenemos

$$\frac{1}{\theta^*(\beta)} = \frac{\partial}{\partial r} \lambda_{r,\beta} \Big|_{r=r^*(\beta)} = \frac{\partial}{\partial r} \widehat{\lambda}_{r,\beta} \Big|_{r=\widehat{r}^*(\beta)}.$$

Si $\beta \rightarrow \infty$, el lado derecho de esta expresión tiende a $\frac{\partial}{\partial r} \frac{1}{2}e^r$ evaluado en $r = \log 2$ lo cual es 1 y por tanto

$$\lim_{\beta \rightarrow \infty} \theta^*(\beta) = 1.$$

■

En [5] *R. van der Hofstad and F. den Hollander - Scaling for a random polymer 1995* se prueba que $\theta^*(\beta) \sim C\beta^{\frac{1}{3}}$ para alguna $C > 0$ cuando $\beta \rightarrow 0$, por lo que se tiene la figura 4.1.

Capítulo 5

Comentarios finales

En resumen, definimos una distribución de probabilidad \mathbb{P}_n^β para la configuración espacial de un polímero de longitud n en dimensión 1, el cual no tiene prohibidas las intersecciones pero estas son penalizadas por un factor $e^{-\beta}$; a esto se le conoce como *polímero suave con fuerza de auto-repelencia β* o en términos de las caminatas aleatorias como *caminata aleatoria auto-evasiva débil*.

Primero nos restringimos al caso en el que el polímero es un *puente*, es decir; este se encuentra entre sus puntos extremos, y vimos que esta condición no afecta significativamente la definición de \mathbb{P}_n^β .

Al número de intersecciones en cada trayectoria le asociamos una cadena de Markov en términos del número de visitas a cada estado y cuyas probabilidades de transición están relacionadas con una caminata aleatoria con deriva θ , ya que si se condiciona a que el polímero esté en el estado $\lceil \theta n \rceil$ al tiempo n , la probabilidad de este evento puede ser aproximada por medio de la distribución de probabilidad de esta caminata.

Encontramos una equivalencia entre el número de visitas a cada estado y la medida empírica por pares L_N^2 asociada a la cadena de Markov y derivamos el principio de grandes desviaciones para $\mathbb{P}_n^{\beta, \text{puente}}$ a través del principio de grandes desviaciones para L_N^2 . Caracterizamos la función tasa en términos de una familia de aigevalores $\lambda_{r, \beta}$ asociados a cierta matriz $A_{r, \beta}$ de $\mathbb{N} \times \mathbb{N}$ y vimos que tiene la siguiente gráfica

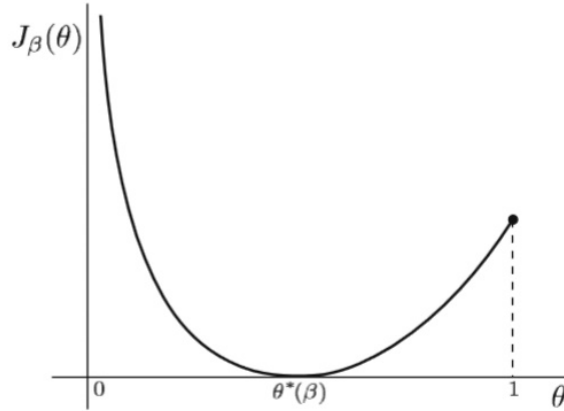


Figura 5.1: Función tasa J_β para los polímeros puente [3].

Después se removió la condición de puente y se estableció un resultado similar para $\mathbb{P}_n^{+, \beta}$ y resulta que la función tasa asociada a este principio de grandes desviaciones tiene la misma forma que en el resultado anterior salvo para valores de θ menores a una *velocidad crítica* θ^{**} , esto es porque cuando la trayectoria tiene velocidades menores a este valor crítico, la penalización por permanecer como un puente es menor que la penalización por realizar lazos alrededor de sus puntos inicial y final.

Este principio de grandes desviaciones nos ayuda a probar que

$$\lim_{n \rightarrow \infty} \mathbb{P}_n^\beta \left(\left| \frac{1}{n} S_n - \theta^*(\beta) \right| < \epsilon \mid S_n \geq 0 \right) = 1, \quad \forall \epsilon > 0$$

es decir, con probabilidad 1 para cada fuerza de autorepelencia β existe una velocidad a la que el polímero tiende a alejarse de su punto inicial conforme la cantidad de monómeros que lo constituye aumenta, y con ayuda de la gráfica de $\theta \rightarrow \theta^*(\beta)$

podemos ver que si a fuerza de autorepelencia es muy pequeña, la velocidad a la que el polímero se aleja tiende a cero, es decir; el polímero tiende a formar pequeñas esferas, y si esta fuerza aumenta, la velocidad tiende a 1, la velocidad con la que se aleja es mayor (la trayectoria da un paso de longitud 1 lejos de su punto inicial por cada unidad de tiempo).

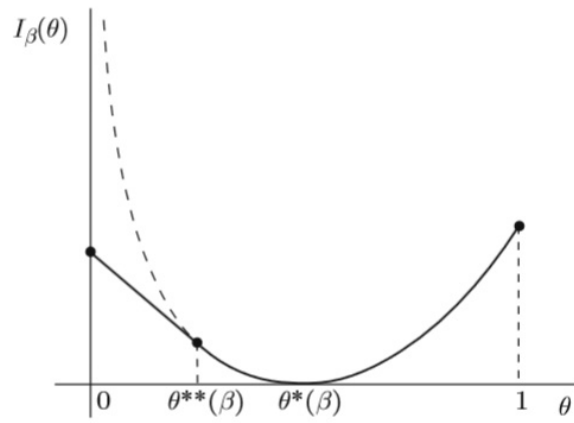


Figura 5.2: Función tasa I_β en comparación con J_β de los polímeros puente [3].

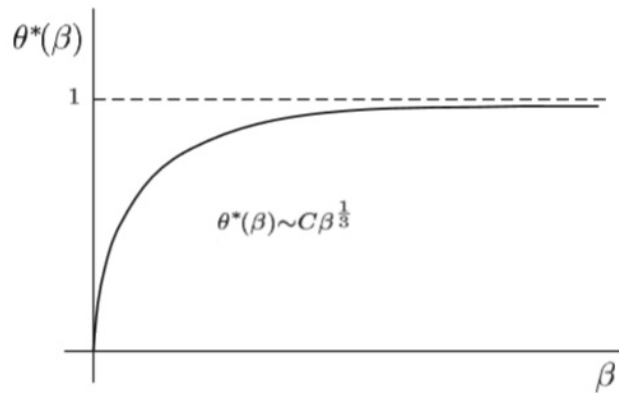


Figura 5.3: Velocidad lineal de un polímero suave [3].

Los polímeros amorfos como el poliestireno se pueden ver como ejemplos de este comportamiento, a temperaturas altas se encuentran en forma de un líquido viscoso y al enfriarlos se convierten en sólidos duros, rígidos y frágiles; lo que sucede es que conforme la temperatura disminuye el polímero se contrae porque las moléculas se mueven menos y por tanto se atraen más.

Apéndice

Sea $(S_n)_{n \in \mathbb{N}_0}$ una cadena de Markov en \mathbb{Z} , $p_{ij}(n)$ la probabilidad de ir del estado i al estado j en n pasos y $f_{ij}(n)$ la probabilidad de que estando en el estado i la caminata visite el estado j por primera vez en exactamente n pasos, es decir;

$$p_{ij}(n) = \mathbb{P}(S_n = j \mid S_0 = i)$$

$$f_{ij}(n) = \mathbb{P}(S_n = j, S_{n-1} \neq j, \dots, S_1 \neq j \mid S_0 = i)$$

y se define $f_{ij}(0) = 0$, incluyendo el caso $i = j$.

PROPOSICIÓN. para cada $n \geq 1$, $p_{ij}(n) = \sum_{k=1}^n f_{ij}(k)p_{jj}(n-k)$.

DEMOSTRACIÓN. Consideremos los siguientes eventos

$$A_1 = (S_n = j, S_1 = j, S_0 = i)$$

$$A_2 = (S_n = j, S_2 = j, S_1 \neq j, S_0 = i)$$

$$A_3 = (S_n = j, S_3 = j, S_2 \neq j, S_1 \neq j, S_0 = i)$$

\vdots

$$A_n = (S_n = j, S_{n-1} \neq j, \dots, S_2 \neq j, S_1 \neq j, S_0 = i)$$

Notemos que $A_i \cap A_j = \emptyset$ para todo $i \neq j$ y

$$A_1 \cup A_2 \cup \dots \cup A_n = \{S_n = j, S_0 = i\},$$

como $(S_n)_{n \in \mathbb{N}_0}$ es una cadena de Markov

$$\mathbb{P}(A_k) = \mathbb{P}(S_n = j, S_k = j, S_{k-1} \neq j, \dots, S_1 \neq j, S_0 = i)$$

$$\begin{aligned}
&= \mathbb{P}(S_n = j \mid S_k = j, S_{k-1} \neq j, \dots, S_1 \neq j, S_0 = i) \mathbb{P}(S_k = j, S_{k-1} \neq j, \dots, S_1 \neq j, S_0 = i) \\
&= \mathbb{P}(S_n = j \mid S_k = j) \mathbb{P}(S_k = j, S_{k-1} \neq j, \dots, S_1 \neq j \mid S_0 = i) \mathbb{P}(S_0 = i) \\
&= \mathbb{P}(S_{n-k} = j \mid S_0 = i) f_k \mathbb{P}(S_0 = i) = p_{n-k} f_k \mathbb{P}(S_0 = 0)
\end{aligned}$$

por lo que

$$\mathbb{P}(S_n = 0, S_0 = 0) = \mathbb{P}\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n \mathbb{P}(A_k) = \sum_{k=1}^n p_{n-k} f_k \mathbb{P}(S_0 = 0)$$

por lo tanto

$$p_n = \sum_{k=0}^n f_k p_{n-k}$$

□

LEMA DE ABEL. Si la serie $\sum_{k=0}^{\infty} a_k$ converge, entonces

$$\lim_{t \rightarrow 1^-} \sum_{k=0}^{\infty} a_k t^k = \sum_{k=0}^{\infty} a_k.$$

OBSERVACIÓN. $Q_n(a) = \max_{k \geq an} \binom{n}{\lceil an \rceil}$.

Esto se ve realizando distintas iteraciones y utilizando el triángulo de Pascal.

PROPOSICIÓN. Sea X una variable aleatoria con función generadora de momentos $\varphi(t) < \infty$ para cada $t \in (-s, s)$, para algún $s > 0$; entonces

1. Todos los momentos de X son finitos.

2. $\varphi(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbb{E}[X^n]$.

3. $\varphi(t)$ tiene derivadas continuas de cualquier orden en $(-s, s)$, y se cumple que

$$\left. \frac{d^n}{dt^n} \varphi(t) \right|_{t=0} = \mathbb{E}[X^n].$$

DEMOSTRACIÓN.

1. La prueba se basa en las siguientes identidades

$$\mathbb{E}[|X|^n] = n \int_0^\infty (1 - F(x))x^{n-1}dx + n \int_{-\infty}^0 F(x)|x|^{n-1}dx,$$

$$\varphi(t) = 1 + t \int_0^\infty (1 - F(x))e^{tx}dx - t \int_{-\infty}^0 F(x)e^{tx}dx,$$

en donde, por hipótesis, las dos integrales de $\varphi(t)$ son finitas para todo $t \in (-s, s)$. Veremos que cada integral en la expresión de $\mathbb{E}[|X|^n]$ es menor o igual a la correspondiente integral de $\varphi(t)$.

Para el caso $x > 0$ se toma cualquier $t \in (0, s)$, y entonces

$$\frac{(tx)^n}{n!} \leq e^{tx} \quad \text{es decir,} \quad x^n \leq \left(\frac{n!}{t^n}\right) e^{tx}.$$

De modo que, salvo constantes, la primer integral de $\mathbb{E}[|X|^n]$ es menor o igual a la primer integral de $\varphi(t)$, siendo esta última finita, la primera también lo es.

Para el caso $x < 0$ conviene tomar $t \in (-s, 0)$, pues en tal caso $tx > 0$ y entonces

$$\frac{|tx|^n}{n!} \leq e^{|tx|} = e^{tx} \quad \text{es decir,} \quad |x|^n \leq \left(\frac{n!}{|t|^n}\right) e^{tx}.$$

Ahora la segunda integral de $\mathbb{E}[|X|^n]$ es menor o igual a la segunda integral de $\varphi(t)$, siendo esta última finita, la primera también lo es.

De esta forma todos los momentos de X existen cuando $\varphi(t)$ es finita en algún intervalo no trivial alrededor de cero.

2. Se usa la fórmula

$$\mathbb{E}[X^n] = n \int_0^\infty (1 - F(x))x^{n-1}dx - n \int_{-\infty}^0 F(x)x^{n-1}dx.$$

FUNCIÓN TASA PARA ALGUNAS DISTRIBUCIONES CONOCIDAS

1. Si X tiene una distribución *Poisson* de parámetro $\lambda > 0$

$$\begin{aligned}\varphi(t) &= \mathbb{E} [e^{tX}] = \sum_{x=0}^{\infty} e^{tx} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{-\lambda} e^{\lambda e^t} \\ &\Rightarrow \varphi(t) = e^{\lambda(e^t-1)}, \quad \forall t \in \mathbb{R}\end{aligned}$$

Sea $g(t) = zt - \log \varphi(t)$, en este caso

$$g(t) = zt - \log e^{\lambda(e^t-1)} = zt - \lambda(e^t - 1) \quad \text{y} \quad g'(t) = z - \lambda e^t$$

entonces $g'(t) = 0$ implica que $z - \lambda e^t = 0$ por lo que $t_0 = \log z - \log \lambda$ es un punto crítico de g .

Además $g''(t) = -\lambda e^t$, entonces

$$g''(t_0) = -\lambda e^{\log z - \log \lambda} = -\lambda \left(e^{\log \frac{z}{\lambda}} \right) = -\lambda \left(\frac{z}{\lambda} \right) = -z < 0, \quad \text{si } z > 0$$

es decir; $t_0 = \log z - \log \lambda$ es un punto máximo de g si y sólo si $z > 0$, por tanto

$$\begin{aligned}I(z) &= \sup_{t \in \mathbb{R}} [zt - \log \varphi(t)] = \sup_{t \in \mathbb{R}} [zt - \lambda(e^t - 1)] = zt_0 - \lambda(e^{t_0} - 1) \\ &= z(\log z - \log \lambda) - \lambda \left(e^{\log z - \log \lambda} - 1 \right) = z \log z - z \log \lambda - \lambda \left(\frac{z}{\lambda} - 1 \right) \\ &= z \log z - z \log \lambda - z + \lambda\end{aligned}$$

$$\therefore I(z) = \begin{cases} z \log z - z \log \lambda - z + \lambda & \text{si } z \geq 0 \\ \infty & \text{en otro caso} \end{cases}$$

Note que

$$I'(z) = \log z + z \left(\frac{1}{z} \right) - \log \lambda - 1 = \log z - \log \lambda$$

entonces $I'(z) = 0$ implica que $\log z = \log \lambda$ por lo que $z_0 = \lambda$ es un punto crítico de I .

Además $I''(z) = \frac{1}{z}$, entonces $I''(z_0) = \frac{1}{\lambda} > 0$, es decir; $z_0 = \lambda = \mathbb{E}[X]$ es un punto mínimo de I .

Luego $I(z_0) = I(\lambda) = \lambda \log \lambda - \lambda \log \lambda - \lambda + \lambda = 0$, por tanto

$I(z) \geq 0$, $\forall z \geq 0$ cumpliéndose la igualdad si y sólo si $z = \lambda = \mathbb{E}[X]$

$$I''(\lambda) = \frac{1}{\lambda} = \frac{1}{\text{Var}(X)}$$

2. Si X tiene una distribución *Exponencial* de parámetro $\lambda > 0$

$$\begin{aligned} \varphi(t) &= \mathbb{E}[e^{tX}] = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} e^{x(t-\lambda)} dx = \lambda \left(\frac{1}{t-\lambda} e^{(t-\lambda)x} \Big|_0^{\infty} \right) \\ &= \begin{cases} \lambda \left(\frac{-1}{t-\lambda} \right) & \text{si } t-\lambda < 0 \\ \infty & \text{si } t-\lambda \geq 0 \end{cases} \\ \Rightarrow \varphi(t) &= \begin{cases} \frac{\lambda}{\lambda-t} & \text{si } t < \lambda \\ \infty & \text{si } t \geq \lambda \end{cases} \end{aligned}$$

Sea $g(t) = zt - \log \varphi(t)$ con $\varphi(t) < \infty$, en este caso

$$g(t) = zt - \log \left(\frac{\lambda}{\lambda-t} \right) = zt - \log \lambda + \log(\lambda-t) \quad \text{y} \quad g'(t) = z - \frac{1}{\lambda-t}$$

entonces $g'(t) = 0$ implica que $z - \frac{1}{\lambda-t} = 0$ por lo que $t_0 = \lambda - \frac{1}{z}$ es un punto crítico de g .

Además $g''(t) = \frac{-1}{(\lambda-t)^2}$, entonces

$$g''(t_0) = \frac{-1}{\left(\lambda - \lambda + \frac{1}{z}\right)^2} = -z^2 < 0$$

es decir; $t_0 = \lambda - \frac{1}{z}$ es un punto máximo de g
por tanto

$$I(z) = \sup_{t \in \mathbb{R}} [zt - \log \varphi(t)] = \sup_{t \in \mathbb{R}} [zt - \log \lambda + \log(\lambda-t)] = zt_0 - \log \lambda + \log(\lambda-t_0)$$

$$\begin{aligned}
&= z \left(\lambda - \frac{1}{z} \right) - \log \lambda + \log \left(\lambda - \lambda + \frac{1}{z} \right) = z\lambda - 1 - \log \lambda + \log \left(\frac{1}{z} \right) \\
&= z\lambda - 1 - \log \lambda - \log z
\end{aligned}$$

$$\therefore I(z) = \begin{cases} \lambda z - 1 - \log \lambda - \log z & \text{si } z > 0 \\ \infty & \text{en otro caso} \end{cases}$$

Note que

$$I'(z) = \lambda - \frac{1}{z}$$

entonces $I'(z) = 0$ implica que $\frac{1}{z} = \lambda$ por lo que $z_0 = \frac{1}{\lambda}$ es un punto crítico de I .

Además $I''(z) = \frac{1}{z^2}$, entonces $I''(z_0) = \lambda^2 > 0$, es decir; $z_0 = \frac{1}{\lambda} = \mathbb{E}[X]$ es un punto mínimo de I .

Luego $I(z_0) = I\left(\frac{1}{\lambda}\right) = \lambda\left(\frac{1}{\lambda}\right) - 1 - \log \lambda - \log\left(\frac{1}{\lambda}\right) = 0$, por tanto

$I(z) \geq 0$, $\forall z > 0$ cumpliéndose la igualdad si y sólo si $z = \frac{1}{\lambda} = \mathbb{E}[X]$

$$I''\left(\frac{1}{\lambda}\right) = \lambda^2 = \frac{1}{\text{Var}(X)}$$

3. Si X tiene una distribución *Normal* de parámetros μ, σ^2

$$\begin{aligned}
\varphi(t) &= \mathbb{E}[e^{tX}] = \int_{-\infty}^{\infty} e^{xt} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x^2-2x\mu+\mu^2-2\sigma^2xt)}{2\sigma^2}} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x^2-2(\mu+\sigma^2t)x+\mu^2)}{2\sigma^2}} dx = e^{\frac{(\mu+\sigma^2t)^2-\mu^2}{2\sigma^2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-(\mu+\sigma^2t))^2}{2\sigma^2}} dx \\
&= e^{\frac{\mu^2+2\mu\sigma^2t+\sigma^4t^2-\mu^2}{2\sigma^2}} = e^{\frac{2\mu\sigma^2t+\sigma^4t^2}{2\sigma^2}} \Rightarrow \varphi(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}, \forall t \in \mathbb{R}
\end{aligned}$$

Sea $g(t) = zt - \log \varphi(t)$, en este caso

$$g(t) = zt - \log e^{\mu t + \frac{\sigma^2 t^2}{2}} = (z - \mu)t - \frac{\sigma^2 t^2}{2} \quad \text{y} \quad g'(t) = z - \mu - \sigma^2 t$$

entonces $g'(t) = 0$ implica que $z - \mu - \sigma^2 t = 0$ por lo que $t_0 = \frac{z-\mu}{\sigma^2}$ es un punto crítico de g .

Además $g''(t) = -\sigma^2$ para todo $t \in \mathbb{R}$; en particular $g''(t_0) = -\sigma^2 < 0$ por lo que $t_0 = \frac{z-\mu}{\sigma^2}$ es un punto máximo de g .

Por tanto

$$\begin{aligned} I(z) &= \sup_{t \in \mathbb{R}} [zt - \log \varphi(t)] = \sup_{t \in \mathbb{R}} \left[(z - \mu)t - \frac{1}{2}\sigma^2 t^2 \right] = (z - \mu)t_0 - \frac{1}{2}\sigma^2 t_0^2 \\ &= (z - \mu) \left(\frac{z - \mu}{\sigma^2} \right) - \frac{1}{2}\sigma^2 \left(\frac{z - \mu}{\sigma^2} \right)^2 = \frac{(z - \mu)^2}{\sigma^2} - \frac{1}{2} \frac{(z - \mu)^2}{\sigma^2} \end{aligned}$$

$$\therefore I(z) = \frac{1}{2\sigma^2}(z - \mu)^2, \quad \forall z \in \mathbb{R}$$

Note que

$$I'(z) = \frac{z - \mu}{\sigma^2}$$

entonces $I'(z) = 0$ implica que $z = \mu$ por lo que $z_0 = \mu$ es un punto crítico de I .

Además $I''(z) = \frac{1}{\sigma^2}$, entonces $I''(z_0) = \frac{1}{\sigma^2} > 0$, es decir; $z_0 = \mu = \mathbb{E}[X]$ es un punto mínimo de I .

Luego $I(z_0) = I(\mu) = \frac{1}{2\sigma^2}(\mu - \mu)^2 = 0$, por tanto

$I(z) \geq 0$, $\forall z \in \mathbb{R}$ cumpliéndose la igualdad si y sólo si $z = \mu = \mathbb{E}[X]$

$$I''(\mu) = \frac{1}{\sigma^2} = \frac{1}{\text{Var}(X)}$$

Bibliografía

- [1] F. den Hollander, *Large Deviations*, American Mathematical Society, Fields Institute Monographs 14, Providence, RI, 2000.
- [2] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, (2nd. ed.), Springer, New York, 1998.
- [3] F. den Hollander, *Random Polymers*, Lecture Notes in Mathematics 1974, Springer-Verlag Berlin Heidelberg, 2009.
- [4] A. Greven and F. den Hollander, *Variational characterization of the speed of a one-dimensional self-repellent random walk*, Ann. Appl. Probab. 3 (1993) 1067-1099.
- [5] J. Baillon, Clément, P. Greven and F. den Hollander, *On a variational problem for an infinite particle system in a random medium*, Part I: *The global growth rate*. Preprint, 1991.
- [6] R. van der Hofstad and F. den Hollander, *Scaling for a random polymer*, Commun. Math. Phys. 169 (1995) 397-440.
- [7] N Madras and G. Slade, *The Self-Avoiding Random Walk*, Birkhäuser, Boston, 1993.
- [8] F. Spitzer, *Principles of Random Walk* (2nd. ed.), Springer, New York, 1976.
- [9] Hoel P. G., Port S. C., Stone C.J., *Introduction to stochastic processes*, Houghton Mifflin Company, 1972.
- [10] K. Petersen, *Ergodic Theory*, Cambridge University Press, Cambridge, 1983.

- [11] L. Rincón, *Introducción a los procesos estocásticos*, Las prensas de ciencias, UNAM, 2012.
- [12] Greenwood, E.A. Milne, *Natural High Polymers*, Contemporary Science Paperbacks 18, Oliver & Boyd, Edinburgh, 1968.