

Trabajo Especial de Licenciatura en Física

**ANÁLISIS DE LAS DISTRIBUCIONES DE
RETORNOS Y EFECTOS DE TAMAÑO
FINITO PARA UN MODELO DE URNAS**

Autora: María José Domenech

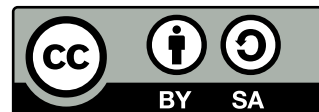
Directora: Dra. Silvia Menchón



Universidad Nacional de Córdoba
Facultad de Matemática, Astronomía, Física y Computación

Marzo 2024

Esta obra está bajo una licencia [Creative Commons](https://creativecommons.org/licenses/by-sa/4.0/)
“Atribución-CompartirIgual 4.0 Internacional”.



Resumen

En este trabajo se presenta un modelo de urnas que combina dos modelos clásicos: el de Ehrenfest y el del votante, ambos modelos estocásticos que han sido ampliamente utilizados para describir sistemas físicos. En cada uno se consideran dos urnas y un número finito N de elementos distribuidos en ellas. Cada modelo define una evolución temporal diferente que modifica de manera aleatoria la distribución de los elementos. En particular, en cada paso temporal, hay una probabilidad α de que la evolución sea la correspondiente al modelo de Ehrenfest y $1 - \alpha$ de que sea del modelo del votante. Se supone que para este sistema la distribución de los tamaños de fluctuaciones respecto al estado de equilibrio sigue una ley de potencias cuando N tiende a infinito. En este trabajo verificamos numéricamente dicho comportamiento para diferentes valores de N y α . Además analizamos la distribución de retornos, que se define como la distribución de la diferencia de fluctuaciones consecutivas. Esta distribución está asociada a distribuciones del tipo q -Gaussiana cuando la distribución de las fluctuaciones sigue una ley de potencias. Relacionando el parámetro q con el exponente de la ley de potencias y considerando los efectos de tamaño finito, comprobamos numéricamente que las distribuciones de retorno pueden relacionarse con una combinación de Gaussianas y q -Gaussianas, donde el aporte Gaussiano disminuye a medida que N aumenta.

Palabras Clave: Distribuciones de retorno, Caminatas aleatorias, Modelos de Urna, Ley de Potencia, Tamaño Finito.

Agradecimientos

Quiero agradecer el haber sido parte de una institución tan linda, que sin dudas lo es gracias al esfuerzo y la vocación de las personas que la componen. A Silvia, mi directora, por trabajar conmigo incluso a la distancia. A mis profesores y sus incontables clases hermosas, donde disfruté de sentir curiosidad y ganas de aprender. Me siento orgullosa de formar parte de la universidad pública argentina, excelente y humana.

Agradezco a mis amigos y compañeros, tan importantes, verlos fue mi motivación para seguir en los tiempos más cansadores. También son responsables de los momentos más divertidos, que alegran el corazón. A mis amigas además, decirles que son compañeras de la vida de una manera única, pilar fundamental. A Andrés, mi compañero de infinitas cosas, le agradezco porque nada hubiese sido lo mismo sin él. Estuvo siempre a mi lado con muchísimo amor.

Gracias a mi mamá y papá, que les debo todo. Por eso, este trabajo también es suyo. También a mis hermanas, que crecimos juntas en esta etapa y aprendimos mucho de la vida y la convivencia, pero también de la compañía.

A mis hermanas, primas, mamá, tía, amigas, cuñada, suegra, abuelas... mujeres de mi vida, una más fuerte que la otra, decirles que son la fuente de mi inspiración.

Inmensamente feliz y agradecida.

Dedicado a mi abuela, la Cris.

«Este trabajo utilizó recursos computacionales del CCAD de la Universidad Nacional de Córdoba (<https://ccad.unc.edu.ar/>), que forman parte del SNCAD del MinCyT de la República Argentina.»

Contenidos

Introducción	9
1 Marco Teórico	11
1.1 Modelos de Urnas	11
1.1.1 Planteo del problema general	12
1.1.2 Modelo de Ehrenfest	13
1.1.3 Modelo del Votante	14
1.2 Caminatas Aleatorias	16
1.2.1 Caminatas aleatorias como cadenas de Markov	17
1.2.2 Distribuciones estacionarias de probabilidad	19
1.2.3 El modelo de Ehrenfest y del votante como caminatas aleatorias	19
1.2.4 Probabilidad de primer pasaje y distribución de retornos	22
1.3 Comportamiento tipo ley de potencia	24
2 Modelo Físico	26
2.1 Noisy Voter Model	26
2.1.1 Ejemplo	28
2.1.2 El modelo como una caminata aleatoria	29
3 Resultados y Discusión	33
3.1 Resultados en la región supracrítica ($\alpha > 1/N$)	33
3.2 Resultados en el pseudo punto crítico ($\alpha = 1/N$)	41
Conclusiones	46

Introducción

Los sistemas complejos se caracterizan por estar formados por una gran cantidad de componentes interactuantes que, como un conjunto, presentan conductas macroscópicas que son difíciles de deducir y/o relacionar con sus correlaciones microscópicas. Estos sistemas están presentes en campos tan diversos como, por ejemplo, el estudio de redes neuronales, de ecosistemas, de opiniones en una comunidad, de flúidos o del comportamiento de la bolsa de valores. En muchos casos, un modelo definido con simples reglas de interacción entre sus elementos, puede constituir un sistema complejo y llevarnos a observar conductas no evidentes y difíciles de determinar.

En este trabajo, planteamos un modelo de urna que es combinación de otros dos modelos para los cuales ya conocemos su comportamiento: el modelo de Ehrenfest y el modelo del votante. El nuestro se compone de N elementos distribuidos en dos urnas A y B . La distribución de los elementos cambia de manera aleatoria en el tiempo según las reglas de evolución del modelo de Ehrenfest o las del votante, con probabilidad α y $1 - \alpha$, respectivamente. El mecanismo de evolución específico aplicado en cada caso será explicado en los capítulos siguientes. El modelo a considerar en este trabajo será denominado *Noisy Voter Model* porque es un caso particular del modelo planteado por Kirman en 1993 [1].

Los modelos de urna constituyen una herramienta simple pero poderosa que busca representar el comportamiento de un sistema complejo a través de la evolución de estados discretos a lo largo del tiempo. Son mayormente empleados en el estudio de sistemas que involucran variables aleatorias, como la dinámica de precios el mercado y en modelos de opinión. En el primer caso, cada urna podría representar, por ejemplo, la riqueza de un individuo y las reglas de transición podrían describir cómo esa riqueza cambia con el tiempo debido a factores como el intercambio comercial. En el segundo caso, cada urna podría representar una opinión, y las reglas de transición podrían modelar cómo esas opiniones cambian a medida que las personas interactúan entre sí. En este sentido, estudiando el proceso estocástico involucrado podemos inferir propiedades de la dinámica del mismo, que nos sirven para comprender e intentar predecir la conducta de sistemas reales.

En este trabajo también caracterizamos nuestro modelo de urnas con una caminata aleatoria unidimensional, lo que constituye una herramienta crucial para comprender y analizar su dinámica. Nos permite estudiar distribuciones de probabilidad relacionadas

a magnitudes del sistema que serán detalladas en las secciones posteriores. Una de las distribuciones que estudiaremos, tendrá comportamiento tipo ley de potencia. Este tipo de comportamiento es particularmente relevante, porque emerge en muchos sistemas complejos cuya dinámica y propiedades microscópicas son muy diferentes entre sí, lo cual, sugiere la existencia de un origen generalizado que aún no se conoce en la actualidad. Otras de las distribuciones que estudiaremos en este trabajo tienen la forma de q -Gaussianas. Las mismas son de interés porque son distribuciones estándar en el estudio recientemente desarrollado de la entropía no aditiva de Tsallis [2], cuyo enfoque busca generalizar las herramientas de la mecánica estadística para otros tipos de sistemas más complejos y fuera del equilibrio termodinámico usual. Entender la conexión de este tipo de distribuciones puede significar un camino hacia la comprensión de la complejidad inherente de este tipo de sistemas, revelando regularidades que pueden escapar a un análisis superficial.

En definitiva, la simplicidad de este tipo de modelos puede limitar su capacidad para capturar todos los aspectos de la realidad, pero a la vez pueden ofrecer intuiciones valiosas sobre los mecanismos subyacentes que impulsan el comportamiento observado. Además, pueden servir como punto de partida para investigaciones más sofisticadas y análisis más detallados.

Este trabajo final se organiza de la siguiente manera: en el primer capítulo, introduciremos el marco teórico necesario para caracterizar el comportamiento del modelo de Ehrenfest, del votante y el *Noisy Voter Model*. Allí expondremos herramientas de la teoría de caminatas aleatorias que utilizaremos, junto a otras que completan la descripción y pueden ser de utilidad para futuros trabajos. En el segundo capítulo, describiremos nuestro modelo en detalle, utilizando desarrollos teóricos de otros autores y mostraremos la presencia de un pseudo punto crítico para el valor de $\alpha = 1/N$, que marca dos regiones con patrones de comportamiento diferentes. En el tercer capítulo presentaremos los resultados computacionales obtenidos junto a un análisis de los mismos, donde relacionaremos el exponente de la ley de potencias de la distribución de primer retorno, con el parámetro q de la q -Gaussiana relacionada a la denominada distribución de retornos. Además, la observación del comportamiento asintótico de estas distribuciones nos permitirá sacar conclusiones sobre cómo influye el tamaño finito del sistema, y cómo el mismo podría comportarse en el caso límite de $N \rightarrow \infty$.

Capítulo 1

Marco Teórico

1.1 Modelos de Urnas

Los denominados modelos de urnas se han utilizado desde épocas muy antiguas y constituyen un enfoque comúnmente utilizado en teoría de probabilidad y mecánica estadística. Como señalamos en la introducción, resultan muy útiles para la comprensión de algunos sistemas dinámicos complejos [3]. Su nombre proviene de la analogía que se establece entre las componentes del sistema a describir y sus estados individuales, con urnas o cajas que contienen elementos que pueden ser de distinta naturaleza. En principio, estos modelos pueden componerse por cualquier número de urnas y de elementos. Esos elementos pueden tener diferentes características; y en general, esas diferencias se asocian con colores. A través de procesos aleatorios y reglas específicas se seleccionan los elementos de las urnas y se establecen acciones sobre ellos, como puede ser la adición o extracción de los mismos de alguna urna, y/o cambios en su color. La elección de estas reglas definen la dinámica del modelo y de ellas dependerá qué sistema físico puede representarse. Si bien las posibilidades son muy diversas, en este trabajo nos centramos en una clase específica: modelos de solo dos urnas con transferencia de elementos entre ellas.

Un número considerable de resultados de la teoría de probabilidad pueden ser derivados de forma simple utilizando modelos de urna. Por ejemplo, supongamos que tenemos una sola urna que contiene N elementos, de los cuales n_1 son rojos y n_2 son azules. Si definimos como regla sistemática el sacar un elemento al azar de la urna y volverlo a colocar en la misma, y lo repetimos n veces, obtenemos una secuencia *RARAA...* de colores de los elementos. Podemos considerar como variable de interés la cantidad x de elementos rojos que salen al cabo de repetir el experimento n veces. Es decir, nos interesa la distribución de probabilidad de la variable aleatoria X que representa el número de elementos rojos x entre los n ¹. Esta distribución es muy conocida y se obtiene al tener en cuenta que los N elementos tienen la misma probabilidad de ser tomados: la probabilidad de que un

¹La notación utilizada es la usual en estadística: sea X una variable aleatoria o estocástica que puede tomar valores x que pertenecen a un conjunto numerable (en este caso finito) denominado espacio muestral $S = \{x_1, x_2, \dots\}$, es decir, $X(S) = x_1, x_2, \dots$. La distribución de probabilidad $P(x)$ de X es la probabilidad de que X tome el valor x , y viene dada por la suma de las probabilidades de todos los puntos muestrales de S para los cuales X toma el valor x y satisface $0 \leq P(x) \leq 1 \quad \forall x$.

elemento individual sea tomado es $1/N$, la probabilidad de tomar uno rojo es $p = n_1/N$ y la de uno azul $q = n_2/N = 1 - p$. Así, teniendo en cuenta también las formas de ordenar la secuencia $RARAA\dots$ con el coeficiente binomial, obtenemos:

$$P(X = x) = \binom{n}{x} p^x q^{n-x},$$

que es la conocida distribución binomial.

A continuación detallamos la clase de modelos de urna que utilizaremos en este trabajo, para luego plantear desde esta perspectiva al modelo de Ehrenfest y al del votante.

1.1.1 Planteo del problema general

En nuestro caso particular consideramos un sistema compuesto por dos urnas, las cuales llamamos A y B ; y por N elementos idénticos repartidos en las mismas. Si bien en este caso los elementos son todos iguales, no son indistinguibles, ya que los mismos están numerados. Decimos que están rotulados con los números $1, 2, 3, \dots, N$. Esto significa que cada elemento es único y solo puede estar ubicado en la urna A o en la B , lo cual puede pensarse como un sistema cuántico de dos niveles de N partículas distinguibles. Denotamos al estado del elemento i por s_i , con $i = 1, \dots, N$. Como solo hay dos estados posibles para cada elemento, dado que puede estar colocado o bien en la urna A o bien en la urna B , decimos que s_i toma el valor 1 si el elemento está ubicado en la urna A y que s_i toma el valor 0 si está ubicado en la urna B . Así, el estado total del sistema queda caracterizado por un conjunto específico de los estados s_i , que representa una configuración concreta de los elementos i en las urnas. Definimos el *vector de ocupación*, $\mathbf{v} = (s_1, s_2, \dots, s_N)$, como el vector que describe el estado total del sistema.

En todos los modelos que presentamos en este trabajo, la primera variable de interés considerada es la cantidad de elementos en cada urna. Denotamos por n_A y n_B la cantidad de elementos en la urna A y B , respectivamente. Como la cantidad total N de elementos es constante, solo una de estas variables es independiente ($n_B = N - n_A$). Por lo tanto, utilizamos n_A para describir esta característica del sistema. Dada la definición de los estados s_i , podemos obtener n_A a través de la suma de todos los estados s_i , es decir,

$$n_A = \sum_{i=1}^N s_i \tag{1.1}$$

Finalmente, definimos para cada modelo un conjunto de acciones que se aplican a cada paso temporal Δt , el cual, sin pérdida de generalidad, es considerado como una unidad temporal ($\Delta t = 1$). La implementación sucesiva de varios pasos temporales, donde se aplica el conjunto de acciones, constituye la dinámica del sistema. De esta forma, las cantidades que caracterizan al sistema son variables del tiempo ($\mathbf{v}(t)$, $n_A(t)$) y su variación $n_A(t+1) - n_A(t)$ dependerá exclusivamente de la dinámica específica considerada.

En nuestro trabajo realizamos simulaciones numéricas de la evolución temporal de los modelos de urna. Independientemente de las reglas a implementar, o proceso estocástico a

considerar, en general vamos a inicializar nuestras simulaciones desde un estado con igual cantidad de elementos en cada urna, lo cual implica que $\mathbf{v}(t = 0) = (1, \dots, 1, 0, \dots, 0)$. Esto significa que ubicamos los primeros $N/2$ elementos en la urna A , es decir que tienen estado $s_i = 1$; y colocamos los segundos $N/2$ elementos en la urna B , y en consecuencia su estado es $s_i = 0$. En este sentido es conveniente considerar N como un número par. La inicialización del sistema puede ser implementada como mostramos en el siguiente pseudocódigo, donde se genera el vector \mathbf{v} con las condiciones iniciales mencionadas.

Algoritmo Inicialización del sistema

Input N número natural par.

Output \mathbf{v} vector cuya entrada i , $v[i]$, se corresponde con s_i , el estado del elemento i .

$v[1], v[2], \dots, v[\frac{N}{2}] = 1$ \triangleright Los $N/2$ primeros elementos están en la urna A .
 $v[\frac{N}{2} + 1], \dots, v[N] = 0$ \triangleright Los $N/2$ segundos elementos están en la urna B .

1.1.2 Modelo de Ehrenfest

El modelo de Ehrenfest, formulado en 1907 por Tatiana y Paul Ehrenfest, es considerado un clásico en la teoría de la probabilidad [4]. Puede ser descrito utilizando los términos presentados en la sección anterior, dado que consiste de dos urnas A y B con N elementos idénticos rotulados. La motivación detrás de este modelo fue estudiar la transferencia de calor entre dos cuerpos aislados separados por una membrana permeable. En esta representación, las urnas simbolizan los dos cuerpos y las moléculas que pueden atravesar la membrana representan los elementos que se transfieren entre las dos urnas. El proceso estocástico asociado a este modelo puede resumirse mediante la siguiente acción: en cada instante de tiempo t , se selecciona un elemento de manera aleatoria y se cambia de urna. Es decir, en cada paso temporal ocurre lo siguiente:

1. Se sortea un número entero m entre 1 y N .
2. El elemento con rótulo m se cambia de urna.

En este caso, el cambio en la cantidad de elementos en la urna A , $n_A(t)$, solo puede aumentar o disminuir en una unidad. Es decir,

$$n_A(t + 1) - n_A(t) = \pm 1$$

Es conocido para este sistema que, independientemente de la condición inicial¹, el mismo evoluciona a un estado de equilibrio donde las urnas tienen la misma cantidad de elementos cada una, con pequeñas fluctuaciones alrededor del mismo. En otras palabras, luego de un tiempo de transición, la variable $n_A(t)$ fluctúa alrededor del estado de equilibrio $N/2$. Este comportamiento fue utilizado para explicar la interpretación estadística de la segunda ley de la termodinámica, la cual afirma que la entropía de un sistema cerrado

¹La distribución inicial de los elementos, o $n_A(t = 0)$.

solo puede incrementar². La deducción de estos resultados tiene que ver con teoremas que serán presentados en la próxima sección.

En la Fig. 1.1, se muestra el comportamiento mencionado para tres tamaños de sistema N . En estos casos, la condición inicial se corresponde con el estado en el que todos los elementos están ubicados en la urna A ($n_A(0) = N$). Podemos observar en la figura un decaimiento exponencial de la variable $n_A(t)$ hacia el valor medio $N/2$, donde la tasa de decaimiento depende del tamaño de sistema N . Las fluctuaciones de $n_A(t)$ respecto al valor de equilibrio también son sensibles al tamaño de sistema N , ya que su amplitud (relativa a N) se reduce a medida que el tamaño de sistema aumenta.

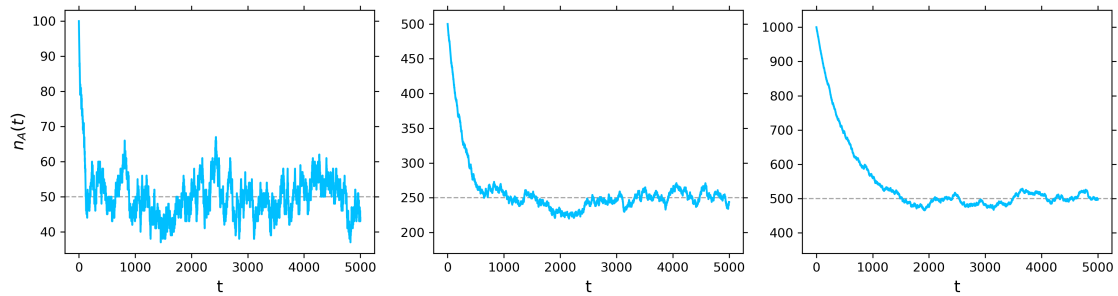


Figura 1.1: Evolución temporal de la cantidad de elementos en la urna A , $n_A(t)$, para 5000 pasos temporales y tamaños de sistema de $N = 100$ (izq.), $N = 500$ (centro), $N = 1000$ (der.).

El modelo de Ehrenfest puede ser resumido como se muestra en el siguiente pseudocódigo. Recordemos que la componente $v[i]$ del vector de ocupación es el estado s_i del elemento con rótulo i , por lo que conociendo $\mathbf{v}(t)$ y tomando la suma de sus componentes al tiempo t , podemos conocer $n_A(t)$.

Algoritmo Paso temporal de Ehrenfest

Input $\mathbf{v}(t)$ $\triangleright \mathbf{v}(t)$ es el vector de ocupación al tiempo t .
Output $\mathbf{v}(t+1)$
Sortear número m entre 1 y N
if $v[m] = 1$ **then**
 $v[m] = 0$ \triangleright cambiar el elemento m de la urna A a la urna B .
else if $v[m] = 0$ **then**
 $v[m] = 1$ \triangleright cambiar el elemento m de la urna B a la urna A .
end if

1.1.3 Modelo del Votante

El modelo del votante ha sido empleado para explicar la evolución de opiniones en una población hacia el consenso. Su versión clásica es otro caso del que hemos presentado: dos urnas A y B con N elementos totales. A diferencia del modelo de Ehrenfest, las reglas de

²Sin importar cómo estén distribuidas inicialmente las partículas en los dos cuerpos, si el sistema es cerrado y la membrana es permeable, las partículas se difundirán ocupando ambos cuerpos.

evolución consideran una interacción entre los elementos, donde se busca modelar cómo los agentes de una sociedad con opiniones distintas, cambian de opinión mediante una comunicación entre ellos. El proceso estocástico definido por las acciones que se aplican en cada paso temporal, puede expresarse de la siguiente manera:

1. Se sortean dos números enteros m_1 y m_2 , ambos entre 1 y N .
2. El elemento cuyo rótulo coincide con m_2 es colocado en la urna del elemento con rótulo m_1 .

En este caso es posible que los elementos que salen sorteados tengan el mismo estado, es decir sean tomados de la misma urna. Cuando eso ocurre, ninguno de los elementos cambia de estado. Esto significa que la cantidad de elementos en la urna A puede aumentar o disminuir una unidad, pero también puede permanecer constante al cabo de un paso temporal. Esto es:

$$n_A(t+1) - n_A(t) = \pm 1, 0$$

El comportamiento a tiempos largos de este modelo es conocido y exhibe, durante un tiempo de transición, una competición entre la población de elementos en la urna A y los de la urna B . Luego evoluciona hacia uno de los estados de total ocupación, es decir, un estado donde todos los elementos estén ubicados en la urna A o todos en la urna B . En términos de los agentes de una sociedad con dos opiniones A y B , significa que la sociedad evoluciona a un consenso global, donde todos los agentes tienen la misma opinión, ya sea la A o la B . Una vez alcanzado ese punto de equilibrio, el sistema permanece en ese estado para siempre. Por este motivo decimos que los dos estados de ocupación total son absorbentes para el sistema.

En la Fig. 1.2 se puede observar a través de la variable $n_A(t)$ el estado del sistema, para diferentes tamaños N . La condición inicial en los tres casos es que las dos urnas tengan la misma cantidad de elementos, es decir, el caso de una sociedad con máxima polarización³. La dinámica del sistema muestra un tiempo de transición donde $n_A(t)$ fluctúa hasta alcanzar el valor de uno de los bordes 0 o N , donde permanece constante. Esto ocurre porque una vez que se alcanza uno de los bordes no hay posibilidad de cambiar elementos de una urna a otra con las reglas definidas, ya que todos los elementos están en la misma.

³En otras palabras, la mitad de los individuos con opinión A y la otra mitad con opinión B

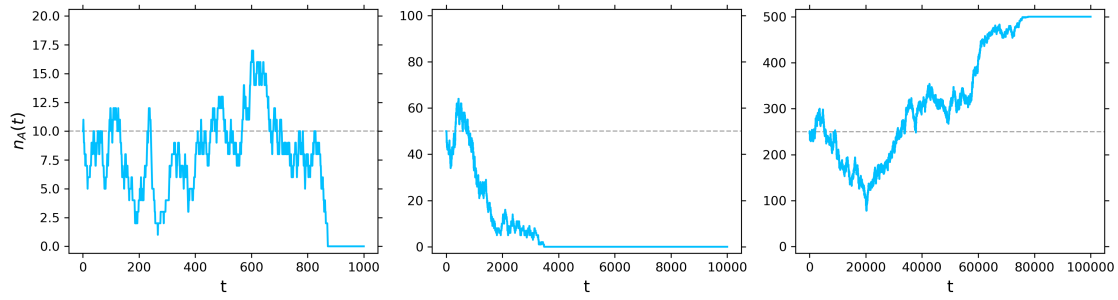


Figura 1.2: Evolución temporal de la cantidad de elementos en la urna A , $n_A(t)$, para distintos tamaños de sistema y cantidad de pasos temporales: $N = 20$ y 1000 pasos (izq.), $N = 100$ y 10000 pasos (centro), $N = 500$ (der.) y 100000 pasos.

El modelo del votante puede ser descripto de manera simplificada por el siguiente pseudocódigo.

Algoritmo Paso temporal modelo del votante

Input $\mathbf{v}(t)$ $\triangleright \mathbf{v}(t)$ es el vector de ocupación al tiempo t .

Output $\mathbf{v}(t+1)$

Sortear número m_1 entre 1 y N

repeat

Sortear número m_2 entre 1 y N

until $m_2 \neq m_1$

$\triangleright m_1$ y m_2 son dos elementos diferentes

if $v[m_2] \neq v[m_1]$ **then**

$v[m_2] = v[m_1]$

\triangleright cambiar a m_2 a la urna de m_1

end if

1.2 Caminatas Aleatorias

En esta sección estudiaremos los sistemas ya expuestos desde el formalismo de caminatas aleatorias. El objetivo es caracterizar a la variable $n_A(t)$ con la posición de un caminante aleatorio que realiza su caminata en una red discreta y finita, para poder valernos de resultados conocidos. Resulta natural relacionar cada paso temporal del estado de un modelo de urnas con el paso de un caminante aleatorio, porque en los mismos el tiempo se toma como una variable discreta.

Una caminata aleatoria discreta a tiempos discretos, es la trayectoria que realiza un caminante dando pasos aleatorios sucesivos en una grilla definida. Cada nodo de la grilla es un valor del espacio muestral disponible para una variable aleatoria Y , que en nuestro caso es la cantidad de elementos en la urna A , $n_A(t)$. En este contexto, la grilla es una red unidimensional con $N + 1$ puntos equiespaciados, del 0 al N .

La posición del caminante luego de n pasos y con condición inicial Y_0 , queda definida por la variable aleatoria Y_n , que dado el tamaño de la red puede tomar valores desde 0 hasta N . Así, la secuencia $\{Y_n\}$ con $n = 0, 1, 2, \dots$ conforma la caminata aleatoria, que

nos da la evolución en el tiempo de la cantidad de elementos en la urna A , es decir, $n_A(t)$ para $t = 0, 1, 2, \dots$. En general, se define a la variable aleatoria Y_n como:

$$Y_0 = n_A(0)$$

$$Y_n = n_A(t_n) = \sum_{i=1}^n X_i = Y_{n-1} + X_n, \quad n = 1, 2, \dots \quad (1.2)$$

donde X_i es una variable aleatoria que denota el paso i -ésimo que da el caminante y genera la caminata aleatoria con $Y_n = X_1 + X_2 + \dots + X_n$.

Todos nuestros casos de interés se restringen a caminatas con saltos a primeros vecinos y con (o sin) posibilidad de permanencia en el mismo sitio al cabo de un paso temporal, por lo que X_i solo puede tomar los valores $\{\pm 1, 0\}$. En otras palabras, si el caminante está en un sitio de la red, solo es no nula la probabilidad de que salte a uno de los sitios contiguos o la de que permanezca en el mismo.

Si el caminante se encuentra en la posición i de la red al cabo de n pasos temporales ($Y_n = i$), decimos que puede saltar a la derecha con probabilidad a_i , a la izquierda con probabilidad c_i o permanecer en la misma posición con probabilidad b_i , donde $a_i + b_i + c_i = 1$, $\forall i$ entre 1 y $N - 1$. En otras palabras, a_i es la probabilidad de que $X_{n+1} = 1$, b_i de que $X_{n+1} = 0$ y c_i de que $X_{n+1} = -1$. Esto se esquematiza en la Fig. 1.3.

La expresión explícita de las probabilidades a_i, b_i y c_i depende del sistema en particular considerado, junto a las condiciones de contorno para $i = 0, N$, pero es importante destacar que en general las mismas no son constantes a lo largo de la red, es decir, dependen de la posición i en la que se encuentra el caminante.

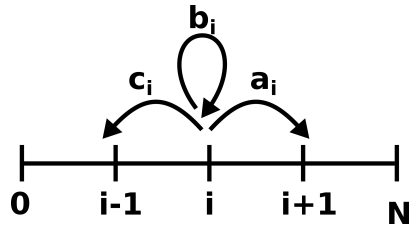


Figura 1.3: Esquema de caminata aleatoria unidimensional, con saltos a primeros vecinos.

1.2.1 Caminatas aleatorias como cadenas de Markov

Una de las cosas que nos podría interesar conocer, es qué probabilidad hay de que la urna A tenga una cantidad específica de elementos al cabo de n pasos temporales. Es decir, que por ejemplo $n_A(t_n) = j$, con j algún valor entre 0 y N . A continuación vemos que algo que en general es una pregunta difícil de responder, para un sistema que es una cadena de Markov se simplifica.

Comenzamos definiendo el vector de distribución de probabilidad del sistema al tiempo t_n como el vector $\mathbf{P}(t_n)$. Este vector $\mathbf{P}(t_n) = (P_0(t_n), P_1(t_n), \dots, P_N(t_n))$, tiene $N + 1$ componentes $P_i(t_n)$, donde cada una de ellas representa la probabilidad de que el

caminante esté en la posición i al cabo de n pasos, o lo que es lo mismo, la *probabilidad de ocupación* $\mathcal{P}(Y_n = i)$ de i luego de n pasos, dado que comenzó en Y_0 .⁴

Conocer las componentes de $\mathbf{P}(t_n)$ puede resultar una tarea muy pesada, ya que habría que tener en cuenta todas las trayectorias que llegan al punto j en n pasos partiendo de Y_0 , sin embargo, en los casos estudiados en este trabajo esta probabilidad se simplifica porque las caminatas aleatorias consideradas cumplen la propiedad de Markov, y por lo tanto se denominan cadenas de Markov. Esta propiedad implica que la probabilidad de que el caminante salte de un sitio a otro es independiente del camino que trazó y depende solo de la posición del caminante. En términos de probabilidades condicionales, la propiedad de Markov se lee:

$$P(Y_n = j | Y_{n-1} = i, Y_{n-2} = k, Y_{n-3} = l, \dots, Y_0 = m) = P(Y_n = j | Y_{n-1} = i) = M_{ij} \quad (1.3)$$

$P(Y_n = j | Y_{n-1} = i) = P(j|i)$ representa la probabilidad condicional de que Y_n tome el valor j dado que Y_{n-1} toma el valor i . La cantidad M_{ij} es la *probabilidad de transición* del estado i al j y puede ordenarse conformando la *matriz de transición* $M \in \mathbb{R}^{(N+1) \times (N+1)}$ cuyas entradas vienen dadas por:

$$M = \begin{bmatrix} M_{00} & M_{01} & \dots & M_{0N} \\ M_{10} & M_{11} & \dots & M_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ M_{N0} & M_{N1} & \dots & M_{NN} \end{bmatrix} \quad (1.4)$$

Como en nuestro caso las caminatas que consideraremos solo tienen saltos a primeros vecinos, $M_{ij} = 0$ para todo $|i - j| > 1$. O sea que en los casos considerados, la matriz será a lo sumo tridiagonal, con

$$\begin{aligned} M_{i,i+1} &= a_i \\ M_{ii} &= b_i \\ M_{i,i-1} &= c_i \end{aligned} \quad (1.5)$$

Si la matriz de transición M es aplicada al vector $\mathbf{P}(t_n)$, se obtiene la distribución de probabilidad del sistema al cabo de un paso temporal,

$$M\mathbf{P}(t_n) = \mathbf{P}(t_n + 1) \quad (1.6)$$

donde $t_{n+1} = t_n + 1$. De esta manera, la distribución de probabilidad del sistema al tiempo t_n , se obtiene de aplicar n veces la matriz M a $\mathbf{P}(0)$.

⁴La condición inicial Y_0 , en términos del vector de distribución de probabilidad, se lee como $P(0) = (0, 0, \dots, 1, \dots, 0)$, lo que significa que al tiempo $t = 0$, el caminante está en una posición de la red con total certeza.

$$M^n \mathbf{P}(0) = \mathbf{P}(t_n) \quad (1.7)$$

Escribir la ecuación (1.6) para una componente $P_i(t)$, nos da la *ecuación maestra* del sistema, la cual contiene toda la información de la evolución del mismo. En nuestro caso, la misma toma la forma:

$$\begin{aligned} P_i(t) &= M_{i-1,i}P_{i-1}(t) + M_{i,i}P_i(t) + M_{i+1,i}P_{i+1}(t) \\ &= P(i|i-1)P_{i-1}(t) + P(i|i)P_i(t) + P(i|i+1)P_{i+1}(t) \\ &= a_{i-1}P_{i-1}(t) + b_iP_i(t) + c_{i+1}P_{i+1}(t) \end{aligned} \quad (1.8)$$

Conocer la ecuación maestra es de utilidad porque a través de una expansión de tamaño del sistema la misma toma la forma de una ecuación diferencial. Al final de la sección se muestran estos resultados para el modelo de Ehrenfest y del votante.

1.2.2 Distribuciones estacionarias de probabilidad

Si bien de esta manera es posible conocer la distribución de probabilidad del sistema al tiempo t_n , dada una distribución inicial; también es interesante conocer su comportamiento a tiempos largos. Se define la *distribución estacionaria* de probabilidad como la solución a la ecuación

$$M\pi = \pi \quad (1.9)$$

junto a la condición de normalización para π . Si estamos interesados en el comportamiento del sistema a tiempos largos, tenemos que considerar la distribución estacionaria de probabilidad. Es posible obtener esta distribución analíticamente para algunos casos particulares, tomando el límite de manera adecuada para $t_n \rightarrow \infty$, y relacionando a $P_i(t_n)$ con un conjunto de polinomios ortogonales. En este trabajo, utilizamos las distribuciones estacionarias conocidas para el modelo de Ehrenfest, el modelo del votante y nuestro propio modelo, para analizar el comportamiento de las simulaciones realizadas.

1.2.3 El modelo de Ehrenfest y del votante como caminatas aleatorias

Para el modelo de Ehrenfest, la probabilidad de saltos a primeros vecinos viene dada por:

$$\begin{aligned} a_i &= P(i|i+1) = \frac{N-i}{N} \\ c_i &= P(i|i-1) = \frac{i}{N} \end{aligned} \quad (1.10)$$

donde a_i es la probabilidad condicional de elegir un elemento de la urna B (dado que tiene $N - i$ elementos) entre N elementos totales, y c_i es la probabilidad condicional de elegir

un elemento de la urna A (dado que tiene i elementos) entre los N totales. En este caso, las condiciones de contorno son reflectantes, porque cuando el caminante llega a uno de los extremos de la red, en el siguiente paso vuelve al nodo anterior inevitablemente. Esto implica que $M_{0,1} = 1 = M_{N,N-1}$, por lo que

En este caso la matriz de transición viene dada por:

$$M = \begin{bmatrix} 0 & 1 & & \dots & 0 \\ c_1 & 0 & a_1 & & \vdots \\ & c_2 & 0 & a_3 & \\ & & \ddots & \ddots & \ddots \\ \vdots & & & c_{N-1} & 0 & a_{N-1} \\ 0 & \dots & & & 1 & 0 \end{bmatrix}, \quad (1.11)$$

ya que $b_i = 0 \quad \forall i$, y por lo tanto, $a_0 = 1 = c_N$. La ecuación maestra, en este caso es:

$$P_i(t) = a_i P_{i-1}(t) + c_i P_{i+1}(t), \quad (1.12)$$

para i entre 1 y $N - 1$, con condiciones de contorno $P_0(t) = P_1(t)$ y $P_N(t) = P_{N-1}(t)$. La distribución estacionaria para Ehrenfest es una distribución binomial que tiene la forma [5]:

$$\pi_j = \binom{N}{j} \left(\frac{1}{2}\right)^N \quad (1.13)$$

Esta expresión se grafica en la Fig. 1.4, donde se puede observar cómo se afina la región de mayor probabilidad al incrementar el valor de N . Que la distribución estacionaria tenga un máximo en este punto significa que el sistema estará con mayor probabilidad en valores cercanos al valor medio $N/2$, luego del tiempo de transición que le lleva al sistema estabilizarse en su comportamiento característico. El hecho de que la distribución estacionaria tenga un decaimiento más abrupto a medida que N aumenta está directamente relacionado a que las fluctuaciones respecto a $N/2$ disminuyan su amplitud relativa al aumentar el tamaño de sistema. Este es el comportamiento que se expuso en la Fig. 1.1. $N/2$ se denomina un *punto de atracción* del sistema, porque el mismo alcanza un equilibrio fluctuando alrededor de ese punto.

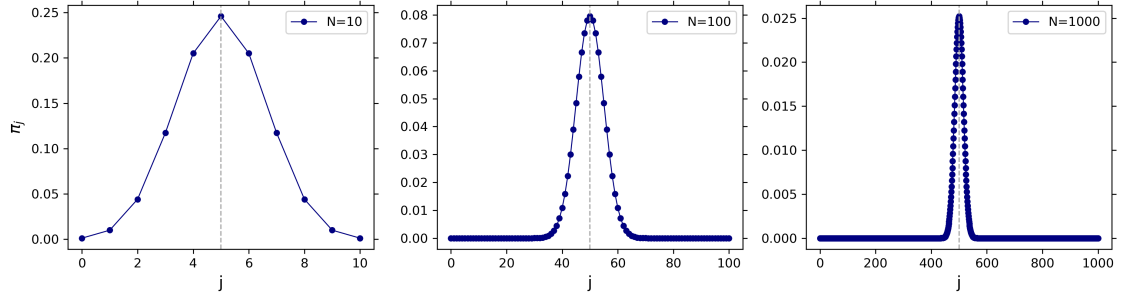


Figura 1.4: Distribución estacionaria para el modelo de Ehrenfest para tres tamaños de sistema: $N = 10, 100, 1000$.

Para el modelo del votante, las probabilidades de saltos a primeros vecinos son:

$$\begin{aligned}
 a_i &= P(i|i+1) = \frac{i}{N} \frac{N-i}{N-1} \\
 b_i &= P(i|i) = 1 - 2 \frac{i}{N} \frac{N-i}{N-1} \\
 c_i &= P(i|i-1) = \frac{i}{N} \frac{N-i}{N-1},
 \end{aligned} \tag{1.14}$$

donde $P(i|i \pm 1)$ es el producto de la probabilidad de elegir un elemento de la urna A de los N elementos totales, i/N , y la probabilidad de elegir un elemento de la urna B de los $N-1$ restantes, $(N-i)/(N-1)$. Mientras que $P(i|i) = 1 - P(i|i+1) - P(i|i-1)$. La matriz de transición resulta:

$$M = \begin{bmatrix} 1 & 0 & & \dots & 0 \\ c_1 & b_1 & a_1 & & \vdots \\ & c_2 & b_2 & a_2 & \\ & & \ddots & \ddots & \ddots \\ \vdots & & & c_{N-1} & b_{N-1} & a_{N-1} \\ 0 & \dots & & 0 & 1 \end{bmatrix} \tag{1.15}$$

En este caso la condición de contorno en los bordes de la red es absorbente, ya que una vez que el caminante llega a uno de los extremos, finaliza su caminata. Por lo tanto $b_0 = b_N = 1$ y $a_0 = c_N = 0$.

La ecuación maestra es la (1.8) para i entre 1 y N , y con condiciones de contorno $P_0 = 1 = P_N$. La distribución estacionaria para el modelo del votante, está dada por una en la que todos los elementos están en la urna A , o bien todos están en la urna B , sin fluctuar. Es decir que el caminante fue absorbido por uno de los bordes y por lo tanto luego de un tiempo de transición, ese será el estado en el que encontraremos al sistema.

1.2.4 Probabilidad de primer pasaje y distribución de retornos

Como mencionamos anteriormente, en este trabajo vamos a estudiar algunas distribuciones de probabilidad asociadas al modelo que detallaremos en el capítulo siguiente. Una de ellas, es la distribución de primer pasaje y la otra se denomina distribución de retornos. Estas distribuciones son de interés ya que describen la probabilidad de que un sistema, que evoluciona en el tiempo como un proceso estocástico, regrese a ciertos estados específicos. Esta información es importante para comprender el comportamiento a largo plazo del sistema y cómo puede fluctuar.

Una manera de definir la probabilidad de primer pasaje $f^{(n)}$ es a través del tiempo, es decir, la cantidad de pasos temporales que el caminante tiene que realizar para volver a un punto por primera vez una vez que salió del mismo. Sea T_0 el tiempo hasta que el caminante aleatorio vuelve al punto de partida por primera vez. Es decir, dado $Y_n = Y_{n-1} + X_n$ $n = 1, 2, \dots$ y considerando $Y_0 = i$,

$$T_i = \min\{n \geq 1 : Y_n = i\} \quad (1.16)$$

Definimos a T_i como el tiempo de primer pasaje al punto i . Dado que, en este ejemplo, el caminante salió del punto i y volvió al mismo punto, solemos referirnos a T_i como el tiempo de *primer* retorno al punto i . De la misma manera se puede definir el tiempo de primer pasaje utilizando puntos arbitrarios de la red. Por ejemplo, si el caminante sale del punto i , el tiempo de primer pasaje por el punto j viene dado por:

$$T_{ij} = \min\{n \geq 0 : Y_0 = i, Y_n = j\}, \quad (1.17)$$

es decir, T_{ij} es el tiempo de primer pasaje al punto j , dado que el caminante comenzó en el punto i .

En este trabajo estamos interesados en las distribuciones de probabilidad de tiempos de primer pasaje o la distribución de primer retorno. En particular, estudiaremos el primer pasaje al punto $Y_n = N/2$, dado que $Y_0 = N/2$. Para obtener la distribución de probabilidad, se calcularán sucesivos tiempos de primeros pasajes $T_{N/2}$ a través de simulaciones numéricas, donde el sistema evolucionará en el tiempo modificando su estado $n_A(t)$. En esta evolución, cada vez que el caminante sale del punto $n_A(t) = N/2$ y vuelve, el tiempo que tarda en realizarlo constituye un punto para el armado de la distribución de frecuencias de $T_{N/2}$. En este sentido, es natural considerar a los valores $T_{N/2}$ como el tamaño temporal de las *fluctuaciones* del sistema respecto al valor $n_A(t) = N/2$. Llamaremos λ al tamaño temporal de dicha fluctuación. Así, la distribución de probabilidades de primer pasaje $f_{N/2}^{(n)}$ es la misma que la distribución de los tamaños temporales de las fluctuaciones λ respecto al valor medio.

En la Fig. 1.5, se esquematiza una evolución temporal del sistema y los valores λ que forman parte de la distribución de primer pasaje. Numéricamente, el valor de λ se obtiene como $\lambda = t' - t''$, donde $n_A(t') = n_A(t'') = N/2$, y no ocurre que $n_A(t) = N/2$ para ningún t entre t' y t'' .

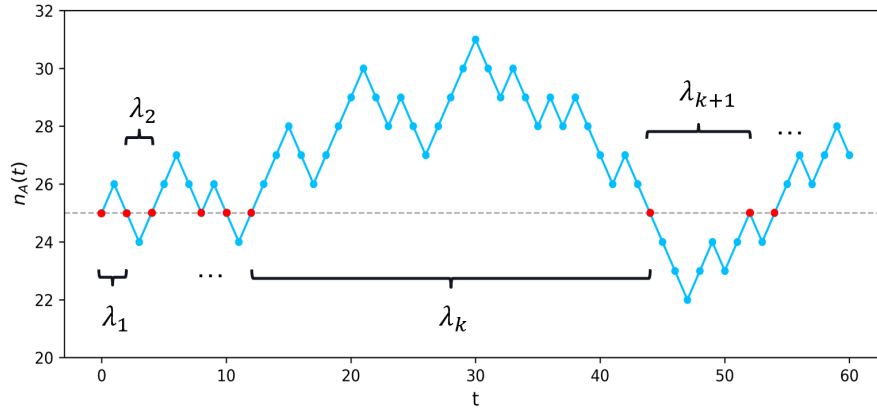


Figura 1.5: Esquema de la trayectoria de la caminata aleatoria Y_n para $n = 0, 1, \dots$. Los puntos rojos son los puntos donde $Y_n = N/2$.

La segunda distribución de interés que mencionamos en la introducción, es la *distribución de retornos*⁵. Ésta se define como la distribución de probabilidad de las diferencias entre tamaños de fluctuaciones consecutivas, esto es:

$$\Delta\lambda(k) = \lambda(k+1) - \lambda(k)$$

donde k son el conjunto de los valores de tiempo t tal que $n_A(k) = N/2$.

Si bien en este trabajo el estudio de estas distribuciones es a partir de un enfoque numérico, es posible obtener la distribución de primer pasaje $f^{(n)}$ de manera analítica para algunos casos resolviendo el problema de polinomios ortogonales relacionado a la Ec. (1.7). Este en general es un problema complejo, dado que hay que conocer dos familias de polinomios ortogonales: los que están relacionados con la caminata aleatoria y los denominados asociados de aquellos. Un ejemplo de este enfoque está dado en el trabajo [6], donde se obtienen expresiones para $f^{(n)}$ en caminatas aleatorias como las de nuestro trabajo, pero con probabilidades de saltos a primeros vecinos constantes en toda la red, es decir, para diferentes conjuntos de valores $a_i = a, b_i = b, c_i = c \ \forall i$. Más adelante veremos que las conclusiones de este trabajo cobran relevancia en el análisis de nuestros resultados.

A continuación se muestra el algoritmo con el pseudocódigo que realiza la evolución temporal del sistema y recopila los valores de λ y $\Delta\lambda$.

⁵Notar que a pesar de su denominación, es diferente a la distribución de prime retorno, la cual hace referencia a la de primeros pasajes o de tamaños λ .

Algoritmo Cálculo de primeros pasajes y retornos

Input $\mathbf{v}(0), T$ $\triangleright T$ es la cantidad de pasos temporales a realizar
Output $\mathbf{v}(T), \lambda, \Delta\lambda$
for t in $[0, T]$ **do**
 Ehrenfest($\mathbf{v}(t)$) o Votante($\mathbf{v}(t)$) \triangleright Se realiza un paso temporal, obtenemos $\mathbf{v}(t+1)$
 if $t+1$ es tal que $n_A(t+1) = N/2$ **then**
 lista _{t} .agregar($t+1$) \triangleright La lista con los tiempos donde el sistema está en $N/2$
 end if
end for
for i in rango(largo(lista _{t}) - 1) **do**
 λ .agregar(lista _{t} [$i+1$] - lista _{t} [i]) \triangleright El arreglo λ contiene los primeros pasajes
end for
for i in rango(largo(λ) - 1) **do**
 $\Delta\lambda$.agregar(λ [$i+1$] - λ [i]) \triangleright El arreglo $\Delta\lambda$ contiene los retornos
end for

1.3 Comportamiento tipo ley de potencia

Para finalizar con el marco teórico, mencionaremos algunas propiedades que emergen cuando un sistema se comporta siguiendo una ley de potencia. Este comportamiento es común y se observa en las distribuciones de probabilidad de varias cantidades de la naturaleza. Como mencionamos en la sección anterior, en este trabajo estudiaremos las distribuciones de primer pasaje y la de retornos para el modelo que será propuesto en el capítulo 2. En particular observaremos si la distribución de primer pasaje sigue un comportamiento del tipo ley de potencia.

En términos matemáticos, una ley de potencia se da cuando la variable dependiente cambia proporcionalmente como una potencia de la variable independiente. Para una distribución de probabilidad, en general, es de interés ver cuando el comportamiento asintótico de la misma tiene esta forma, lo cual se expresa como:

$$P(X \geq x) \sim x^{-\alpha},$$

donde α es una constante positiva que caracteriza a la ley de potencia y x algún valor de la distribución de probabilidad. La constante α es el exponente de la misma, y determina la tasa de decaimiento de la probabilidad a medida que x aumenta. En la Fig. 1.6 se puede observar una distribución con esta conducta.

Una característica distintiva del comportamiento tipo ley de potencia es que la distribución tiene una cola pesada, lo que significa que hay una probabilidad significativa de observar eventos extremadamente raros o de gran magnitud. Otra característica es la ausencia de una escala característica que determine el tamaño típico de los eventos. Esto significa también que la distribución se mantiene similar cuando se la observa a diferentes escalas, es decir, al observarse eventos a diferentes niveles de magnitud. Este tipo de comportamiento, es distintivo respecto al comportamiento "gaussiano" encontrado en la mayoría de los procesos aleatorios y que se encuentra más comúnmente en la naturaleza,

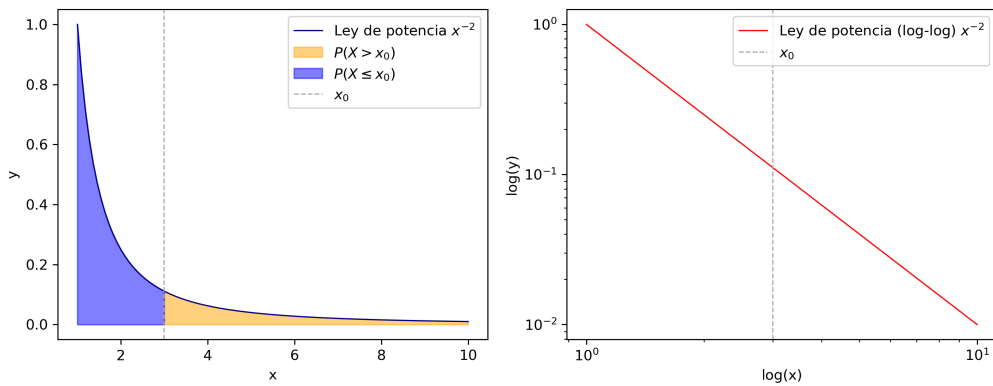


Figura 1.6: Un ejemplo de comportamiento tipo ley de potencia, en escala lineal (izq.) y log-log (der.). La función graficada corresponde a $y = x^{-2}$.

donde las distribuciones presentan colas cortas, que decaen exponencialmente. En el caso gaussiano el tamaño de los eventos queda caracterizado por un rango posible, ya que la probabilidad de que ocurra uno fuera de esa escala es despreciable. Tienen una escala característica definida por la media y la desviación estándar, lo que indica la tendencia central y la dispersión de los datos. En cambio, las distribuciones tipo ley de potencia pueden no tener media y varianza finitas en algunos casos, especialmente cuando la cola es extremadamente pesada.

Aunque en los modelos de urnas que estudiamos en este trabajo la cantidad de elementos N se mantiene finita y por lo tanto siempre observaremos efectos de tamaño finito, veremos en el capítulo 3 que tanto la distribución de primer pasaje como la de retornos presentan un comportamiento tipo ley de potencia para valores grandes de N . Esto nos indica, por ejemplo, que existe una probabilidad baja pero no despreciable de observar fluctuaciones de tamaño λ extremadamente grandes. En la Fig. 1.7, se puede ver cómo se manifiesta una ley de potencia, comparada a un decaimiento exponencial.

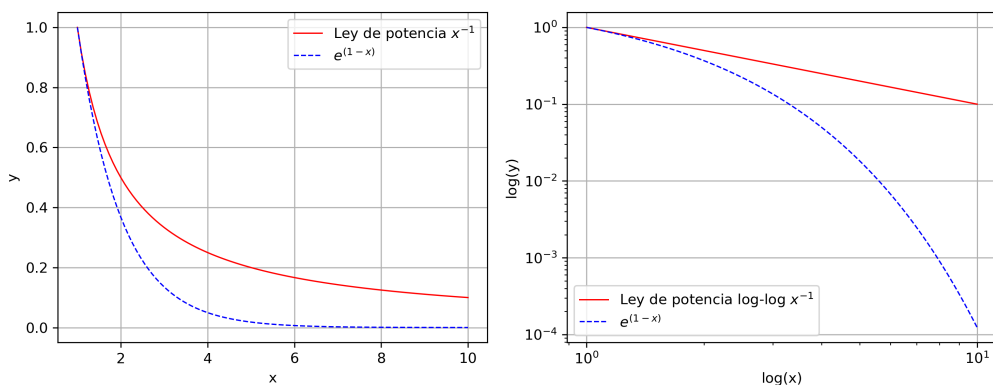


Figura 1.7: Un ejemplo de comportamiento tipo ley de potencia comparado a un decaimiento exponencial, en escala lineal (izq.) y log-log (der.). La relación funcional graficada corresponde a $y = x^{-1}$ y e^{1-x} .

Capítulo 2

Modelo Físico

En este capítulo, presentaremos el modelo físico desarrollado en este trabajo y lo describiremos utilizando los conceptos previamente expuestos en las secciones anteriores. En primer lugar, lo caracterizaremos como un modelo de urnas que combina el modelo de Ehrenfest y el del votante a través de la introducción de un parámetro α . Analizaremos las similitudes que tiene con cada uno de ellos y veremos cómo el valor de dicho parámetro influye en su comportamiento. Luego, estudiaremos simulaciones de evoluciones temporales del mismo, comentando respecto a los comportamientos cualitativos del sistema, utilizando la distribución estacionaria de probabilidad del mismo para justificarlos. Por último presentaremos el algoritmo con el pseudocódigo completo, que inicializa el sistema, lo evoluciona y computa la distribución de tamaños de fluctuaciones y de retornos, las cuales conforman los resultados del próximo capítulo.

2.1 Noisy Voter Model

En el capítulo anterior introdujimos dos conocidos modelos de urna: el de Ehrenfest y el del votante. La propuesta de este trabajo final consiste en una combinación de los mismos. Vimos que ambos se componen de N elementos rotulados del 1 al N distribuidos en dos urnas A y B , pero que se diferencian en las reglas de interacción entre los elementos. Nuestro modelo tiene las mismas componentes, pero su dinámica queda definida combinando las reglas de evolución de los dos modelos ya expuestos. En particular, en cada paso temporal nuestro sistema tiene la posibilidad de realizar un paso temporal de Ehrenfest o uno del votante. Esto se realiza mediante la definición de un parámetro α que representa la probabilidad de que en cada paso temporal ocurra un paso de Ehrenfest. En otras palabras, α es la probabilidad de que a un dado paso temporal la dinámica sea de Ehrenfest, mientras que $1 - \alpha$ es la probabilidad de que a un dado paso temporal se utilice la dinámica del modelo del votante. Así, las acciones que definen el proceso estocástico subyacente son las siguientes:

- Se sortea un número real l del 0 al 1.
- Si $l \leq \alpha$, se realiza un **paso de Ehrenfest**:

1. Se sortea un número entero m entre 1 y N .
 2. El elemento con rótulo m se cambia de urna.
- Si $l > \alpha$, se realiza un **paso del votante**:
 1. Se sortean dos números enteros m_1 y m_2 , ambos entre 1 y N .
 2. El elemento cuyo rótulo coincide con m_2 se cambia a la urna del elemento con rótulo m_1 .

En este caso, la variación en la cantidad de elementos en la urna A luego de un paso temporal viene dado por:

$$n_A(t+1) - n_A(t) = \pm 1, 0,$$

por lo que en cada paso temporal, la cantidad de elementos en la urna A puede mantenerse constante, aumentar o disminuir una unidad. Los valores ± 1 pueden ocurrir tanto si el paso temporal fue de Ehrenfest o del votante, mientras que el valor 0 solo puede ocurrir si el paso temporal fue del votante. Más adelante veremos que esta propiedad cobrará relevancia en el comportamiento del sistema, para valores de α cercanos a 1. En particular, si $\alpha = 1$ recuperamos el modelo de Ehrenfest, y si $\alpha = 0$ volvemos a tener el del votante.

El modelo que presentamos es un caso particular de uno ya propuesto por Alan Kirman en 1993 [1], conocido como *Noisy Voter Model*. Kirman abordó este modelo como una modificación al modelo del votante, en el cual se agregan parámetros que juegan el rol de ruido. Esto lo hizo con el objetivo de evitar los bordes absorbentes del modelo del votante. De esta manera, el sistema ya no evoluciona hacia un estado absorbente de equilibrio, porque la consecuencia inmediata de incluir ruido es que si el sistema alcanza el estado de consenso absoluto, hay una probabilidad no nula de abandonarlo. En nuestro caso, α puede ser considerado un parámetro que juega un papel equivalente al ruido. Kirman propone un modelo más generalizado, con dos parámetros ϵ y δ , aunque analiza en detalle un caso particular para el cual δ toma el valor de 2ϵ . Nuestro modelo se corresponde con el modelo de Kirman tomando $\delta = \epsilon$.

En este trabajo consideramos más intuitivo abordar el *Noisy Voter Model* como una combinación de los dos modelos estudiados en el capítulo anterior. Este enfoque ya fue propuesto en el trabajo de Perachia y Menchón [6], donde se lo plantea en términos de un modelo de opinión. Desde esa perspectiva, es posible afirmar que si luego de un paso temporal ocurre que $n_A(t+1) - n_A(t) = \pm 1$ debido al modelo del votante, el individuo cambió de opinión (cambió de urna) por interacción con otro individuo. Por otro lado, si esto ocurre debido a un paso de Ehrenfest, se puede decir que el cambio de opinión fue espontáneo. En otras palabras, se combinan los cambios de opinión por interacción entre individuos con peso $1 - \alpha$ con los cambios de opinión espontánea con peso α . Este enfoque representa la combinación de un proceso que tiende naturalmente hacia el consenso (todos los elementos en una de las urnas), con otro que tiende naturalmente hacia la polarización (igual cantidad de elementos en ambas urnas).

El algoritmo que incluye el parámetro α en la evolución temporal puede ser representado por el siguiente pseudocódigo.

Algoritmo Paso temporal *Noisy Voter Model*

Input $\mathbf{v}(t)$ $\triangleright \mathbf{v}(t)$ es el vector de ocupación al tiempo t .

Output $\mathbf{v}(t+1)$

Sortear número l entre 0 y 1

if $l \leq \alpha$ **then**

Ehrenfest($\mathbf{v}(t)$) \triangleright con probabilidad α se realiza un paso de Ehrenfest.

else

Votante($\mathbf{v}(t)$) \triangleright con probabilidad $1 - \alpha$ se realiza un paso del votante.

end if

2.1.1 Ejemplo

Supongamos que tenemos $N = 10$ elementos rotulados en las urnas A y B , y que la condición inicial es $n_A(t_0) = 5$, es decir, la mitad de los elementos en A y la mitad en B . Supongamos que fijamos el parámetro $\alpha = 0.5$ y que la distribución de los elementos es como en la Fig. 2.1.

Para obtener $n_A(t_0 + 1)$ se sortea un número l (paso temporal I en la Fig. 2.1). Suponiendo que obtenemos $l = 0.2$ y dado que $\alpha > 0.2$ se efectúa el paso temporal de Ehrenfest: se sortea n_1 . Suponiendo que en el sorteo obtenemos $n_1 = 3$, como el elemento con el rótulo igual a 3 está en la urna A y estamos haciendo un paso de Ehrenfest, se lo cambia a la urna B , y así, al tiempo $t_0 + 1$ tenemos $n_A(t_0 + 1) = 4$, donde $n_A(t_0 + 1) - n_A(t_0) = -1$.

Luego, para obtener $n_A(t_0 + 2)$ se sortea un nuevo número l (paso temporal II en la Fig. 2.1), digamos que ahora $l = 0.8 > 0.5 = \alpha$. En este caso se realiza un paso temporal del votante: es decir se sortea n_1 y n_2 . Suponiendo que n_1 toma el valor 8 y n_2 el valor 2, como el elemento con la etiqueta 8 está en la urna B y el elemento con la etiqueta 2 en la A , se coloca el elemento 8 en la urna A . Así, $n_A(t_0 + 2) = 5$, donde $n_A(t_0 + 2) - n_A(t_0 + 1) = +1$.

Si iteramos un paso más (paso temporal III en la Fig. 2.1), para obtener $n_A(t_0 + 3)$ se sortea $l = 0.7 > 0.5 = \alpha$, por lo que nuevamente se realiza un paso temporal del votante: si en el sorteo obtenemos $n_1 = 1$ y $n_2 = 8$, como el elemento con la etiqueta 1 está en la urna A y el elemento con la etiqueta 8 también, no se cambia de urna ningún elemento y se obtiene $n_A(t_0 + 3) = 5$, donde $n_A(t_0 + 3) - n_A(t_0 + 2) = 0$.

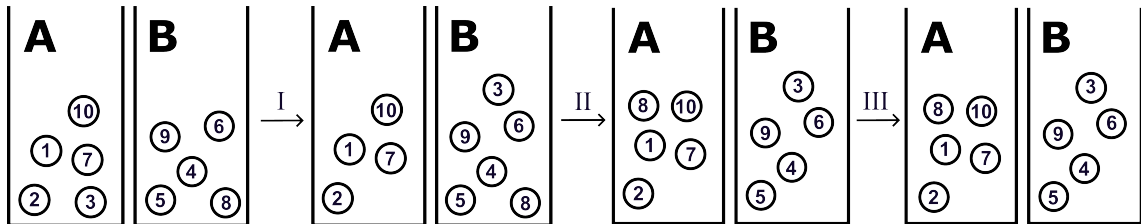


Figura 2.1: Esquema de evolución temporal con tres pasos temporales de las urnas A y B .

La evolución temporal al cabo de M pasos se obtiene haciendo $M - 3$ iteraciones más.

2.1.2 El modelo como una caminata aleatoria

Como se mencionó en la sección 1.2 es posible relacionar la cantidad $n_A(t)$ con una caminata aleatoria unidimensional con saltos a primeros vecinos. El *Noisy Voter Model* es un problema análogo al de una caminata aleatoria a tiempo discreto en una red de $N + 1$ nodos equiespaciados del 0 al N . Las condiciones de contorno en 0 y N dependen del parámetro α . En particular, si $\alpha = 1$ son reflectantes, ya que recuperamos las condiciones de Ehrenfest, y si $\alpha = 0$ son absorbentes, ya que tenemos el caso del votante. En los casos intermedios, los bordes son denominados "pegajosos", y su nivel de adherencia depende del peso que tenga el modelo del votante, es decir son más pegajosos a medida que α toma valores más cercanos a 0.

En nuestro modelo, la probabilidad de saltos a primeros vecinos está dada por:

$$\begin{aligned} a_i &= P(i|i+1) = \alpha \frac{N-i}{N} + (1-\alpha) \frac{i}{N} \frac{N-i}{N-1} \\ b_i &= P(i|i) = (1-\alpha) \left(1 - 2 \frac{i}{N} \frac{N-i}{N-1}\right) \\ c_i &= P(i|i-1) = \alpha \frac{i}{N} + (1-\alpha) \frac{i}{N} \frac{N-i}{N-1}, \end{aligned} \quad (2.1)$$

donde a_i, b_i y c_i constituyen la suma pesada de las probabilidades de saltos a primeros vecinos que antes teníamos para Ehrenfest y el votante. Esto es así porque nuestro modelo es la combinación de dos procesos estocásticos independientes¹. La matriz de transición M tiene la misma forma que la Ec. (1.4), con los coeficientes de la Ec. (2.1).

Este sistema tiene un pseudo punto crítico² en $\alpha = 1/N$ [6]. En este punto, existe un balance entre el modelo del votante y el de Ehrenfest, mientras que si $\alpha > 1/N$ ($\alpha < 1/N$), el modelo de Ehrenfest (votante) tiene más peso que el del votante (Ehrenfest). Cuando $N \rightarrow \infty$, el sistema tiene una transición entre un estado de orden a uno de desorden para $\alpha_c = 0$.

Para analizar el comportamiento en las regiones que divide el pseudo punto crítico, utilizamos la distribución de probabilidad estacionaria del sistema. En este caso la distribución tiene la forma [6]:

$$\pi_j = \binom{N}{j} \frac{\beta(a+j+1, N+a+1-j)}{\beta(a+1, a+1)}, \quad (2.2)$$

donde $\beta(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$ es la función beta y $\Gamma(x)$ es la función gamma, con $a = (\alpha N - 1)/(1 - \alpha)$. En particular, si $\alpha = 1$ recuperamos la ecuación (1.13) del modelo de Ehrenfest, la cual se obtiene tomando correctamente el límite para $\alpha \rightarrow 1$ ($a \rightarrow \infty$).

¹sortear el número l es independiente del sorteo de m o m_1 y m_2

²Es decir que el sistema para $N \rightarrow \infty$ tiene un punto crítico.

Para $\alpha = 1/N$, se obtiene que $\pi_j = 1/(N + 1)$, por lo que la distribución estacionaria es uniforme. Si $\alpha = 0$ y $n_A(0) = 0$, $\pi_j = 0$ para j entre 1 y $N - 1$, y $\pi_0 = 1/2 = \pi_N$, que es el caso del modelo del votante.

En la Fig. 2.2 se grafica esta distribución para distintos valores de α . Observamos que para valores de α cercanos a 1, la distribución se parece a la de Ehrenfest con un único máximo en $N/2$, y a medida que disminuye su valor (pero todavía $\alpha > 1/N$), el valor máximo de la distribución es cada vez menor. En $\alpha = 1/N = 0.1$ podemos ver la distribución uniforme, y para $\alpha < 1/N$ se observa que hay dos máximos en los valores extremos $j = 0$ y $j = N$.

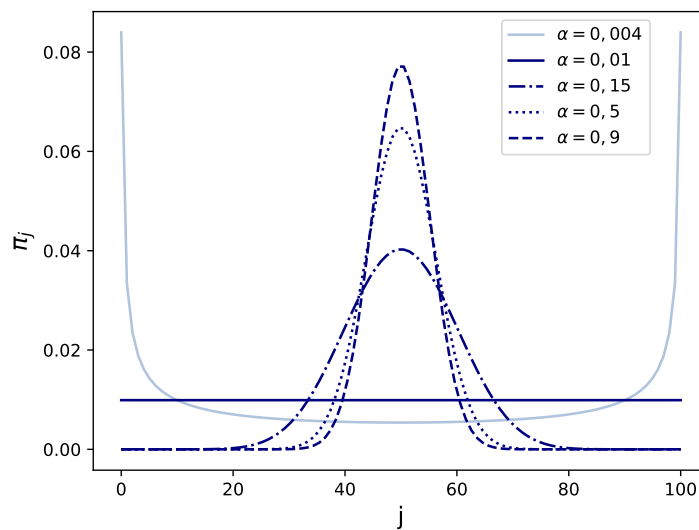


Figura 2.2: Distribuciones estacionarias para un sistema de $N = 100$ elementos.

Esta información permite caracterizar tres regiones distintas, cada una con patrones de comportamiento diferentes. En primer lugar, tenemos el pseudo punto crítico en $\alpha = 1/N$, donde la distribución es uniforme y el sistema no tiene mayor probabilidad de encontrarse en un punto en particular. En términos de $n_A(t)$, significa que no hay manera de predecir su valor luego de un tiempo característico. Como mencionamos hace un momento, este pseudo punto crítico divide una fase bi-modal de una unimodal. La primera fase que denominaremos *región supracrítica*, se corresponde con $\alpha > 1/N$. En esta región la distribución estacionaria tiene su máximo en $N/2$, por lo que el mismo es un evidente atractor del sistema, haciendo que fluctúe alrededor del mismo. Notamos que la atracción más fuerte hacia $N/2$ se da para el caso de Ehrenfest ($\alpha = 1$), y que va disminuyendo a medida que se acerca al valor del pseudo punto crítico. En términos de los modelos de urna, esta región se identifica con que las urnas están ambas parcialmente llenas. En un modelo de opinión, esto se traduce en la presencia de dos poblaciones macroscópicas coexistiendo (opiniones A y B). La segunda fase la denominaremos *región subcrítica*, y es donde $\alpha < 1/N$. En este caso la distribución de probabilidad tiene dos máximos ubicados en los extremos, lo cual genera que los bordes sean "pegajosos". A

medida que $\alpha \rightarrow 0$, la intensidad de la adherencia aumenta, ya que nos acercamos al caso del modelo del votante. En términos de los modelos de urna, casi todos los elementos están en, por ejemplo, la urna A por un tiempo determinado y luego cambian a estar casi todos en la urna B . Desde la perspectiva de un modelo de opinión, esto implica que existe solo una población macroscópica con una opinión definida, que luego de un tiempo determinado puede invertirse.

En la Fig. 2.3 se observan diferentes evoluciones para distintos valores de α . Podemos observar en la evolución superior el caso de Ehrenfest, que ya habíamos observado en la Fig. 1.1. Vemos que el valor de $n_A(t)$ nunca llega a los bordes y que el sistema fluctúa alrededor del valor $N/2$. La segunda evolución, contando de arriba a abajo, es para un valor de α de la región supracrítica. Observamos el mismo comportamiento que en Ehrenfest pero con fluctuaciones un poco más grandes. En la tercer evolución, que corresponde al valor de α correspondiente al pseudo punto crítico, se puede ver que el sistema ya no tiende a fluctuar muy cerca de $N/2$. La cuarta evolución corresponde a la región subcrítica, y allí se ven con claridad los bordes "pegajosos". La última evolución corresponde al modelo del votante, y vemos como al llegar a uno de los bordes $n_A(t)$ permanece constante en ese valor.

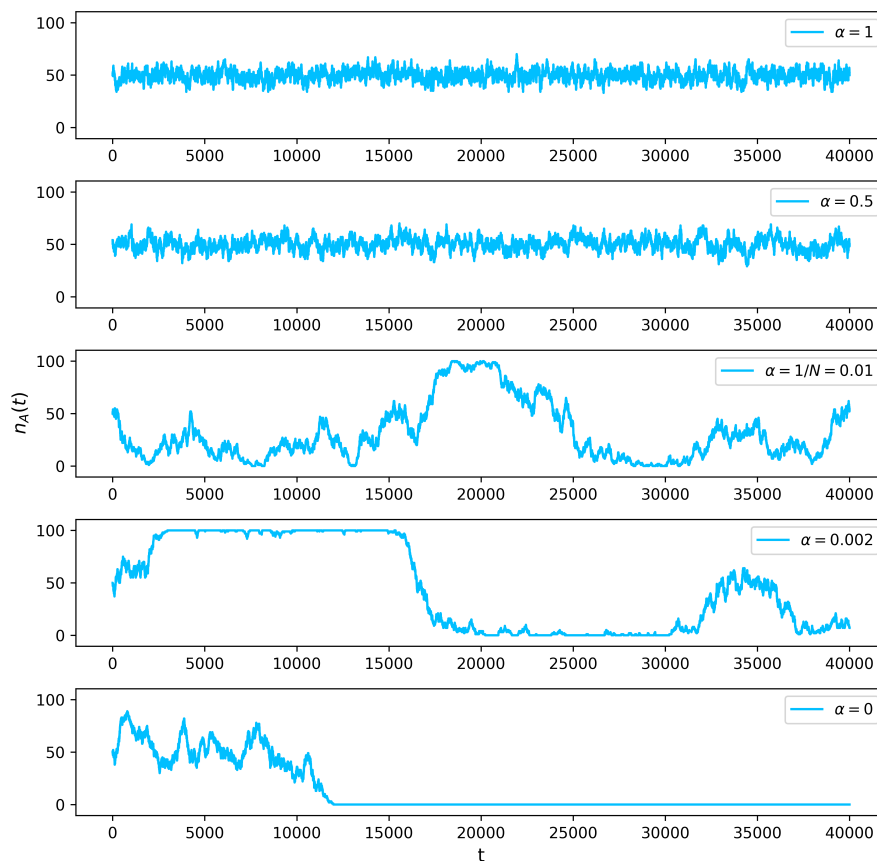


Figura 2.3: Evolución temporal de $n_A(t)$ para diferentes valores de α .

Si bien a partir del valor $\alpha \leq N$ la distribución estacionaria ya no tiene un máximo

en $N/2$, este punto sigue siendo importante, ya que si tomamos cualquier otro punto i de la red, el tiempo $T_{N/2,i}$ que el sistema tarda en ir de $N/2$ a ese punto es menor que el tiempo $T_{i,N/2}$ de ir de i a $N/2$. Es decir, $N/2$ sigue siendo un punto de atracción [6]. Por este motivo, para todas las simulaciones que realizamos en el capítulo 3, calculamos los tiempos de primer retorno hacia $N/2$.

Para finalizar la sección, a continuación se presenta el pseudocódigo completo, que realiza una cantidad n de ciclos de T pasos temporales, y donde además se calculan los vectores λ y $\Delta\lambda$ que guardan los valores de la distribución de primer pasaje y de la de retorno.

Algoritmo Evolución temporal

Input $\alpha, \mathbf{v}(0), T, l$.

Output $\mathbf{v}(n \times T), \text{lista}_\lambda, \text{lista}_{\Delta\lambda}$

$\text{lista}_\lambda = [0]$ \triangleright condición inicial para $t = 0$

$n_A(t) = [\sum_{i=1}^N (v[i])]$ \triangleright condición inicial para $t = 0$

$\text{lista}_\lambda = []$

$\text{lista}_{\Delta\lambda} = []$ \triangleright las listas con los tamaños de fluctuaciones λ y retornos están vacías.

while $\text{largo}(\text{lista}_\lambda) < l + 1$ **do**

for t in $[0, T]$ **do** \triangleright Se realiza un paso temporal

 Sortear número l entre 0 y 1

if $l \leq \alpha$ **then** \triangleright con probabilidad α se realiza un paso de Ehrenfest.

 Sortear número m entre 1 y N

if $v[m] = 1$ **then**

$v[m] = 0$ \triangleright cambiar el elemento m de la urna A a la urna B.

else if $v[m] = 0$ **then**

$v[m] = 1$ \triangleright cambiar el elemento m de la urna B a la urna A.

end if

else \triangleright con probabilidad $1 - \alpha$ se realiza un paso del votante.

 Sortear número m_1 entre 1 y N

repeat

 Sortear número m_2 entre 1 y N

until $m_2 \neq m_1$ \triangleright m_1 y m_2 son dos elementos diferentes

if $v[m_2] \neq v[m_1]$ **then**

$v[m_2] = v[m_1]$ \triangleright cambiar a m_2 a la urna de m_1

end if

end if

if $t + 1$ es tal que $n_A(t + 1) = N/2$ **then**

$\text{lista}_t.\text{agregar}(t + 1)$ \triangleright La lista con los tiempos donde el sistema está en $N/2$

end if

end for

for i in $\text{rango}(\text{largo}(\text{lista}_t) - 1)$ **do**

$\lambda.\text{agregar}(\text{lista}_t[i + 1] - \text{lista}_t[i])$ \triangleright El arreglo λ contiene los primeros pasajes

end for

for i in $\text{rango}(\text{largo}(\lambda) - 1)$ **do**

$\Delta\lambda.\text{agregar}(\lambda[i + 1] - \lambda[i])$ \triangleright El arreglo $\Delta\lambda$ contiene los retornos

end for

end while

Capítulo 3

Resultados y Discusión

En este capítulo presentaremos los resultados obtenidos del estudio computacional del *Noisy Voter Model* que describimos en el capítulo 2. Como mencionamos anteriormente, el sistema presenta tres regímenes característicos en función del parámetro α : la región supracrítica ($\alpha > 1/N$), el pseudo punto crítico ($\alpha = 1/N$) y la región subcrítica ($\alpha < 1/N$). Nuestro objetivo principal es investigar a través de las distribuciones de primeros pasajes y de retornos el comportamiento del sistema al aumentar su tamaño N . En ese sentido, dividiremos los resultados de esta manera: los obtenidos en la región supracrítica ($\alpha > 1/N$) y los obtenidos en el pseudo punto crítico ($\alpha = 1/N$). En el primer caso, primero replicaremos los resultados obtenidos en un trabajo previo [5], en el que se exploró el comportamiento para varios α de la región. Luego nos centraremos en nuestros resultados principales, que son las distribuciones para diferentes valores de N para $\alpha = 0.9$ y $\alpha = 1/N$. Para estos dos casos, replicaremos la técnica utilizada por Bakar y Tirnakli [7] en el modelo de Ehrenfest, donde se utiliza una relación entre el exponente τ de la ley de potencia de la distribución de primer pasaje y el parámetro q de la q -Gaussiana asociada a la distribución de retornos. Luego realizaremos el ajuste de la función propuesta en [8] sobre las distribuciones de retorno obtenidas, con el parámetro q calculado *a priori*. No presentaremos resultados de la región subcrítica por el alto costo computacional que significa simular un buen número de fluctuaciones respecto a $N/2$ en este caso.

3.1 Resultados en la región supracrítica ($\alpha > 1/N$)

En esta región vimos que la distribución estacionaria de probabilidad tiene un máximo en $N/2$ y por lo tanto ese punto es un atractor del sistema, que fluctuará al rededor del mismo. Como estamos interesados en los primeros pasajes a $N/2$ (o lo que es lo mismo, los tamaños temporales λ de fluctuaciones respecto a $N/2$), comenzamos todas las simulaciones con condición inicial en ese punto, es decir, con $n_A(0) = N/2$ ¹. Como se mencionó en la sección 1.2.4, para computar el tamaño λ de una fluctuación, se registra la cantidad de pasos temporales entre que el sistema sale de la posición de equilibrio

¹Recordemos que esto significa que las dos urnas comienzan con la misma cantidad de elementos cada una.

$N/2$ y vuelve. Luego, para obtener los retornos, calculamos la diferencia entre valores λ consecutivos: $\Delta\lambda = \lambda(k+1) - \lambda(k)$, donde el índice k indica el orden en el que fueron obtenidas las fluctuaciones de tamaño λ .

Lo primero que notamos al realizar las evoluciones fue que la cantidad de puntos λ y $\Delta\lambda$ obtenidos con una cantidad específica de pasos temporales es diferente para distintos valores de α y N . Esto era esperado porque vimos que la probabilidad de que el sistema salga y se acerque a $N/2$ cambia en cada caso. Por este motivo, para obtener una cantidad de puntos de λ y $\Delta\lambda$ que permita comparar todos los casos, fijamos como orden de magnitud aceptable una cantidad de entre 10^6 (simulaciones donde comparamos diferentes α) y 10^7 (simulaciones donde comparamos para distintos N) puntos. Para satisfacer ese requisito, se utilizó el Algoritmo *Evolución temporal* en todas las simulaciones. Este algoritmo aplica ciclos de $T = 3 \times 10^5$ pasos temporales cada uno, hasta que se satisface la condición requerida.

En la Fig. 3.2 se observan las distribuciones de tamaños λ y $\Delta\lambda$ obtenidas, en escala logarítmica, para $N = 100$ y $\alpha = 1$. Este es el caso donde nuestro sistema se reduce al modelo de Ehrenfest. Podemos observar, al igual que en el trabajo de Bakar y Tirnakli [7], que la distribución de tamaños λ presenta una ley de potencia que se extiende hasta $\lambda \sim 10^2$, y que para valores más grandes de λ presenta un decaimiento exponencial debido al efecto de tamaño finito del sistema. Sobre esta distribución realizamos un ajuste lineal y obtuvimos un valor de $\tau = 1.5$, donde definimos τ como el exponente de la ley de potencia.

En la misma figura podemos ver también la distribución de retornos $\Delta\lambda$. La misma no puede ser ajustada por una gaussiana, pero tampoco presenta la cola ancha de una q-Gaussiana, debido al decaimiento por tamaño finito. El valor máximo de la distribución de retornos se da para $\Delta\lambda = 0$, por la gran contribución debido a pequeñas fluctuaciones. Estos resultados coinciden con los reportados en el trabajo mencionado.

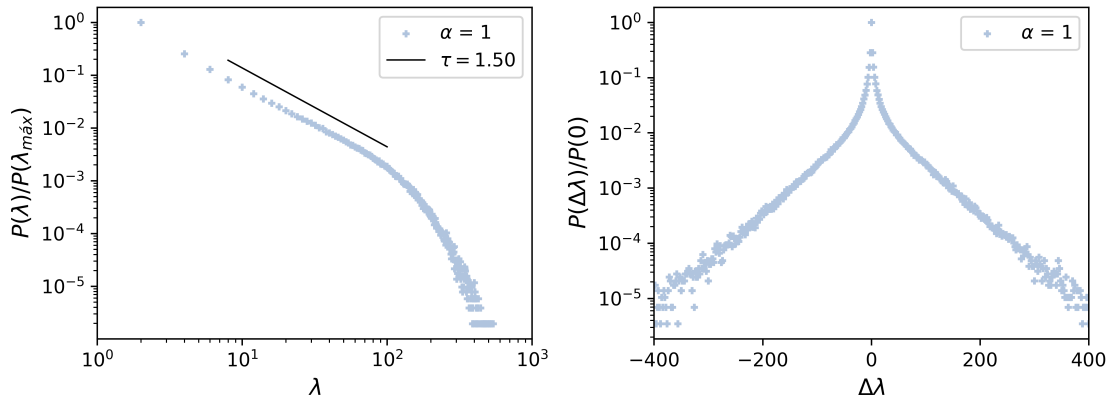


Figura 3.1: Distribución de tamaño de fluctuaciones λ , para el caso de Ehrenfest ($\alpha = 1$) y $N = 100$, normalizada al valor máximo, junto a ajuste lineal para el cálculo del exponente τ (Izq.). Distribución de retornos para el mismo caso, normalizada al valor de $\Delta\lambda = 0$. (Der.) Cada distribución tiene aproximadamente 10^6 puntos.

Luego, en la Fig. 3.2 tenemos las distribuciones de tamaños λ para otros valores de

α , también para $N = 10^2$. Un aspecto que resalta inmediatamente, a diferencia del caso de Ehrenfest, es la división en dos ramas cuando α se aproxima a 1. Esta distinción ocurre porque, si consideramos una fluctuación, el caminante puede tardar una cantidad par o impar de pasos en volver a $N/2$ una vez que salió del mismo punto, ya que hay una probabilidad no nula de permanecer en el mismo sitio al cabo de un paso temporal. En el modelo de Ehrenfest, vimos que el cambio $n_A(t+1) - n_A(t)$ solo puede ser ± 1 , lo que se traduce en que los λ solo pueden ser pares. Al incluir la posibilidad de que $n_A(t+1) - n_A(t) = 0$ en nuestro modelo, podemos tener λ impares. Dado que el modelo del votante tiene muy poca probabilidad de ocurrencia si α es cercano a 1, la probabilidad de tener λ impares es menor. Esto nos da como resultado dos ramas en la distribución: la de mayor probabilidad corresponde a los λ pares y la de menor probabilidad a los λ impares.

Además, en la figura podemos notar que la diferencia entre la rama par e impar es más pronunciada para valores pequeños de λ . Esto se debe a que fluctuaciones pequeñas el sistema sale y regresa al equilibrio en una menor cantidad de pasos, lo que reduce las oportunidades de que ocurra un paso del modelo del votante. Este comportamiento es similar al observado en otros sistemas, como por ejemplo en el estudiado en [6], donde se considera una caminata aleatoria con probabilidades de saltos a primeros vecinos constantes y donde la probabilidad de permanencia en el sitio disminuye considerablemente.

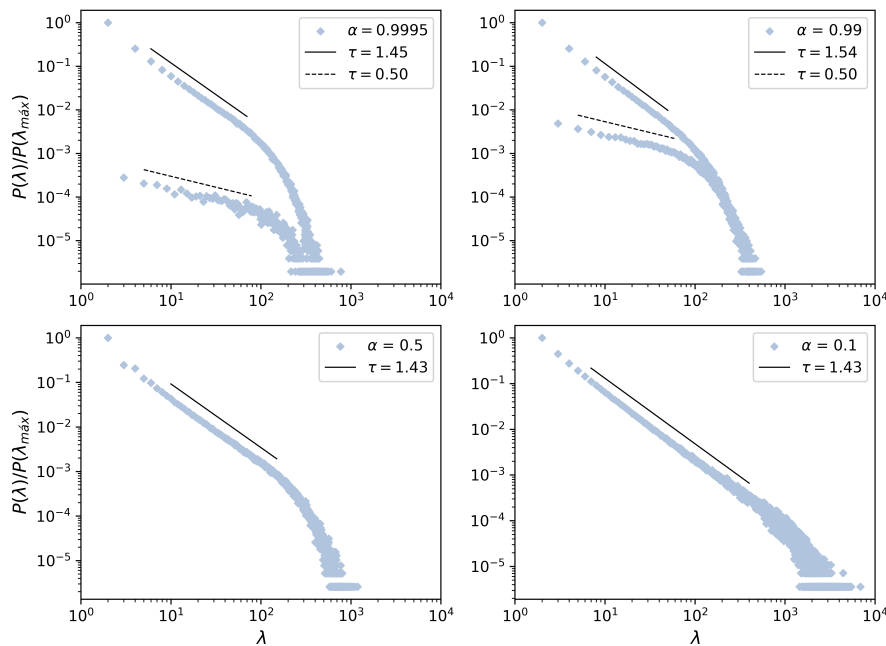


Figura 3.2: Distribuciones de tamaño de fluctuaciones λ , para distintos valores de α , para $N = 10^2$ y ajustes lineales para el cálculo del exponente τ . Están normalizadas respecto al valor máximo de cada una. Cada distribución tiene aproximadamente 10^6 puntos.

Respecto al comportamiento tipo ley de potencia, en la Fig. 3.2 podemos observar su presencia en todos los casos, junto al decaimiento exponencial por efecto de tamaño finito.

Notamos que la ley de potencia se extiende para valores más grandes de λ a medida que α disminuye y se acerca al valor del pseudo punto crítico. Además, observamos que a medida que α se acerca a 1, la rama impar disminuye gradualmente en comparación con la rama par, manteniendo su pendiente.

Los ajustes lineales realizados sobre las distribuciones dan un valor $\tau = 1.4$, para los casos donde no hay separación en dos ramas. Por otro lado, cuando existe la división, se obtiene $\tau_{par} = 1.5$ y $\tau_{impar} = 0.5$. En el trabajo previo de Perachia [5], se obtuvo el mismo comportamiento, con los mismos exponentes.

Posteriormente, en la Fig. 3.3 se presenta el resultado obtenido para las distribuciones de retornos $\Delta\lambda$, correspondientes a los mismos casos de la Fig. 3.3. Podemos observar que para los α que había división en λ pares e impares, las distribuciones de retorno también se separan en dos ramas. La rama superior corresponde a los retornos pares, mientras que la rama inferior los impares.

Es importante tener en cuenta que, dado que los retornos $\Delta\lambda$ son la diferencia de dos λ consecutivos, un retorno par se forma con la diferencia entre dos λ pares o dos λ impares, mientras que un $\Delta\lambda$ impar solo se forma con la diferencia de un λ par y otro impar. En este caso la rama par de retornos es mayor que la impar porque siempre es más probable tener un λ par seguido de otro par que de uno impar, y claramente es menos probable tener dos λ impares consecutivos.

Por último, notamos que a medida que α se acerca al valor del punto crítico, la cola de la distribución se vuelve cada vez más ancha.

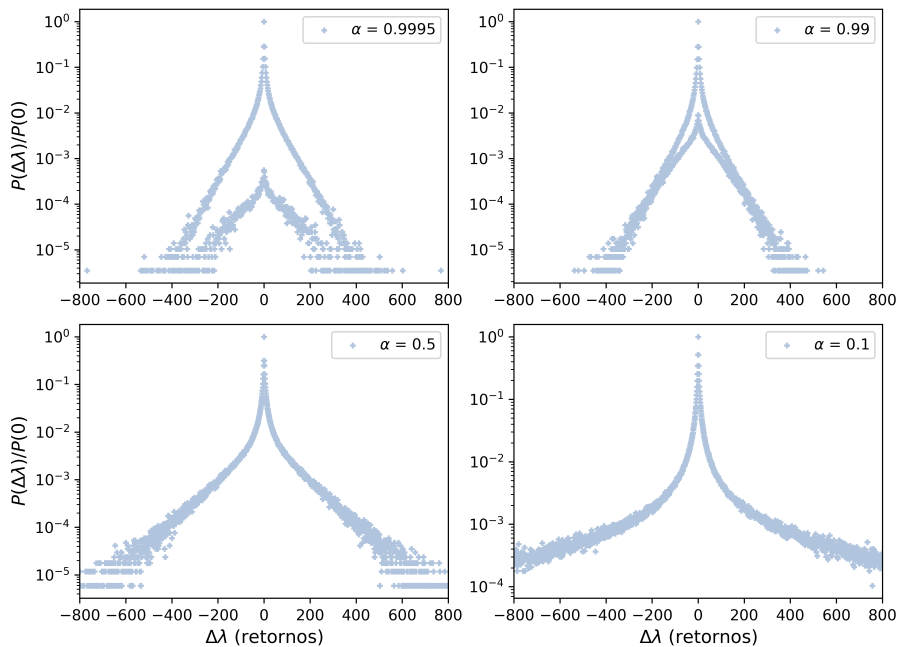


Figura 3.3: Distribuciones de retornos $\Delta\lambda$, para distintos valores de α , para $N = 10^2$. Están normalizadas respecto al valor máximo de cada una ($P(0)$). Cada distribución tiene aproximadamente 10^6 puntos.

Resultados para $\alpha = 0.9$

Como anticipamos en la introducción de este capítulo, luego de caracterizar el comportamiento de α en toda la región supracrítica nos enfocamos en aumentar el tamaño de sistema para el valor de $\alpha = 0.9$.

En la Fig. 3.4, se presenta la distribución de tamaños de fluctuaciones λ para los valores de $N = 10^2, 10^3, 5 \times 10^3$ y 10^4 . En todos los casos se observa una ley de potencia con la misma pendiente $\tau = 1.45 \pm 0.05$. Para este valor de α la división en dos ramas se da para λ pequeños. Como la separación no varía al cambiar el N , podemos afirmar que no se debe a un efecto de tamaño finito.

Notamos que a medida que aumentamos el sistema, el decaimiento exponencial se pospone hacia valores más grandes de λ . Esto se debe a que los bordes del sistema se encuentran relativamente más cerca del punto de equilibrio $N/2$ si el sistema es pequeño. Si los bordes están más lejos, hay posibilidades de que el caminante aleatorio tarde más en volver a $N/2$ una vez que salió, lo que aumenta la probabilidad de obtener fluctuaciones más grandes. Con estos resultados, podemos suponer que si $N \rightarrow \infty$, la ley de potencia se extenderá para todo λ .

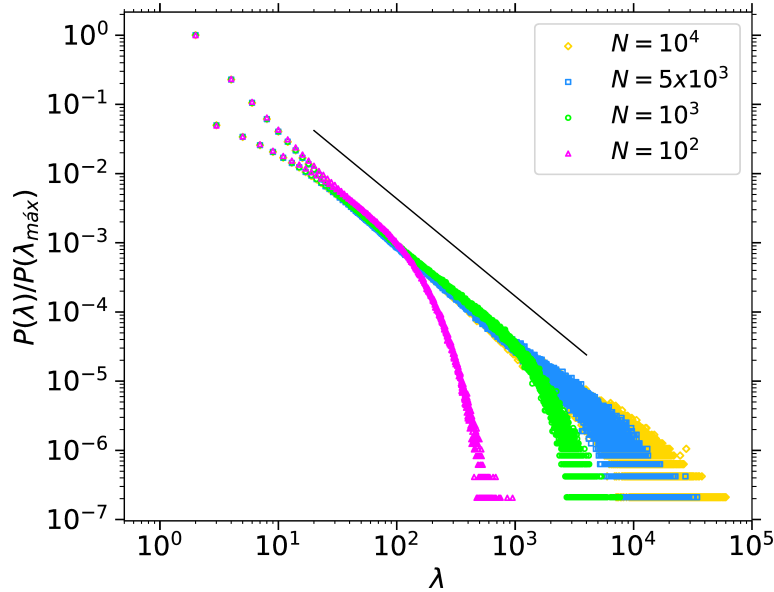


Figura 3.4: Distribuciones de tamaño de fluctuaciones λ en escala logarítmica para $N = 10^2, 10^3, 5 \times 10^3, 10^4$, normalizadas respecto al valor máximo de cada una. Cada distribución tiene aproximadamente 10^7 puntos. Ajuste lineal de pendiente $\tau = 1.45 \pm 0.05$.

En la Fig. 3.5, presentamos la distribución de retornos para los mismos valores de N . Tanto en el gráfico **a** como en el **b**, se puede observar que las distribuciones tienden a tener colas cada vez más anchas a medida que aumenta el tamaño de sistema. Este resultado confirma nuestra expectativa de que, a medida que disminuye el efecto de tamaño finito en la distribución de primer pasaje, su distribución de retornos asociada se asemeja cada vez más a una q-Gaussiana.

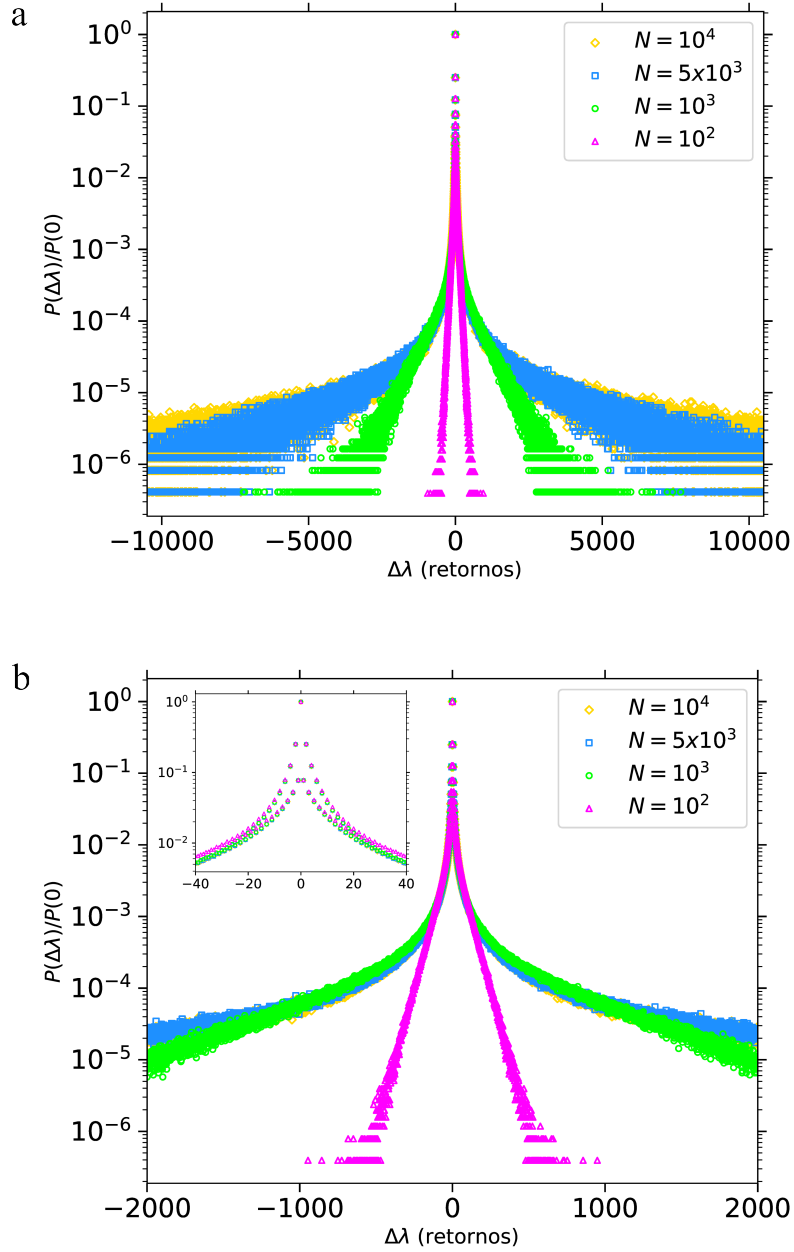


Figura 3.5: Distribuciones de retorno en escala logarítmica para $N = 10^2, 10^3, 5 \times 10^3, 10^4$, normalizadas respecto al valor máximo de cada una. Cada distribución tiene aproximadamente 10^7 puntos.

Una q-Gaussiana se puede definir en términos de los parámetros q y β de la siguiente manera:

$$P(x) = P(0) \left[1 - (1 - q)\beta x^2 \right]^{\frac{1}{1-q}}, \quad (3.1)$$

donde $x = \Delta\lambda - \langle \Delta\lambda \rangle$, y $\langle \Delta\lambda \rangle$ es el valor medio de los retornos, que en nuestro caso resulta ser igual a 0. El parámetro β caracteriza el ancho de la distribución, mientras que q es el

índice de mecánica estadística no aditiva. Si $q \rightarrow 1$, se recupera la distribución gaussiana, mientras que si $q \neq 1$, el comportamiento difiere del gaussiano.

En el trabajo de Bakar y Tirnakli [7], [9], que se centró en el análisis del modelo de Ehrenfest, establecen una relación entre la pendiente τ de la distribución de primeros pasajes y el valor q de la q-Gaussiana que ajusta a la distribución de retornos. Esta relación es importante porque permite obtener uno de los parámetros de la q-Gaussiana *a priori*. Por otro lado, en la investigación de Celikoglu y Tirnakli [8], que abordó un problema similar denominado *Coherent Noise Model*, se obtuvo una fórmula exacta para esta relación:

$$q = \frac{(\tau + 2)}{\tau}, \quad (3.2)$$

la cual emplearemos en este trabajo para ahorrar un parámetro de ajuste para las distribuciones de retorno.

En nuestro modelo, observamos que las distribuciones de retornos no tienen la forma de una gaussiana ni de una q-Gaussiana, y notamos que su forma depende del tamaño N del sistema. Por lo tanto, utilizaremos como función de ajuste la propuesta en los trabajos [7] y [8]. Esta función combina una gaussiana con una q-Gaussiana, e incluye dos parámetros a_q y a_1 , además del valor de q que obtenemos a priori a partir de la relación (3.2):

$$f(x) = \left[1 - \frac{a_q}{a_1} + \frac{a_q}{a_1} e^{(q-1)a_1 x^2}\right]^{1/(1-q)}, \quad (3.3)$$

donde una vez más $x = \Delta\lambda - \langle \Delta\lambda \rangle$, y $\langle \Delta\lambda \rangle = 0$.

Esta ecuación es solución a la ecuación diferencial:

$$\begin{aligned} \frac{dy}{d(x^2)} &= -a_r y^r - (a_q - a_r) y^q \\ [a_q \geq a_r \geq 0, q > r, y(0) = 1], \end{aligned} \quad (3.4)$$

con $r = 1$ y $q > 1$. Si $a_1 = 0$ la solución es una q-Gaussiana, mientras que si $a_q = a_1$ (y por lo tanto $q = 1$) la solución es gaussiana. En la región intermedia, para $a_q > a_1 > 0$ y $q > 1$, se da una combinación de ambas.

En las Fig. 3.6 y 3.7 se presentan los ajustes lineales realizados sobre las distribuciones de tamaños λ y los ajustes de la función (3.3) sobre las distribuciones de retorno. En todos los casos, se calculó el parámetro q utilizando la Ec. (3.2). Para los ajustes lineales, consideramos para cada N una región relevante del eje λ : se dejaron afuera las fluctuaciones más pequeñas y las más grandes, evitando la separación en dos ramas y el decaimiento exponencial. En cuanto a los ajustes sobre las distribuciones de retornos, se consideraron solamente los puntos de la rama de retornos par, ya que el τ obtenido se relaciona a los λ pares. Además para realizar los ajustes, tomamos el logaritmo de los

datos y ajustamos el logaritmo de la Ec. (3.3), ya que de otra manera los puntos en las colas no pesaban lo suficiente.

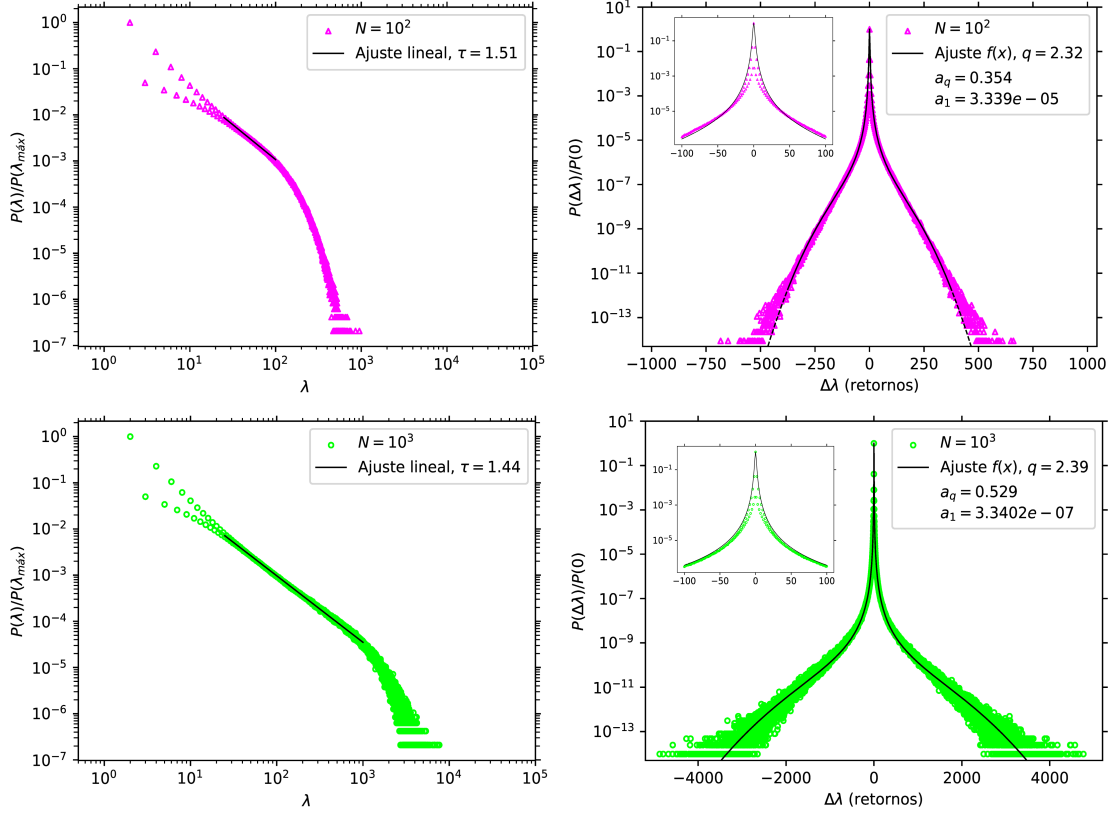


Figura 3.6: Ajuste lineal sobre la distribución de tamaños de fluctuaciones λ en el intervalo $[25, 100]$ para $N = 10^2$ y $[25, 1000]$ para $N = 10^3$ (izq.). Ajuste de la función $f(x)$ (Ec. (3.3)) sobre la distribución de retornos con valor q calculado a priori, en los intervalos $[-370, 370]$ para $N = 10^2$ y $[-3800, 3800]$ para $N = 10^3$ (der.).

Podemos verificar a través de los ajustes, que la componente gaussiana a_1 disminuye al aumentar el tamaño de sistema, por lo que la distribución queda caracterizada cada vez más por una q -Gaussiana a medida que N crece. Notemos que si $(q - 1)a_1x^2 \ll 1$, la Ec. (3.3) tiende a una q -Gaussiana, $f(x) \sim [1 - (1 - q)a_qx^2]^{1/(1-q)}$. Con esta notación, el parámetro a_q cumple el rol del parámetro β de la Ec. (3.1). Por otro lado, si $(q - 1)a_1x^2 \gg 1$ el término con la exponencial prevalece, por lo que asintóticamente va como una gaussiana: $f(x) \sim (a_q/a_1)^{1/(1-q)} \exp(-a_1x^2)$.

Con base en estos resultados, podemos concluir que a medida que la ley de potencia en la distribución de primeros pasajes se extiende, la q -Gaussiana apropiada domina cada vez más en la distribución de retornos.

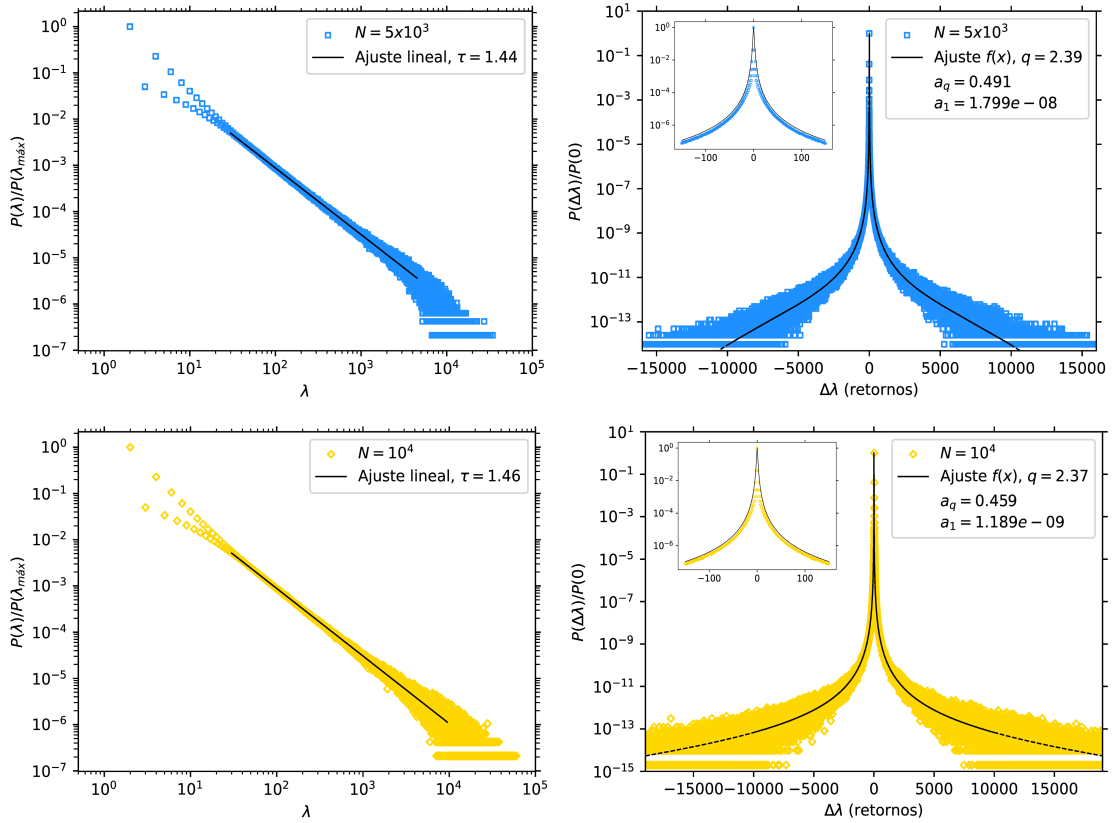


Figura 3.7: Ajuste lineal sobre la distribución de tamaños de fluctuaciones λ en el intervalo $[30, 4500]$ para $N = 5 \times 10^3$ y $[30, 9500]$ para $N = 10^4$ (izq.). Ajuste de la función $f(x)$ (Ec. (3.3)) sobre la distribución de retornos con valor q calculado a priori, en los intervalos $[-9950, 9950]$ para $N = 5 \times 10^3$ y $[-9950, 9950]$ para $N = 10^4$ (der.).

3.2 Resultados en el pseudo punto crítico ($\alpha = 1/N$)

En esta sección también realizamos simulaciones aumentando el tamaño N de sistema. Recordemos que para este valor de α , la distribución estacionaria de probabilidad es uniforme. En este caso, al simular el sistema para valores de N más grandes, surgieron problemas de complejidad computacional. Al estar el sistema en el pseudo punto crítico, la probabilidad de tener tamaños grandes de fluctuaciones λ es mayor. Además, con sistemas de mayor tamaño, es posible que se requiera una cantidad considerable de pasos temporales para que el sistema regrese a su estado de equilibrio en $N/2$ una vez que salió del mismo, lo que implica un tiempo de cómputo considerable para cada fluctuación individual.

Como mencionamos en la sección anterior, establecimos como criterio aceptable la generación de aproximadamente 10^7 fluctuaciones. Sin embargo, al intentar alcanzar este objetivo para los casos de $N = 5 \times 10^3$ y $N = 10^4$, el tiempo de ejecución se volvió inmanejable. Para solucionar este problema realizamos optimizaciones del código, pasando de Python a Julia. Dado el enfoque científico de Julia, esta transición proporcionó mejoras significativas en los tiempos de ejecución. Aún así, para obtener una cantidad aceptable

de puntos en las distribuciones, tuvimos que recurrir a la paralelización del código. Esto significó realizar 1000 evoluciones de 10^4 puntos en paralelo, que luego se unieron. Es importante destacar que el efecto de esta práctica en los resultados es despreciable en comparación con la cantidad de datos obtenidos. De esta manera, pudimos obtener 10^7 puntos en la distribución de tamaños λ y de retornos.

En la Figura 3.8, se presentan las distribuciones de tamaños λ en escala logarítmica para $N = 10^2, 10^3, 5 \times 10^3$ y 10^4 . Se destaca que, a diferencia del caso supracrítico, no existe separación en dos ramas. Además, se observa que para todos los valores de N , la distribución sigue una ley de potencia de exponente $\tau = 1.5$. Podemos ver que para $N = 10^2$, hay un decaimiento exponencial en el rango de valores de $\lambda \sim 10^4 - 10^5$. Para los valores de N más grandes no observamos decaimiento exponencial. Sin embargo, a partir de $\lambda \sim 10^4$ la distribución en estos casos se torna ruidosa, con probabilidades muy bajas de ocurrencia. Con esto podemos concluir que el efecto de tamaño finito deja de existir en rango observado. Lo que esperamos, es que al aumentar la cantidad de puntos en la distribución el decaimiento ocurra eventualmente para estos valores de N , pero que al igual que en el caso supracrítico, si $N \rightarrow \infty$ la ley de potencia se extienda para todo valor de λ . Con las técnicas de paralelización podrían llevarse a cabo mayor cantidad de simulaciones en trabajos futuros.

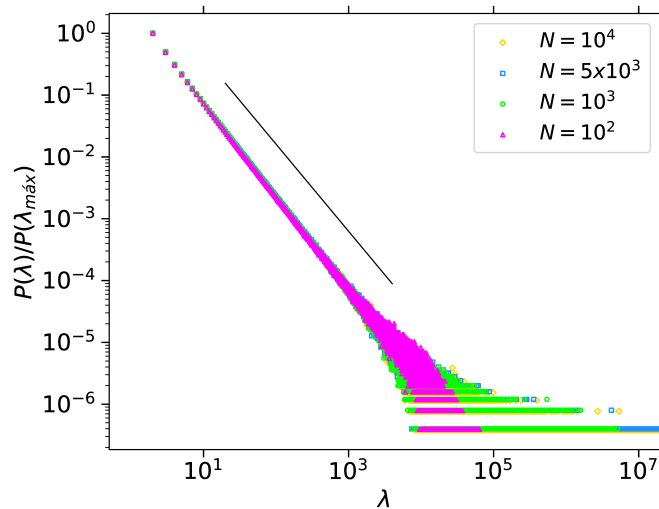


Figura 3.8: Distribuciones de tamaño de fluctuaciones λ en escala logarítmica para $N = 10^2, 10^3, 5 \times 10^3, 10^4$, normalizadas respecto al valor máximo de cada una. Cada distribución tiene aproximadamente 10^7 puntos. El ajuste lineal tiene pendiente $\tau = 1.5$.

En la Fig.3.9, se presenta la distribución de retornos obtenida. Podemos notar que no hay diferencias significativas en la forma de las distribuciones. Se puede observar que para los $N > 10^2$, las distribuciones tienen colas más anchas, pero a la vez más ruidosas.

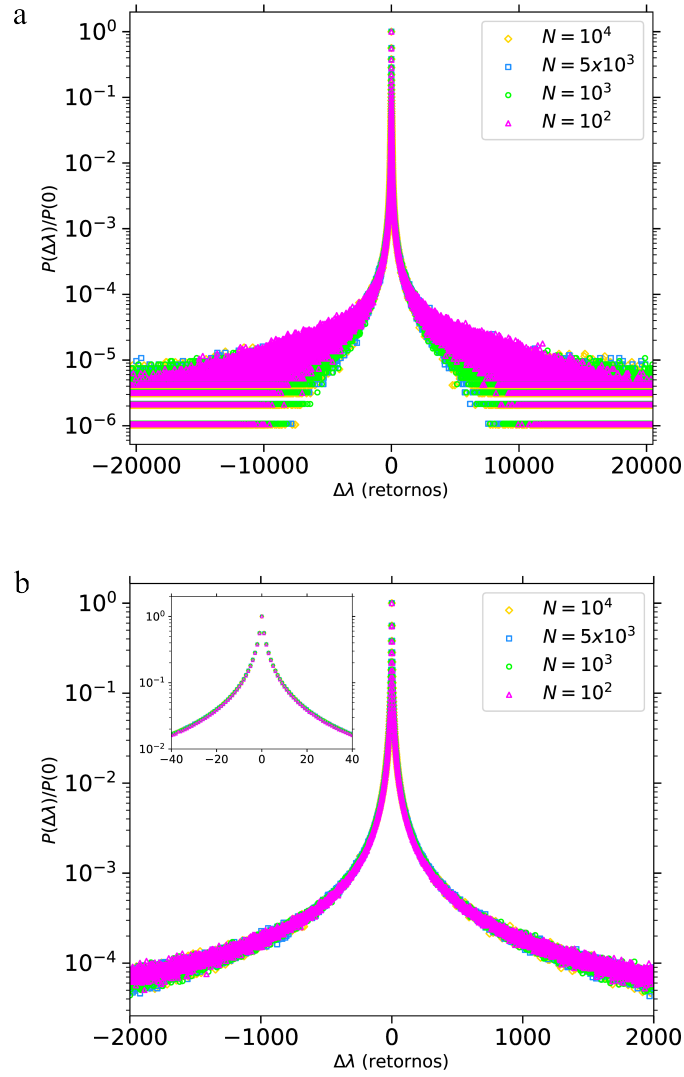


Figura 3.9: Distribuciones de retorno en escala logarítmica para $N = 10^2, 10^3, 5 \times 10^3, 10^4$, normalizadas respecto al valor máximo de cada una. Cada distribución tiene aproximadamente 10^7 puntos.

En las Fig. 3.10 y 3.11 se muestran los ajustes lineales realizados sobre las distribuciones de tamaños de λ y las funciones ajustadas sobre la distribución de retornos. El ajuste lineal se realizó con el mismo criterio que en la sección anterior. Como el valor del parámetro $a_1 \ll 1$ en el caso de $N = 10^2$, para los demás valores de N ajustamos una q-Gaussiana. Reescribiendo la Ec. (3.1) de la siguiente manera:

$$g(x) = \left[1 - (1 - q)\beta x^2 \right]^{1/(1-q)}, \quad (3.5)$$

donde $x = \Delta\lambda$ y calculando q con la Ec. (3.2), realizamos los ajustes sobre la distribución de retornos para $N = 10^3, 5 \times 10^3, 10^4$.

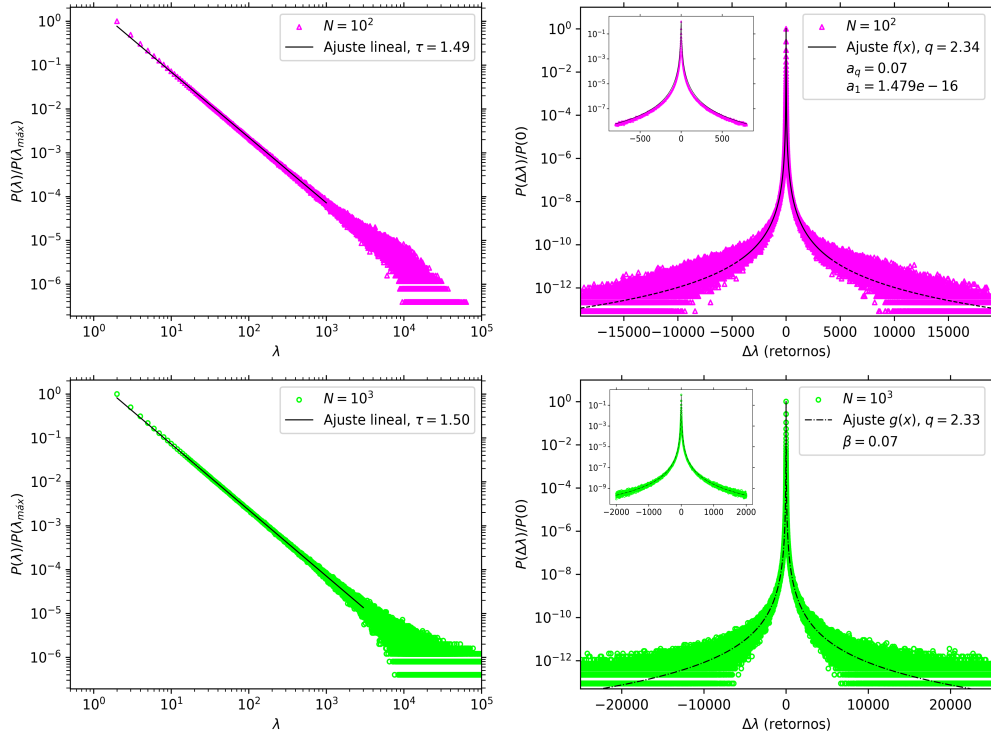


Figura 3.10: Ajuste lineal sobre la distribución de tamaños de fluctuaciones λ para $N = 10^2$ y $N = 10^3$ (izq.). Ajuste sobre la distribución de retornos de la función $f(x)$ (Ec. (3.3)) ($N = 10^2$) y $g(x)$ (Ec. (3.5)) ($N = 10^3$) con valor q calculado a priori (der.).

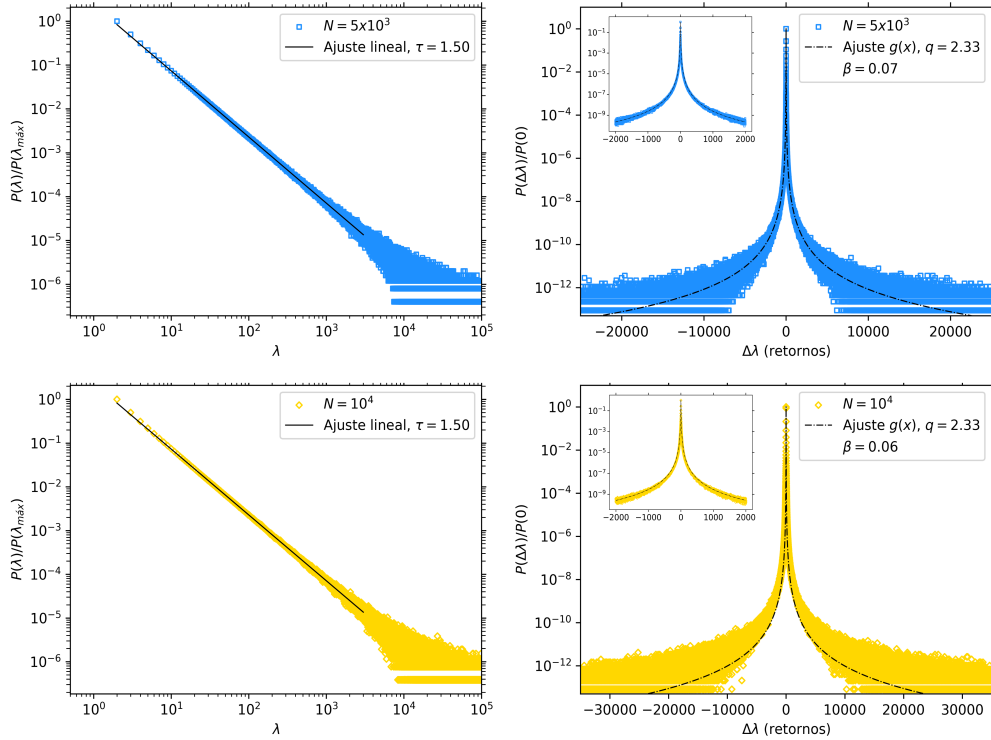


Figura 3.11: Ajuste lineal sobre la distribución de tamaños de fluctuaciones λ para $N = 5 \times 10^3$ y $N = 10^4$ (izq.). Ajuste sobre la distribución de retornos de la función $g(x)$ con valor q calculado a priori (der.).

Es importante remarcar que el parámetro a_q del caso $N = 10^2$ coincide con el parámetro β de los casos $N = 10^3, 5 \times 10^3$. Con estos resultados, podemos decir que en el caso del pseudo punto crítico, la distribución de retornos se ajusta a una q-Gaussiana de parámetro $\beta \approx 0.07$.

Conclusiones

En este trabajo estudiamos la dinámica temporal de un modelo de urnas que denominamos *Noisy Voter Model*. Vimos que el mismo se puede definir como una combinación de dos modelos ya conocidos: el modelo de Ehrenfest y el del votante. La combinación se logra definiendo a α como la probabilidad de ocurrencia de un paso temporal de Ehrenfest y a $1 - \alpha$ como la del votante. A lo largo del marco teórico, estudiamos el comportamiento de estos dos modelos, para poder entender qué conductas del nuestro se deben a la presencia de cada uno. Entendimos que el caso $\alpha = 1$ se reduce al modelo de Ehrenfest, y el de $\alpha = 0$ al del votante.

Con base en investigaciones previas, hemos caracterizado el comportamiento del sistema en tres regiones distintas: el pseudo punto crítico ($\alpha = 1/N$), la región supracrítica ($\alpha > 1/N$) y la región subcrítica ($\alpha < 1/N$). En el caso del valor pseudo crítico, la distribución de probabilidad estacionaria es uniforme, lo que implica que todas las configuraciones de las urnas son igualmente probables. Para valores menores de α , la configuración más probable cambia, favoreciendo aquellas en las que una urna está llena y la otra vacía en lugar de tener ambas con la misma cantidad de elementos. Esto conlleva a que los estados 0 y N sean más probables que el estado $N/2$. En este sentido, sería interesante realizar un análisis similar centrado en los tiempos de retorno a los estados 0 o N . En este caso un enfoque numérico puede ser muy costoso, pero uno analítico como se realizó en [6] podría ser de utilidad.

Para el caso supracrítico pudimos replicar los comportamientos en función de α obtenidos en los trabajos [5], [7]. Observamos la separación en rama par e impar de la distribución de tamaños de λ en dos leyes de potencia con pendientes $\tau_{par} = 1.4$ y $\tau_{impar} = 0.5$. Sería interesante estudiar en trabajos futuros el comportamiento cuando N aumenta en estos casos, buscando la forma de relacionar los exponentes de cada rama con la distribución de retornos que también se divide en una rama par y otra impar.

Pudimos verificar, para el caso del pseudo punto crítico y supracrítico, que nuestro sistema presenta comportamiento tipo ley de potencia para la distribución de primer pasaje al punto de atracción $N/2$. Observamos que el decaimiento dado por efecto de tamaño finito disminuye gradualmente con el aumento de N en el caso supracrítico. Por otro lado, en el caso pseudo crítico, deja de estar presente en el rango observado de tamaños λ para $N > 10^2$. Estos hallazgos respaldan nuestra hipótesis y sugieren que para $N \rightarrow \infty$, la ley de potencia estará presente en todo el rango de valores de λ .

El valor del exponente obtenido en las leyes de potencia es en todos los casos cercano a 1.5. Este valor es recurrente en muchos sistemas que exhiben leyes de potencia, lo cual es una característica intrigante. Sugiere que la emergencia de este comportamiento podría no depender de la dinámica específica del modelo considerado, sino más bien de alguna propiedad general que aún no se comprende completamente. Este fenómeno plantea preguntas interesantes sobre la universalidad de ciertos comportamientos en sistemas complejos.

Además, realizamos un análisis de las distribuciones de retornos asociadas al sistema y su relación con las distribuciones q -Gaussianas. En todos los casos, obtuvimos el parámetro q a priori, con valores en el rango de 2.3 a 2.4. Para el caso supracrítico, donde el decaimiento por tamaño finito debe ser considerado, ajustamos una función que combina una gaussiana con una q -Gaussiana. Observamos que el parámetro a_1 , que representa el peso gaussiano, disminuye a medida que aumenta N . Esto nos hace pensar que en el límite $N \rightarrow \infty$, la distribución de retornos es una q -Gaussiana. Para el caso del pseudo punto crítico, las distribuciones presentan colas más pesadas, y ajustamos directamente una q -Gaussiana. Determinamos que el parámetro β de la q -Gaussiana es aproximadamente 0.07.

Es interesante explorar un sistema donde las leyes de potencia dan lugar a distribuciones de probabilidad tipo q -Gaussianas, ya que esto sugiere una conexión entre estas distribuciones y la teoría desarrollada por Tsallis. Aunque esto no constituye una prueba formal, abre la puerta a una interesante línea de investigación. Para futuros trabajos, sería beneficioso considerar herramientas analíticas similares a las empleadas en [6], así como aumentar el número de simulaciones utilizando técnicas de paralelización para obtener resultados más precisos y exhaustivos.

Bibliografía

- [1] A. Kirman, “Ants, rationality, and recruitment,” *The Quarterly Journal of Economics*, vol. 108, no. 1, pp. 137–156, 1993. [Online]. Available: <https://EconPapers.repec.org/RePEc:oup:qjecon:v:108:y:1993:i:1:p:137-156..>
- [2] C. Tsallis and U. Tirnakli, “Nonadditive entropy and nonextensive statistical mechanics – some central concepts and recent applications,” *Journal of Physics: Conference Series*, vol. 201, no. 1, p. 012001, Dec. 2010. DOI: [10.1088/1742-6596/201/1/012001](https://doi.org/10.1088/1742-6596/201/1/012001). [Online]. Available: <https://dx.doi.org/10.1088/1742-6596/201/1/012001>.
- [3] N. Johnson and S. Kotz, *Urn Models and Their Application: An Approach to Modern Discrete Probability Theory* (Approach to Modern Discrete Probability Theory). Wiley, 1977, ISBN: 9780471446309. [Online]. Available: <https://books.google.com.ar/books?id=ZBfvAAAAAAAJ>.
- [4] L. Takács, “On an urn problem of paul and tatiana ehrenfest,” *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 86, no. 1, pp. 127–130, 1979. DOI: [10.1017/S0305004100000694](https://doi.org/10.1017/S0305004100000694).
- [5] F. Perachia, “Estudio de la dinámica temporal de un modelo de urnas : Probabilidades de primer pasaje y criticalidad auto-organizada,” Bachelor’s Thesis, Facultad de Matemática, Astronomía, Física y Computación. Universidad Nacional de Córdoba., 2020.
- [6] F. Perachia, P. Román, and S. A. Menchón, “Noisy voter model: Explicit expressions for finite system size,” *Phys. Rev. E*, vol. 106, p. 054155, 5 Nov. 2022. DOI: [10.1103/PhysRevE.106.054155](https://doi.org/10.1103/PhysRevE.106.054155). [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevE.106.054155>.
- [7] B. Bakar and U. Tirnakli, “Return distributions in dog-flea model revisited,” *Physica A: Statistical Mechanics and its Applications*, vol. 389, no. 17, pp. 3382–3386, Sep. 2010, ISSN: 0378-4371. DOI: [10.1016/j.physa.2010.04.037](https://doi.org/10.1016/j.physa.2010.04.037). [Online]. Available: <http://dx.doi.org/10.1016/j.physa.2010.04.037>.
- [8] A. Celikoglu, U. Tirnakli, and S. M. D. Queirós, “Analysis of return distributions in the coherent noise model,” *Phys. Rev. E*, vol. 82, p. 021124, 2 Aug. 2010. DOI: [10.1103/PhysRevE.82.021124](https://doi.org/10.1103/PhysRevE.82.021124). [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevE.82.021124>.

- [9] B. Bakar and U. Tirnakli, “Analysis of self-organized criticality in ehrenfest’s dog-flea model,” *Physical Review E*, vol. 79, no. 4, 2009, ISSN: 1550-2376. DOI: [10.1103/physreve.79.040103](https://doi.org/10.1103/PhysRevE.79.040103). [Online]. Available: <http://dx.doi.org/10.1103/PhysRevE.79.040103>.

Los abajo firmantes, miembros del Tribunal de evaluación de tesis, damos fe que el presente ejemplar impreso se corresponde con el aprobado por este Tribunal.