



Gradu amaierako lana /Trabajo fin de grado
Fisikako gradua/ Grado en Física

Estudio sociofísico de la política española

Egilea/ Autor:
Ander Fernández Rúa
Zuzendaria/Director:
Jesús Martínez Perdiguero

Estudio sociofísico de la política española © 2025 by Ander Fernández
Rúa is licensed under CC BY-SA 4.0

Leioa, 2025eko urtarrilaren 1a / Leioa, 1 de enero de 2025

Índice

1. Introducción y objetivos	3
1.1. Objetivos	3
1.2. Introducción	3
1.3. Historia de la sociofísica	4
2. Marco Teórico	6
2.1. Modelo de Ising	6
2.1.1. Aplicaciones clásicas	8
2.1.2. Aplicación sociofísica: modelo de Sznajd	8
2.2. modelo de Markov	10
3. Entorno de desarrollo y herramientas	12
4. Extracción y procesamiento de datos	14
4.1. Datos históricos de las elecciones al senado	14
4.2. Datos históricos de las elecciones generales	15
5. Experimentos	16
5.1. Universalidad de la dinámica de voto	16
5.2. Modelo de Markov para el comportamiento electoral	19
5.2.1. Resultados experimentales	21
5.3. Modelo de voto	23
5.3.1. Análisis inicial y calibración del modelo	26
5.3.2. Resultados experimentales	29
6. Conclusiones	33

1. Introducción y objetivos

1.1. Objetivos

En este trabajo se ha llevado a cabo un análisis de la política española desde la perspectiva de la sociofísica destacando tres objetivos principales. En primer lugar, evaluar la idoneidad de este ámbito como objeto de estudio sociofísico. En segundo lugar, comprobar la hipótesis de que el sistema electoral español es un sistema con memoria, haciendo uso de modelos de Markov. En tercer lugar, proponer y analizar un modelo físico capaz de reproducir la dinámica electoral y ajustarlo a los datos reales disponibles.

1.2. Introducción

La sociofísica es un campo de estudio que, a primera vista, puede parecer poco convencional: propone utilizar herramientas de la física estadística para comprender y analizar fenómenos sociales y políticos. Aunque la idea de comparar átomos con personas suene atrevida, en realidad se basa en una observación sencilla: en diversos contextos (elecciones, difusión de noticias, comportamientos de consumo, etc), los individuos muestran patrones colectivos que, al menos en cierta medida, pueden describirse mediante modelos similares a los que se emplean en la física de la materia condensada.

Por supuesto, el comportamiento humano es infinitamente más rico y matizado que el de un conjunto de partículas. Sin embargo, es precisamente esa riqueza la que hace que, a menudo, sea imposible abarcarlo todo si no se eligen cuidadosamente algunos rasgos esenciales. En la sociofísica, se opta por simplificar, no para ignorar la complejidad humana, sino para ver si existen regularidades que emergen incluso cuando se toman en cuenta unas pocas variables clave (por ejemplo, la influencia de la opinión de los demás o la experiencia previa de cada individuo).

Al igual que en la física se usan modelos teóricos y simulaciones por ordenador para reproducir fenómenos naturales, en sociofísica se realizan experimentos virtuales: se diseña un conjunto de *agentes* (personas ficticias con ciertas reglas de comportamiento), se deja que interactúen según determinadas condiciones, y luego se comparan los resultados con datos reales. Si el modelo capta rasgos importantes de la realidad, se pueden hacer predicciones razonables sobre procesos como la formación de consenso, la aparición de polarizaciones o el impacto de la memoria histórica en la evolución de la intención de voto.

Esta aproximación no pretende describir con detalle toda la complejidad de la conducta humana; es imposible encajar en una simulación cada matiz cultural, psicológico y económico. Pero tampoco hace falta: la clave de la física estadística está precisamente en buscar patrones globales que se mantienen estables cuando aumentan las interacciones entre los agentes. A menudo, estos patrones pueden resultar sorprendentes o incluso contradictorios con la intuición. Sin embargo, ayudan a explicar por qué ciertos comportamientos colectivos persisten a lo largo del tiempo o por qué se propagan rápidamente algunas creencias o tendencias.

Así, la sociofísica brinda un punto de vista refrescante y hasta cierto punto provocativo: ¿podemos aprender algo valioso sobre la organización social humana observando sistemas físicos relativamente simples? Sus defensores argumentan que al igual que la aparente sencillez de los átomos y moléculas no les impide generar estructuras complejas y tampoco impide a los físicos extraer leyes universales, la sociedad podría regirse en parte por mecanismos reconocibles, y es justo ahí donde la sociofísica realiza su aportación.

Por supuesto, este enfoque puede encontrar escepticismo en otros ámbitos, ya que implica adoptar una perspectiva cuantitativa y reduccionista para estudiar algo tan difícil de medir como el comportamiento humano. No obstante, en la búsqueda de comprensión de fenómenos tan variados como la difusión de la desinformación o el auge de determinados movimientos políticos, las ideas y métodos de la física estadística ofrecen intuiciones valiosas.

A continuación, se repasa el recorrido histórico de la sociofísica: desde sus primeras tentativas—cuando apenas se intuía la utilidad de este cruce disciplinar—hasta los modelos más avanzados de hoy. Con ello, se pretende mostrar cómo ha ido consolidándose esta disciplina y por qué ofrece un marco prometedor para estudiar aspectos esenciales de la política española.

1.3. Historia de la sociofísica

La sociofísica parte de la idea de que los mismos métodos y conceptos que se usan en física estadística también pueden ayudar a entender fenómenos sociales y políticos. Aunque hoy en día sus aplicaciones resultan cada vez más visibles y aceptadas, el camino hasta aquí no ha sido sencillo. Desde los años setenta, esta disciplina ha tenido que abrirse paso enfrentando críticas y dudas sobre su validez académica.

A mediados de la década de 1970, la física de la materia condensada vivía momentos de gran efervescencia gracias a investigadores como Kenneth G. Wilson, quien introdujo la idea de los grupos de renormalización para explicar cómo pequeñas interacciones locales pueden generar cambios radicales en sistemas más grandes. Inspirado por este enfoque, el físico Serge Galam comenzó a preguntarse si dichas herramientas podrían emplearse también para analizar la dinámica colectiva de las sociedades humanas. Sin embargo, no todo el mundo recibió esta propuesta con agrado. Durante los setenta y ochenta, la mayoría de la comunidad científica consideraba que aplicar fórmulas y teorías creadas para átomos o partículas a las personas era un atrevimiento excesivo. Para muchos, los seres humanos y sus complejas interacciones sociales no podían entenderse empleando métodos tan fríos o mecánicos como los de la física.

En 1982, Galam y su equipo crearon el primer modelo sociofísico [1], inspirado en el famoso modelo de Ising (habitualmente usado para estudiar transiciones de fase en física). Este trabajo demostró cómo interacciones locales y factores externos pueden conducir a decisiones colectivas, como las huelgas. Aunque en ese momento pasó bastante desapercibido, hoy se considera un texto fundamental en la historia de la sociofísica. Pese al escepticismo inicial, Galam y otros entusiastas siguieron desarrollando modelos y publicando sus resultados tanto en revistas de física como en algunas de ciencias sociales,

analizando la propagación de opiniones, la formación de consensos [2] e incluso el comportamiento electoral [3]. Estos esfuerzos dieron forma a la base teórica de la sociofísica, aunque todavía faltaba un largo camino para que fuera tomada en serio por la mayoría de los académicos.

Hacia la mitad de la década de 1990, la idea de aplicar métodos de la física a otras áreas ganó popularidad gracias a la llamada *econofísica* [4], una rama que se centra en el estudio de los mercados financieros con herramientas de física estadística. Esto sirvió de respaldo indirecto a la sociofísica y atrajo a nuevos investigadores al campo. Como resultado, la cantidad de publicaciones comenzó a aumentar y se formaron grupos de trabajo dedicados específicamente a estos temas. Con el paso del tiempo, la sociofísica empezó a ser vista por algunos como un puente entre la física y las ciencias sociales, ofreciendo formas novedosas de estudiar el comportamiento de grandes grupos humanos. Galam, tras haber luchado prácticamente en solitario durante años, vio cómo se multiplicaban los trabajos que seguían su línea de pensamiento.

En la actualidad, la sociofísica ha avanzado lo suficiente como para contar con una posición propia dentro de la investigación sobre política, economía y dinámicas sociales. Serge Galam, en sus escritos más recientes, insiste en la importancia de la rigurosidad metodológica para evitar malas interpretaciones o abusos de conceptos físicos en contextos sociales. Su mensaje es claro: la sociofísica puede aportar enfoques valiosos, siempre y cuando se utilice con cuidado y se mantenga un diálogo abierto con otras disciplinas.

2. Marco Teórico

Se procede a realizar un análisis teórico de los modelos empleados en esta investigación. En primer lugar, se analiza el modelo de Ising y su aplicación a la sociofísica: el modelo de Sznajd. Después, se explica el modelo de Markov, su desarrollo matemático y sus aplicaciones más comunes para contextualizarlo y mostrar su utilidad en este trabajo.

2.1. Modelo de Ising

El modelo de Ising [5] (conocido también como modelo de Lenz-Ising), llamado así en honor a los físicos Ernst Ising y Wilhelm Lenz, es un modelo matemático usado para explicar el ferromagnetismo en el campo de la mecánica estadística. En este modelo se utilizan valores discretos para representar los momentos dipolares magnéticos de spins atómicos, cada uno de los cuales puede tomar solo dos estados diferentes(+1 o -1). Estas partículas se organizan en una estructura específica y generalmente en una disposición periódica. Cuando las partículas magnéticas adyacentes están alineadas (coinciden en su orientación), la energía del sistema es menor que cuando están desalineadas; sin embargo, esta tendencia natural de minimizar la energía se ve afectada por el calor presente en el sistema y esto provoca la aparición de diferentes configuraciones estructurales. Este planteamiento ayuda a identificar transiciones de fase como una representación simplificada de los fenómenos reales. Es importante señalar que el modelo de Ising en dos dimensiones en una red cuadrada es uno de los sistemas estadísticos más simples que muestra un cambio de fase.

El físico Wilhelm Lenz creó este modelo en 1920 y le planteó como desafío a su alumno Ernst Ising quien logró resolver el problema unidimensional (1925) basándose en su tesis de 1924.[2] En la versión unidimensional no se observan cambios de fase evidentes; por el contrario, en la variante bidimensional en una red cuadrada resultó ser mucho más complicada y solo fue analizada de manera detallada posteriormente por Lars Onsager en 1944.

Pensemos en un conjunto de posiciones en una red donde cada posición tiene conexiones adyacentes formando una red de múltiples dimensiones. En cada posición de esta red se establece una variable discreta σ_k que puede adoptar uno de los dos valores disponibles (-1 o +1) para representar el spin de dicha posición. Una disposición de spins, $\sigma = \{\sigma_k\}_{k \in \Lambda}$, es una asignación de valores de spin a cada posición de la red.

Para dos posiciones vecinas i y j en una red conectada entre sí existe una relación mutua representada por una interacción J_{ij} entre ellos. Además, cada posición j en la red se ve afectada por un campo magnético externo h_j que interactúa directamente sobre la posición en cuestión. Los niveles energéticos asociados a una configuración específica se definen a través de una función particular, el Hamiltoniano, que varía según las interacciones entre los espines adyacentes y la influencia del campo magnético externo:

$$H(\sigma) = - \sum_{\langle i,j \rangle} J_{ij} \sigma_i \sigma_j - \mu \sum_j h_j \sigma_j.$$

La probabilidad de que el sistema esté una configuración particular, en función de la temperatura inversa $\beta = \frac{1}{k_B T}$, viene determinada por una distribución estadística específica, la distribución de Boltzmann:

$$P_\beta(\sigma) = \frac{e^{-\beta H(\sigma)}}{Z_\beta},$$

donde Z_β , la función de partición, es un factor de normalización que garantiza que las probabilidades estén equilibradas de manera adecuada:

$$Z_\beta = \sum_{\sigma} e^{-\beta H(\sigma)}.$$

Además, en relación con cualquier propiedad asociada al spin $f(\sigma)$, que puede ser considerada como una característica observada del sistema, se establece un valor esperado que refleja su promedio ponderado de acuerdo con las probabilidades de las posibles configuraciones:

$$\langle f \rangle_\beta = \sum_{\sigma} f(\sigma) P_\beta(\sigma).$$

Estas probabilidades $P_\beta(\sigma)$ indican la probabilidad de que, en un estado de equilibrio establecido, el sistema se encuentre en determinada condición.

La convención común es emplear un signo negativo en cada término de la función que representa la energía en los modelos físicos para distinguir entre interacciones ferromagnéticas y antiferromagnéticas, basándose en el signo del término de interacción entre pares: positivo para la primera ($J_{ij} > 0$) y negativo para la segunda ($J_{ij} < 0$). Cuando no hay interacción entre los espines ($J_{ij} = 0$), se consideran independientes unos de otros.

En un sistema de material ferromagnético, los spins suelen orientarse en la misma dirección; es más probable que los spins adyacentes compartan el mismo sentido ($\sigma_i = \sigma_j$). Por otro lado, en un modelo antiferromagnético, los spins adyacentes tienen preferencia por tener sentidos opuestos ($\sigma_i \neq \sigma_j$).

La convención sobre signos también establece cómo un espín en un punto interactúa con un campo magnético exterior. Dependiendo del sentido y la intensidad del campo (h_j), el espín tiende a alinearse en la misma dirección ($h_j > 0$), en dirección contraria ($h_j < 0$) o no se ve afectado si no hay campo presente ($h_j = 0$).

Transiciones de fase

En su tesis doctoral de 1924, Ising resolvió el modelo para el caso de una dimensión, que puede entenderse como una red lineal en la que cada punto interactúa únicamente con sus vecinos inmediatos, a la izquierda y a la derecha. En una dimensión, la solución no

muestra transición de fase. En esta situación, las correlaciones entre los spins disminuyen exponencialmente conforme aumentan las distancias entre ellos:

$$\langle \sigma_i \sigma_j \rangle_\beta \leq C \exp(-c(\beta)|i - j|),$$

manteniendo así el sistema desordenado. A partir de este hallazgo, Ising concluyó erróneamente que este modelo no presentaba transiciones de fase en ninguna dimensión.

El modelo de Ising muestra un cambio de fase entre un estado ordenado y uno desordenado cuando se examina en dos o más dimensiones. Específicamente, el sistema está en un estado caótico para valores bajos de β , mientras que manifiesta una estructura ferromagnética ordenada para valores altos:

$$\langle \sigma_i \sigma_j \rangle_\beta \geq c(\beta) > 0.$$

Este fenómeno fue inicialmente evidenciado por Rudolf Peierls en 1936 mediante una técnica conocida como el razonamiento de Peierls.

2.1.1. Aplicaciones clásicas

La razón principal detrás de la creación del modelo fue el fenómeno del magnetismo ferroso. El hierro tiene propiedades magnéticas y, una vez se magnetiza, conservará esa característica durante un período considerablemente más largo que el de los tiempos atómicos típicos.

En el siglo XIX, se creía que los campos magnéticos eran producidos por corrientes dentro de la materia. Ampère propuso que los imanes permanentes se originaban en corrientes atómicas fijas, pero el movimiento de partículas cargadas clásicas no podía explicar completamente estas corrientes persistentes en los materiales ferromagnéticos. Para que un material sea ferromagnético, se requiere que sus átomos posean momentos magnéticos permanentes no atribuibles al movimiento de cargas clásicas.

Con el descubrimiento del spin del electrón, se comprendió que el magnetismo se origina por la alineación de un gran número de spins de electrones en la misma dirección. Surgió entonces la interrogante de cómo los spins “deciden” alinearse, dado que los electrones en un extremo del imán no interactúan directamente con aquellos en el otro extremo. El modelo de Ising fue concebido para investigar si las interacciones locales entre electrones vecinos podrían explicar este fenómeno.

2.1.2. Aplicación sociofísica: modelo de Sznajd

El modelo de Sznajd, también conocido como *Unidos permanecemos, divididos caemos* (USDF, por sus siglas en inglés), es un modelo de sociofísica introducido en el año 2000 [6] con el objetivo de comprender la dinámica de las opiniones en sistemas sociales. Este

modelo extiende el modelo de spin de Ising y utiliza el concepto de validación social para describir la propagación y cambio de opiniones dentro de una comunidad. En términos simples, el modelo se basa en dos principios fundamentales:

- **Validación social:** Si dos personas comparten la misma opinión, sus vecinos tenderán a adoptar esa opinión.
- **La discordia destruye:** Si dos personas adyacentes tienen opiniones diferentes, sus vecinos tenderán a adoptar opiniones opuestas a las de sus vecinos más cercanos.

Para modelar este sistema, consideremos una comunidad que, una y otra vez, debe tomar una posición sobre un tema, por ejemplo, votar en un sistema bipartidista. Si cada miembro de la comunidad puede adoptar solo dos actitudes posibles (denotadas como A o B), en varias votaciones se espera observar una diferencia m entre los votantes a favor de A y en contra. El objetivo del modelo es analizar la evolución temporal de m . Para ello, se utiliza una cadena de spins de Ising S_i ($i = 1, 2, \dots, N$) con las siguientes reglas dinámicas:

1. Si $S_i S_{i+1} = 1$, entonces S_{i-1} y S_{i+2} adoptan la dirección del par $(i, i+1)$. (r1)
2. Si $S_i S_{i+1} = -1$, entonces S_{i-1} adopta la dirección de S_{i+1} , y S_{i+2} adopta la dirección de S_i . (r2)

Estas reglas describen la influencia de un par de individuos sobre las decisiones de sus vecinos más cercanos. Cuando los miembros de un par comparten la misma opinión, sus vecinos adoptan esa opinión. Por el contrario, si los miembros de un par tienen opiniones opuestas, los vecinos más cercanos de cada uno discrepan con ellos. Estas reglas dinámicas conducen a tres estados estables:

1. El consenso completo (*estado ferromagnético*), donde todos votan por A o por B .
2. El estado de alternancia (*estado antiferromagnético*), donde las opiniones alternan entre A y B .
3. Un estado especial de 50 % de votos por A y 50 % de votos por B .

El tercer estado solo se alcanza de manera muy particular, cuando cada miembro de la comunidad discrepa con su vecino más cercano. Esto es consistente con la idea principal del modelo.

En términos de relevancia, el modelo de Sznajd ha emergido como una variación simple pero significativa del modelo de Ising prototípico, que ha sido crucial en la sociofísica para el estudio de fenómenos colectivos críticos.

2.2. modelo de Markov

En el ámbito de la probabilidad y la estadística surge la noción de una cadena o proceso de Markov. Este concepto describe un suceso estocástico que representa una sucesión de eventos, en la cual la probabilidad de cada evento está determinada únicamente por el estado en el que se encontraba el sistema en el evento anterior. Dependiendo de cómo avanza el tiempo en la secuencia, se pueden distinguir dos tipos principales: una secuencia interminable y contable que avanza entre estados en intervalos discretos de tiempo se denomina cadena de Markov en tiempo discreto (DTMC), mientras que una secuencia en tiempo continuo se conoce como cadena de Markov en tiempo continuo (CTMC). Estos procesos llevan el nombre del matemático ruso Andréi Markov.

Los modelos estadísticos basados en las cadenas de Markov tienen múltiples aplicaciones y son ampliamente utilizados para simular procesos del mundo real y distribuciones complejas de probabilidad. Uno de los métodos más destacados es el método de Monte Carlo de cadenas de Markov, que permite generar muestreros representativos de distribuciones complejas. Este enfoque es particularmente útil en campos como la física y el procesamiento de señales, entre otros.

Los cambios en un sistema dentro de una cadena de Markov se conocen como transiciones de estado, y las probabilidades asociadas a estas transiciones se representan mediante una matriz de transición junto con un estado inicial o una distribución inicial dentro del espacio de estados. La definición del proceso usualmente incluye todos los estados y transiciones posibles, lo que asegura la continuidad del sistema y evita que termine de forma abrupta. En un sistema que opera en tiempo discreto, los cambios de estado ocurren de manera aleatoria en cada paso.

La teoría de Markov establece que la probabilidad condicional de transición hacia un estado futuro, ya sea en el siguiente paso o en pasos más lejanos, está determinada únicamente por el estado actual del sistema y no por los estados previos. Debido al carácter aleatorio inherente del sistema, resulta difícil predecir con exactitud el estado futuro en un momento específico. Sin embargo, es posible anticipar las propiedades estadísticas que caracterizan dicho estado futuro, y en muchas ocasiones estas propiedades estadísticas son de mayor interés para el análisis que la predicción puntual de un estado. Puesto que en este trabajo se emplean cadenas de Markov discretas, se procede a realizar su análisis matemático.

Cadena de Markov en tiempo discreto

Una cadena de Markov en tiempo discreto es una secuencia de variables aleatorias X_1, X_2, X_3, \dots que satisface la propiedad de Markov, lo que significa que la probabilidad de moverse al siguiente estado depende únicamente del estado actual y no de los estados anteriores. Formalmente, esto se expresa como:

$$P(X_{n+1} = x \mid X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = P(X_{n+1} = x \mid X_n = x_n),$$

si ambas probabilidades condicionales están bien definidas y $P(X_1 = x_1, \dots, X_n = x_n) > 0$.

Los posibles valores de X_t forman un conjunto contable S , denominado espacio de estados de la cadena.

Si el espacio de estados es finito, la distribución de probabilidad de transición puede representarse mediante una matriz, llamada matriz de transición, donde el elemento (i, j) de la matriz P está definido como:

$$p_{ij} = P(X_{n+1} = j \mid X_n = i).$$

Dado que cada fila de P suma uno y todos sus elementos son no negativos, P es una matriz estocástica derecha.

Se introduce la noción de distribución estacionaria, es decir, la distribución de probabilidad en el estado estacionario del sistema cuando el tiempo tiende a infinito. Una distribución estacionaria π es un vector fila cuyos elementos son no negativos, suman uno y permanecen invariantes bajo la operación de la matriz de transición P . Formalmente, se define como:

$$\pi P = \pi.$$

Comparando esta definición con la de un autovector, se observa que los dos conceptos están relacionados, y que:

$$\pi = \frac{e}{\sum_i e_i},$$

Donde e es un múltiplo normalizado ($\sum_i \pi_i = 1$) de un autovector izquierdo de la matriz de transición P asociado al autovalor 1. Si existe más de un autovector asociado al autovalor 1, una combinación ponderada de los correspondientes estados estacionarios también es un estado estacionario. Para una cadena de Markov, generalmente se estudia el estado estacionario que es el límite de la secuencia de distribuciones a partir de una distribución inicial.

Los valores de una distribución estacionaria π_i están asociados con el espacio de estados de P y sus autovectores, de manera que sus proporciones relativas se mantienen. Dado que los componentes de π son positivos y cumplen la restricción de que su suma es la unidad ($\sum_i 1 \cdot \pi_i = 1$), π puede representarse como el producto escalar de π con un vector cuyos componentes son todos 1. Esto implica que π pertenece al conjunto de vectores en \mathbb{R}^n cuyos componentes son positivos y suman 1, lo que es característico de distribuciones de probabilidad.

Aplicaciones

Los sistemas Markovianos aparecen extensivamente en termodinámica y mecánica estadística, especialmente cuando las probabilidades se utilizan para representar detalles desconocidos o no modelados de un sistema. Esto es posible bajo el supuesto de que las dinámicas del sistema sean invariantes en el tiempo y que ninguna variable relevante necesite ser considerada fuera de lo ya incluido en la descripción de cada estado.

Por ejemplo, un estado termodinámico opera bajo una distribución de probabilidad que puede ser difícil o costoso de adquirir. Por esta razón, el método de Monte Carlo basado en cadenas de Markov (MCMC) se utiliza para tomar muestras aleatorias de una *caja negra* con el fin de aproximar la distribución de probabilidad de los atributos de un rango de objetos.

Además, las cadenas de Markov se emplean en simulaciones de QCD en redes (*lattice QCD*).

3. Entorno de desarrollo y herramientas

En esta sección se detalla el entorno y las herramientas utilizadas para llevar a cabo este proyecto. También se ofrece una descripción de las mismas y se explica por qué fueron elegidas en primera instancia.

Python es ampliamente reconocido por su simplicidad y facilidad de lectura, lo que permite implementar rápidamente algoritmos complejos y manejar grandes cantidades de datos de manera efectiva. Esto es especialmente beneficioso en proyectos que involucran análisis de datos y simulaciones computacionales. Además, Python cuenta con una extensa comunidad de desarrolladores y una amplia gama de bibliotecas enfocadas en estadísticas, visualización de datos y simulaciones computacionales, lo cual garantiza un acceso abundante a recursos y soporte comunitario. Por estas razones, se ha decidido utilizar Python como el lenguaje principal para este proyecto [7].

En este proyecto se han utilizado varias bibliotecas de Python para simplificar el análisis de datos y la ejecución de simulaciones:

1. **NumPy y Pandas:** Estas bibliotecas son indispensables para la manipulación y análisis de datos. NumPy proporciona herramientas eficientes para realizar operaciones matemáticas avanzadas, mientras que Pandas simplifica la gestión y transformación de grandes conjuntos de datos a través de estructuras como los *DataFrames*. Estas herramientas han sido esenciales para el procesamiento y preparación de los datos empleados en esta investigación [8, 9].
2. **Matplotlib y Seaborn:** Estas bibliotecas de visualización han sido fundamentales para mostrar de forma clara e inteligible los resultados de los análisis y simulaciones. Matplotlib proporciona flexibilidad para crear gráficos personalizados, mientras que Seaborn facilita la generación de gráficos estadísticos con un diseño más estético [10].

3. **SciPy y NetworkX:** SciPy se ha empleado para llevar a cabo cálculos estadísticos y análisis complejos, mientras que NetworkX ha sido utilizado para la modelización y análisis de redes, fundamentales para estudiar cómo interactúan los agentes en sistemas socio-físicos.
4. **SimPy:** Esta biblioteca se ha utilizado para crear y examinar simulaciones basadas en eventos discretos, permitiendo modelar interacciones y dinámicas complejas en sistemas políticos .

La combinación de estas herramientas proporciona un entorno robusto para la interpretación de información y la creación de escenarios virtuales, permitiendo abordar el problema desde múltiples perspectivas: desde el tratamiento preliminar de datos hasta la presentación gráfica y el análisis detallado de las dinámicas simuladas.

Para garantizar la portabilidad y la uniformidad del entorno de desarrollo y ejecución del trabajo, se ha empleado Docker. Esta herramienta que permite empaquetar una aplicación junto con todas sus dependencias en un contenedor virtual, el cual puede ser desplegado en cualquier sistema compatible. Este enfoque resuelve problemas comunes relacionados con las discrepancias entre los entornos de desarrollo y producción, asegurando que la aplicación mantenga un funcionamiento consistente independientemente de dónde se ejecute.

La elección de Docker fue motivada por la necesidad de ejecutar el proyecto en distintas máquinas de forma eficiente, evitando problemas de dependencias o configuraciones de entorno. Entre sus principales beneficios destacan:

- **Uniformidad:** Se asegura que el proyecto funcione de manera consistente en todos los sistemas, eliminando los problemas relacionados con las variaciones en las configuraciones locales.
- **Simplicidad en la configuración:** Todo lo necesario para el proyecto, incluidos los requisitos del sistema y las bibliotecas, se define en el `Dockerfile`, lo que facilita la configuración del entorno.
- **Portabilidad:** La implementación en diferentes infraestructuras se simplifica, permitiendo realizar simulaciones y análisis en máquinas diversas sin necesidad de ajustes adicionales.

El uso de Docker ha resultado crucial para mantener la cohesión entre los diferentes entornos de trabajo y ha desempeñado un papel importante en la eficiencia y éxito del proyecto.

Durante la realización del proyecto, se ha empleado GitHub como herramienta clave para la gestión y organización del código fuente. GitHub es una plataforma basada en Git, un sistema de control de versiones distribuido que facilita el manejo y seguimiento de los cambios realizados en los archivos del proyecto.

El uso de GitHub ha brindado múltiples beneficios importantes para la estructura interna del proyecto:

- **Gestión de versiones:** GitHub simplifica el seguimiento de todas las modificaciones efectuadas en los archivos del proyecto. Cada cambio puede ser revisado y revertido si es necesario, lo que proporciona un control detallado sobre la evolución del software.
- **Documentación:** GitHub ofrece recursos para la creación de documentación del proyecto, como un sistema de *wiki* y la gestión de *issues*, fundamentales para mantener un registro claro de las funcionalidades del proyecto, tareas pendientes y problemas identificados y resueltos.

Estas herramientas se han seleccionado y utilizado para asegurar un progreso organizado y eficiente en el trabajo, mejorando tanto el análisis de datos como las simulaciones, que son parte fundamental del proyecto.

4. Extracción y procesamiento de datos

En esta sección, se abordan los detalles referentes a los conjuntos de datos utilizados.

4.1. Datos históricos de las elecciones al senado

En la primera parte de este estudio se utilizan los datos históricos sobre las elecciones al senado desde 1977 hasta 2023, obtenidos de la pagina web del senado español [11]. Este conjunto de datos dispone de las siguientes variables:

1. **Tipo convocatoria:** Identifica el tipo de elección o evento electoral.
2. **Id convocatoria:** Código único para identificar la convocatoria electoral.
3. **ccaa:** Comunidad autónoma a la que pertenece la circunscripción electoral. En este archivo, los valores están vacíos.
4. **prv:** Provincia de la circunscripción electoral. Esta variable también contiene valores vacíos.
5. **circunscripción:** Código de la circunscripción electoral. Sin datos en este archivo.
6. **municipio:** Código del municipio asociado al evento electoral. Sin datos en este caso.
7. **distrito:** Código del distrito electoral. En este archivo, los valores están vacíos.
8. **candidatura:** Nombre del partido político o agrupación electoral al que pertenece el candidato.
9. **nombre:** Nombre de pila del candidato.
10. **primer apellido:** Primer apellido del candidato.

11. **segundo apellido:** Segundo apellido del candidato.
12. **votos:** Número de votos obtenidos por el candidato en la circunscripción correspondiente.

Este conjunto de datos se centra principalmente en los resultados por candidato, ofreciendo información relevante sobre los votos obtenidos y la afiliación partidista. Permite un análisis detallado del comportamiento electoral a nivel de partido y candidato.

4.2. Datos históricos de las elecciones generales

En la segunda parte se hace uso de los datos históricos de las elecciones generales al congreso, obtenidos de la página web del ministerio del interior español [12]. Este conjunto de datos dispone de las siguientes variables:

1. **ID _ Registro:** Identificador único de cada registro en el conjunto de datos.
2. **Nombre _ Convocatoria:** Nombre de la convocatoria electoral.
3. **Fecha _ Convocatoria:** Fecha de celebración de la convocatoria electoral.
4. **Codigo _ Convocatoria:** Código numérico asignado a la convocatoria electoral.
5. **Tipo _ Eleccion:** Tipo de elección celebrada.
6. **Distrito:** Código del distrito electoral en el que se realiza la votación.
7. **Seccion:** Código de la sección electoral dentro del distrito.
8. **Partido:** Nombre del partido político o agrupación electoral que participa en la elección.
9. **Votos:** Número de votos obtenidos por el partido en la sección electoral correspondiente.

La manipulación y preprocesado de los datos se llevó a cabo utilizando Python, aprovechando librerías especializadas como NumPy. Esta herramienta permitió filtrar valores erróneos o atípicos presentes en los conjuntos de datos y garantizar la limpieza necesaria para el análisis posterior. Además, se realizaron operaciones previas fundamentales, como la normalización de variables, para homogeneizar las escalas y facilitar la interpretación de los resultados. Este preprocesado fue clave para obtener una base de datos fiable y adecuada para los análisis posteriores en el marco del trabajo.

5. Experimentos

Después de haber establecido el marco teórico requerido, nos centraremos ahora en la implementación práctica de estos conceptos: la creación de un modelo que ayude a comprender cómo se forma la opinión social en el ámbito político español. Este paso es crucial para avanzar en la comprensión de esta dinámica y para confirmar la utilidad de la teoría sociofísica en entornos experimentales reales.

5.1. Universalidad de la dinámica de voto

El estudio de los datos electorales ha mostrado tendencias en el comportamiento que van más allá de las diferencias culturales y políticas en diversas situaciones sociales y económicas en las que se desarrollan las elecciones políticas actuales. Uno de los logros más notables en esta área es el hallazgo de patrones comunes en los procesos de votación a nivel universal tal y como se muestra en el estudio realizado por Fortunato y Castellano en 2007 [13] al demostrar que en sistemas electorales proporcionales que utilizan listas abiertas existe una distribución de votos entre los candidatos de un mismo partido que sigue un patrón universal independiente del país y el año. Este descubrimiento sugiere que el comportamiento colectivo de los votantes está influenciado principalmente por dinámicas sociales básicas, como la comunicación y el intercambio de opiniones, y menos por factores particulares del entorno sociopolítico.

El estudio demuestra que al ajustar el número de votos recibidos por los candidatos según el tamaño de sus listas y el total de votos obtenidos por el partido político, se pueden identificar patrones recurrentes que pueden ser modelados de forma muy precisa. Estas regularidades se han observado en elecciones realizadas en diferentes países como Italia, Polonia y Finlandia a lo largo de varias décadas, reflejando dinámicas electorales persistentes. Los autores sugieren que estos patrones son consistentes con modelos basados en procesos ramificados, donde las interacciones locales entre los votantes desempeñan un papel fundamental.

El propósito de este apartado es comprobar si este tipo de dinámica universal se observa en el contexto español mediante el análisis de datos de elecciones proporcionales en España, determinando si los patrones de distribución de votos entre candidatos del mismo partido pueden ser descritos por el modelo propuesto. A través de esta investigación experimental no solo se busca confirmar la presencia de la universalidad en las elecciones españolas, sino también interpretar posibles desviaciones o particularidades que puedan surgir debido al contexto sociopolítico local. En los apartados siguientes se detallarán los métodos empleados y los resultados alcanzados junto a su análisis para ofrecer nuevas visiones sobre cómo pueden aplicarse los enfoques socio-físicos al estudio de la política en España.

Siguiendo el enfoque metodológico propuesto por Fortunato y Castellano, se busca identificar regularidades en la distribución de sufragios entre candidatos pertenecientes a un mismo partido político y determinar si estas regularidades están libres de influencias externas como el trasfondo histórico o cultural del entorno político.

Para lograr esto, se han empleado datos de elecciones proporcionales donde los electores expresan sus preferencias individuales por candidatos dentro de las listas de sus partidos políticos correspondientes. Se ha optado por realizar esta parte del estudio con los datos de las elecciones al senado ya que es el único proceso electoral en España con listas abiertas. Se han analizado aspectos clave como la cantidad de votos recibidos por cada candidato específico (v_i), el tamaño de la lista a la que pertenecían (Q_{l_i}) y el total de sufragios obtenidos por dichas listas en la circunscripción respectiva (N_{l_i}).

El proceso de análisis comienza normalizando de los datos para neutralizar el efecto de variables como el tamaño de los grupos y las listas en los resultados obtenidos. Esto permite crear una métrica relativa definida como:

$$v_0 = \frac{N}{Q}, \quad (1)$$

donde v_0 representa el promedio de votos por candidato en una lista. Basándose en esta métrica, se introduce un ratio rescalado que evalúa el rendimiento de cada candidato en relación con el promedio de su lista:

$$vQ/N = \frac{v}{v_0}. \quad (2)$$

Este ratio permite identificar candidatos con un desempeño por debajo (< 1), igual ($= 1$) o superior (> 1) al promedio de su lista, lo que a su vez posibilita una comparación equitativa entre listas de distinta longitud.

Después de eso, se reúnen los datos de candidatos con ratios parecidos, creando un histograma y se examinan las distribuciones obtenidas. Los resultados de Fortunato indican que estas distribuciones convergen en una curva universal que no depende del tamaño del grupo o de la lista, sino de la variable vQ/N tal que:

$$P(v, Q, N) = F\left(\frac{vQ}{N}\right), \quad (3)$$

donde $F(vQ/N)$ es una función de distribución universal que describe el comportamiento colectivo de los votantes. Estos hallazgos sugieren que los patrones observados en la asignación de votos pueden aplicarse de manera universal y están influenciados más por dinámicas sociales fundamentales que por factores específicos del contexto.

Finalmente, utilizando un análisis estadístico se establece la forma de la distribución universal, la cual se ajusta a una distribución log-normal:

$$F(vQ/N) = \frac{1}{\sqrt{2\pi\sigma^2(vQ/N)}} \exp\left(-\frac{(\log(vQ/N) - \mu)^2}{2\sigma^2}\right), \quad (4)$$

donde los parámetros μ y σ definen la posición y la dispersión de la curva, respectivamente.

En la figura 1 podemos observar la distribución universal obtenida por Fortunato y Castellano junto a la obtenida al analizar los datos del Senado español. El error cuadrático medio calculado en la distribución experimental es del orden de 10^{-1} que en el peor de los casos corresponde a un error de aproximadamente el 3 %, por lo que podemos determinar que es un buen ajuste.

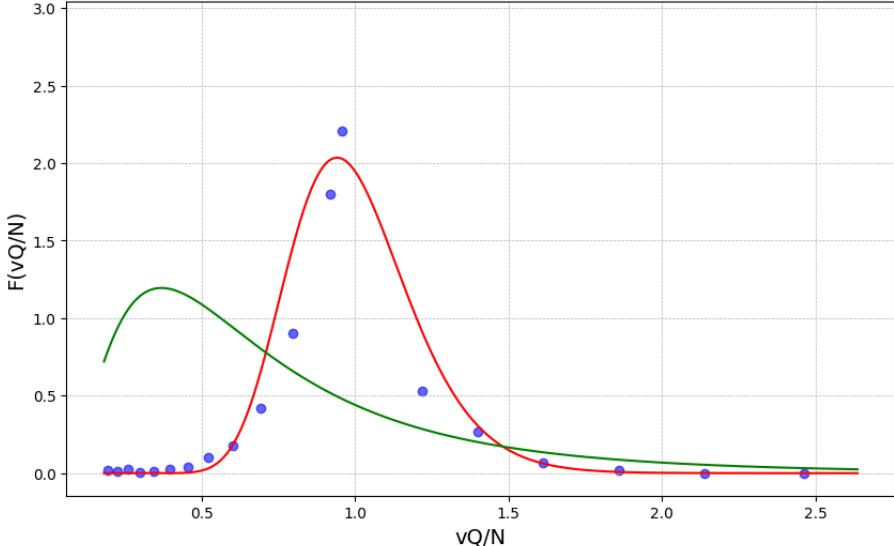


Figura 1: Distribución universal teórica en verde ($\mu = -0.54, \sigma^2 = 1.08$). Valores experimentales obtenidos en azul. Ajuste log-normal experimental en rojo ($\mu = 0.2, \sigma^2 = 0.98$).

Se observa que las distribuciones no solapan, esta situación sugiere que aquí no se sigue la regla general observada en naciones como Italia o Polonia y Finlandia. A pesar de ello, es relevante enfatizar que la distribución de datos en España también se ajusta al patrón log-normal compartido por el resto de los países estudiados. Este descubrimiento indica que a pesar de las diferencias destacadas presentes en ello, el proceso de transmisión de opinión en el ámbito político español comparte rasgos esenciales comunes a los de otras naciones, al menos en lo que respecta a la estructura fundamental de la distribución.

La principal disparidad entre las distribuciones reside en la ubicación del pico máximo; en el caso de España se encuentra notablemente más centrado en la unidad que en la curva universal. Este hecho sugiere que dentro del marco electoral español es menos común que un aspirante destaque de forma sobresaliente frente a sus compañeros de partido en comparación a lo observado en otros entornos políticos similares. Esta particularidad puede encontrarse relacionada directamente a una característica sociopolítica peculiar identificada dentro de España; los votantes parecen estar más influenciados por la afiliación partidista del candidato que por las cualidades individuales particulares que distinguen al propio candidato como persona única e independiente.

Esta tendencia se ajusta a las características del contexto político español, tal y como muestra este artículo [14], en el que los partidos políticos desempeñan un papel destacado en la movilización y comportamiento de los electores. En contraste con esto, en otros países como Italia, Polonia y Finlandia la figura del candidato individual puede tener una influencia más significativa en la decisión del votante. Lo que podría explicar el cambio en el punto máximo de la distribución general.

En este sentido se deduce que los procesos de transmisión de opinión y comportamiento electoral en España siguen patrones universales similares a los investigados en otras naciones. Por lo tanto, se refuerza la idea de que, en esencia, la dinámica electoral está regida por leyes universales que no dependen de variables concretas relacionadas con el entorno social. Si bien pueden haber variaciones regionales, la uniformidad en la distribu-

ción corrobora la utilidad de herramientas de sociofísica en el análisis de acontecimientos políticos en el contexto de España.

5.2. Modelo de Markov para el comportamiento electoral

Una vez se ha determinado la idoneidad de la sociofísica como herramienta para el estudio de la política en España, se intenta obtener un modelo físico que permita el estudio analítico de la dinámica electoral. En primer lugar, se propone la hipótesis de que el sistema electoral pueda tener memoria, es decir, que los resultados de elecciones pasadas podrían influir en las tendencias futuras. Esto sugiere que un modelo adecuado debería considerar los procesos electorales anteriores y sus resultados para capturar adecuadamente la dinámica del sistema. Con el objetivo de comprobar si esta hipótesis es cierta, se plantea un marco matemático basado en un modelo de Markov estocástico.

El modelo de Markov permite analizar la evolución temporal de una elección a la siguiente, de manera que construiremos una cadena de Markov en tiempo discreto donde cada paso temporal será cada proceso electoral. El modelo que se plantea consta de tres estados posibles:

- 1) El líder electo ha sido elegido por primera vez
- 2) El líder electo ha sido reelegido una vez
- 3) El líder electo ha sido reelegido dos o más veces

Para decidir el número de estados en la cadena, se han tomado los resultados obtenidos en [15] que indican que tres estados son suficientes para modelizar el comportamiento de este tipo de sistemas.

Matemáticamente podemos expresar este modelo: Sea $X\mathbf{m}_n \in \mathbb{R}^3$ el vector que representa las tres probabilidades asociadas a un liderazgo democrático:

$$\mathbf{x}_n = \begin{bmatrix} \text{Probabilidad de ser recién elegido,} \\ \text{Probabilidad de ser reelecto una vez,} \\ \text{Probabilidad de ser reelecto múltiples veces} \end{bmatrix}$$

Las probabilidades \mathbf{x}_n se actualizan en cada elección mediante el modelo de Markov basado en la elección previa, según la expresión:

$$\mathbf{x}_n = \mathbf{P}\mathbf{x}_{n-1}, \quad (5)$$

donde \mathbf{P} es una matriz de probabilidades que describe las probabilidades de transición entre los tres estados, dada por:

$$\mathbf{P} = \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} \\ p_{2,1} & p_{2,2} & p_{2,3} \\ 0 & p_{3,2} & p_{3,3} \end{bmatrix}$$

A continuación, se describe brevemente la estructura de la matriz de probabilidades y su interpretación. Al inicio de cada ciclo electoral, siempre se elige un líder democrático. En la elección siguiente, con probabilidad $p_{1,1}$, otro líder recién elegido puede surgir, o el líder actual es reelecto con probabilidad $p_{1,2}$, donde $p_{1,2} = (1 - p_{1,1})$, ya que estas son las dos únicas posibilidades. Una vez que un líder ha sido reelecto, existen tres posibles resultados en la elección siguiente: un nuevo líder es elegido (con probabilidad $p_{1,2}$), el líder es reelecto una vez más (con probabilidad $p_{2,2}$) o el líder es reelecto múltiples veces (con probabilidad $p_{3,2}$). Como estas son las únicas posibilidades, se cumple que $p_{3,2} = (1 - p_{1,2} - p_{2,2})$. La figura 2 representa cada transición

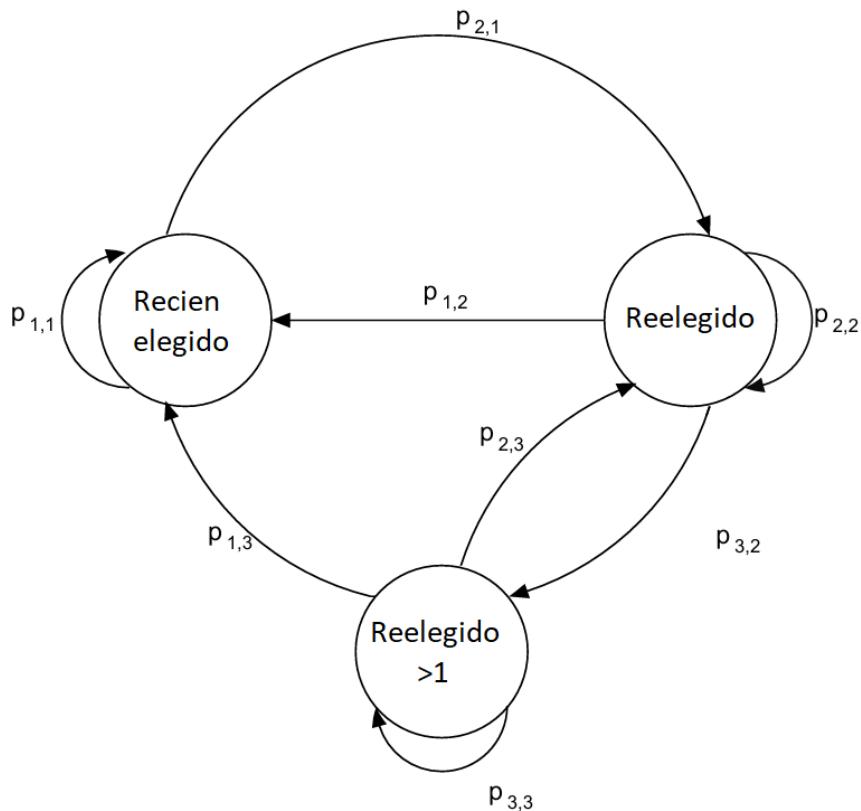


Figura 2: Descripción general del modelo de Markov mostrando los tres estados de un líder/partido y su correspondiente probabilidad de transición. La transición descrita por $p_{2,2}$ modela situaciones en las que un líder reelecto no puede finalizar su mandato, y un líder no elegido previamente asume el cargo. Si este nuevo líder es elegido en la elección siguiente, entonces la probabilidad $p_{2,3}$ captura este caso particular.

Para ajustar el modelo y calcular la matriz de transición de probabilidades \mathbf{P} , se realiza un análisis basado en los datos disponibles. Los elementos de la matriz \mathbf{P} se calculan mediante la siguiente ecuación:

$$P_{ij} = \frac{m_{ij}}{m_i}, \quad (6)$$

donde i y j representan uno de los tres estados, m_{ij} es el número total de líderes que pasaron del estado i al estado j según los datos, y m_i es el número total de líderes en el estado i .

Dado que la matriz \mathbf{P} es estocástica por columnas e irreducible, el teorema del punto fijo garantiza la existencia de un único punto fijo globalmente atractivo. Este punto fijo corresponde a la distribución de equilibrio estacionaria del modelo de Markov cuando el número de elecciones tiende a infinito ($n \rightarrow \infty$), y satisface la relación de estado estacionario:

$$\mathbf{x} = \mathbf{P}\mathbf{x}, \quad (7)$$

donde \mathbf{x} es el vector de probabilidades en estado estacionario. Resolver esta ecuación equivale a encontrar el autovector unitario de \mathbf{P} , normalizando las probabilidades para que sumen uno.

5.2.1. Resultados experimentales

Se procede a realizar los cálculos concretos en el contexto español. En la figura 3 se puede observar el diagrama de todas las presidencias y la transición de estado que se realiza en cada nueva legislatura.

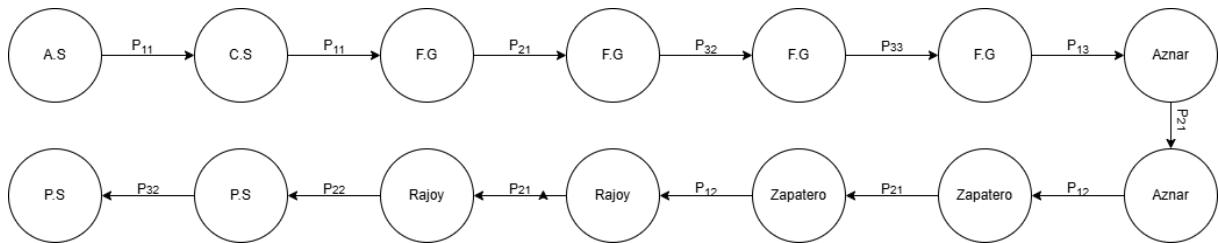


Figura 3: Diagrama de todas las presidencias y la transición de estado que se realiza en cada nueva legislatura

Con estos datos y usando la ecuación 6 podemos calcular la matriz de transición \mathbf{P} .

$$\mathbf{P} = \begin{bmatrix} 0.2 & 0.4 & 0.5 \\ 0.8 & 0.2 & 0 \\ 0 & 0.4 & 0.5 \end{bmatrix}$$

Ahora se calcula el autovector normalizado asociado al autovalor 1 a partir de la ecuación 7, ya que este representa el vector de probabilidades en estado estacionario:

$$\mathbf{x} = [0.38, 0.38, 0.24] \quad (8)$$

Esto implica que en el estado estacionario la probabilidad de estar en el estado 1 es del 38 %, la probabilidad de estar en el estado 2 es del 38 % y la probabilidad de estar en el estado 3 es del 24 % cuando el número de elecciones tiende a infinito ($n \rightarrow \infty$).

Prueba de bondad de ajuste Chi-cuadrado

Para evaluar la validez del modelo de Markov propuesto, se realiza una prueba estándar de bondad de ajuste Chi-cuadrado. Este test permite comparar las frecuencias observadas en los datos con las frecuencias esperadas generadas por el modelo, comprobando si las diferencias entre ambas son estadísticamente significativas.

La hipótesis nula (H_0) plantea que las frecuencias observadas no difieren significativamente de las frecuencias esperadas, lo que implicaría que el modelo describe adecuadamente los datos. La hipótesis alternativa (H_a) establece que existen diferencias significativas, lo que indicaría que el modelo no es adecuado.

El estadístico Chi-cuadrado (χ^2) se calcula mediante la fórmula:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}, \quad (9)$$

donde:

- O_i son las frecuencias observadas.
- E_i son las frecuencias esperadas generadas por el modelo.
- k es el número de categorías o estados considerados.

En este caso, se tienen las frecuencias observadas $\mathbf{O} = (5, 5, 2)$ y las frecuencias esperadas $\mathbf{E} = (4.56, 4.56, 2.88)$, correspondientes a los tres estados definidos en el modelo de Markov.

Sustituyendo estos valores en la ecuación 9, se obtiene:

$$\chi^2 = 0.354.$$

El valor calculado de χ^2 se compara con una distribución Chi-cuadrado con $k - 1 = 3 - 1 = 2$ grados de libertad. El p-valor asociado a este estadístico es $p = 0.838$.

Dado que el p-valor obtenido ($p = 0.838$) es significativamente mayor que un nivel de significancia típico, no se puede rechazar la hipótesis nula. Esto indica que las frecuencias observadas no difieren significativamente de las esperadas, lo cual sugiere que el modelo de Markov propuesto describe adecuadamente los datos.

El resultado de la prueba Chi-cuadrado respalda la validez del modelo de Markov para representar el comportamiento electoral analizado. Esto implica que las probabilidades de transición entre estados estimadas por el modelo son consistentes con las observaciones empíricas, proporcionando evidencia a favor de la hipótesis de que el sistema electoral tiene memoria.

Sabiendo que el sistema tiene memoria, se plantea la pregunta, ¿cuáles son las principales influencias en el comportamiento electoral? Como señalan algunos estudios [16, 17], una teoría que se ha planteado para justificar la conducta de los electores es que estos están influenciados por su entorno familiar, amigos o red social. Otro factor que podría explicar las tasas de reelección de los líderes es la opinión de su reputación entre los votantes. De acuerdo a investigaciones anteriores [18], los electores perciben al líder como exitoso o fracasado, lo cual puede tener diferentes niveles de impacto en su decisión electoral. Partiendo de estas dos teorías, se crea un modelo de votación que combine el elemento de memoria del sistema, planteado en el modelo de Markov de tres estados, junto a la influencia que ejercen la familia, amigos, redes sociales y la imagen del líder político en cada elector de manera individual.

5.3. Modelo de voto

En la sección anterior, se ha analizado el comportamiento electoral en España, mostrando patrones significativos en las decisiones de los votantes. Una posible explicación de las dinámicas observadas radica en la influencia de factores como la familia, los amigos y la reputación del líder sobre los votantes individuales. Para comprender completamente estos datos, se requiere un modelo que incorpore tanto las influencias sociales a nivel individual como los mecanismos que llevan a los votantes a decidir basándose en quién estaba previamente en el poder. Estos mecanismos operan en diferentes etapas del proceso electoral, por lo que es necesario construir un modelo descriptivo que refleje la estructura del sistema de votación en España.

El modelo de votación se divide en tres etapas o niveles: micro, meso y macro. En el nivel micro, los votantes individuales se agrupan en N_c distritos electorales de igual tamaño, con N_v votantes en cada distrito. Los votantes de cada distrito electoral se modelan utilizando una versión 2D modificada del modelo de Sznajd [19], que incorpora las influencias de los vecinos sociales más cercanos, los resultados de elecciones previas y el éxito (β) del líder previamente elegido.

Dado que España es un país con una fuerte tendencia bipartidista, se realizará una simplificación del sistema electoral considerando la existencia de solo dos partidos políticos.

En el nivel micro, el país se divide en distritos electorales donde cada ciudadano emite su voto eligiendo entre las dos opciones disponibles (denotadas por -1 y $+1$). Los votos se contabilizan y el partido que obtiene la mayoría en un distrito gana la representación de ese territorio, asignándosele un escaño. Este proceso refleja un sistema mayoritario en el que el candidato más votado en cada circunscripción obtiene la totalidad de la representación.

En el nivel meso, los representantes elegidos en los distintos distritos se agrupan según su afiliación partidista, conformando los dos bloques parlamentarios principales. Dentro de cada partido, los diputados votan internamente para seleccionar a su líder, quien actuará como principal figura del partido en el proceso de formación de gobierno. Este mecanismo refleja la disciplina partidaria y el proceso de elección interna de liderazgo común en el

sistema político español.

Finalmente, en el nivel macro, el líder del partido con mayor representación parlamentaria es propuesto como candidato a la presidencia del gobierno. Si su partido obtiene la mayoría absoluta de escaños, se le otorga directamente el cargo. β representa el desempeño del mandato del presidente del gobierno que se retroalimenta al nivel micro, influyendo en las siguientes decisiones de los votantes.

El partido político que lidera antes de las elecciones mantiene al líder democrático solo si su mandato es considerado exitoso. El éxito de un mandato se evalúa de manera aleatoria, asignándole una puntuación que determina si fue exitoso o no exitoso. En caso de que el mandato sea considerado no exitoso, el partido realiza una votación para elegir a un nuevo líder. Cada partido elige a su líder basándose en una puntuación de liderazgo asignada a cada parlamentario de su partido, tomada de una distribución normal.

En el proceso de elección de liderazgo, se enfrentan los dos legisladores con las puntuaciones de liderazgo más altas. Dado que la mayoría de los parlamentarios elegidos tienen escaños seguros, suelen tener posturas políticas firmes, por lo tanto, buscarán un líder de partido que también represente estas posturas y viceversa. En consecuencia, todos los demás legisladores votan por el candidato que tiene una orientación política más cercana a la suya.

El líder del partido con el mayor número de votos de legisladores es seleccionado como líder en el nivel macro. Su reputación, que puede ser exitosa o no exitosa, se denota como β_{exito} y β_{noexito} , respectivamente. Esta reputación se determina aleatoriamente comparando si un número generado es menor o mayor que la puntuación de liderazgo del líder, influyendo así en las decisiones de los votantes en las siguientes elecciones.

Solo se consideran dos tipos de reputación, ya que investigaciones previas han demostrado que los votantes solo perciben el mandato del líder como exitoso o no exitoso [18].

Tensión T

En el nivel micro, el comportamiento de los votantes individuales se modela utilizando una versión modificada del modelo 2D de Sznajd. Este modelo introduce un parámetro de tensión, T , que describe la interacción entre la preferencia partidista de un votante y las decisiones de sus vecinos más cercanos. Esta tensión es un claro reflejo del Hamiltoniano que se encuentra en un modelo clásico de Ising. La tensión para cada votante i se define como:

$$T_i = -\frac{J}{2} s_i \sum_{\langle il \rangle} s_l - \mu s_i, \quad (10)$$

donde $s_i = \pm 1$ representa la elección del votante i entre los dos partidos políticos disponibles. El parámetro J describe la intensidad de la influencia social entre los vecinos más cercanos $\langle il \rangle$, restringida al rango de influencia de l vecinos. Si los vecinos comparten

la misma preferencia partidista, la tensión entre ellos es baja; por el contrario, las diferencias en las preferencias generan una tensión mayor. Por su parte, μ representa la influencia externa ejercida por el líder del partido político en el poder, incluyendo su reputación y la memoria electoral del distrito. Un claro reflejo del momento magnético que encontramos en el modelo clásico de Ising, que determinaron qué cantidad interactuarán las partículas en el campo externo.

Influencia externa μ

La influencia externa, μ , refleja el comportamiento previo del distrito electoral, considerándose como:

$$\mu = \text{sign}(\bar{s}^{\text{prev}}) - \beta \text{ sign} \left(\frac{1}{N_c} \sum_{j=1}^{N_c} \bar{s}_j^{\text{prev}} \right), \quad (11)$$

donde \bar{s}^{prev} es el voto promedio del distrito en las elecciones anteriores, y $\frac{1}{N_c} \sum_{j=1}^{N_c} \bar{s}_j^{\text{prev}}$ representa el promedio del voto en todos los distritos. El término μ , por lo tanto, tiene en cuenta si el distrito votó de acuerdo a la mayoría del país o no. En los casos en los que un distrito mantuvo una preferencia partidista coherente con el resultado nacional, la tensión en el sistema disminuye, reflejando estabilidad. Por el contrario, un resultado disonante aumenta la tensión y empuja a un cambio de voto para las siguientes elecciones.

El parámetro β introduce el desempeño del líder previo en el modelo, donde:

$$\beta = \begin{cases} \beta_{\text{exito}} & \text{si el mandato del líder fue exitoso,} \\ \beta_{\text{noexito}} & \text{si el mandato del líder fue no exitoso.} \end{cases}$$

Tanto β_{exito} como β_{noexito} tienen un efecto negativo sobre los votantes si se consideran valores positivos. Para determinar si un mandato fue exitoso o no, se genera un número aleatorio de una distribución uniforme y se compara con la puntuación de liderazgo del líder. Si el número aleatorio es menor que la puntuación de liderazgo, el mandato se considera exitoso; de lo contrario, no exitoso. Este mecanismo conecta directamente el desempeño del líder con las decisiones de los votantes en las siguientes elecciones.

Para determinar la probabilidad de que un distrito electoral esté en una configuración particular, es decir, haya votado de cierta manera, se utiliza la distribución de Boltzmann. La probabilidad se define como:

$$P(s) = \frac{e^{-\sum_{i=1}^{N_v} T_i}}{\sum_{s=\pm 1} e^{-T_i}}, \quad (12)$$

donde N_v es el número de votantes en un distrito electoral. En este modelo, tanto la temperatura como la constante de Boltzmann se han fijado en la unidad. La distribución de Boltzmann se elige porque la configuración más probable de los votantes se obtiene

maximizando la ecuación (11), lo que equivale a minimizar la tensión total ($\sum_{i=1}^{N_v} T_i$) para cada distrito electoral. Corresponde análogamente a reducir la energía (el Hamiltoniano) en un modelo clásico de Ising.

Para las dos primeras elecciones, la influencia externa μ se establece en cero, ya que en las primeras elecciones no existe un líder establecido. A partir de la tercera elección, la influencia externa μ comienza a tener un efecto en los votantes.

5.3.1. Análisis inicial y calibración del modelo

Antes de realizar simulaciones del modelo de votación, es necesario acotar el espacio de parámetros a explorar, ya que no está claro qué valores deberían asignarse al parámetro de influencia social J o a las influencias externas del líder, β_{exito} y β_{noexito} . Para ver la influencia de cada uno de estos parámetros en el promedio de voto \bar{s} se realiza una aproximación de campo medio tamando una red bidimensional periódica (2D), donde los votantes están influenciados por sus cuatro vecinos más cercanos: arriba, abajo, izquierda y derecha denotado $z = 4$.

Se identifican dos comportamientos generales esperados para el modelo. Para un acoplamiento J pequeño, se espera que la tensión de cada votante sea gobernada principalmente por la influencia del líder (μ). En este caso, partiendo de una distribución inicial uniforme de preferencias entre los votantes, el promedio de votos en un distrito electoral (\bar{s}) será aproximadamente cero. Aunque habrá fluctuaciones aleatorias debido a la ecuación 11, en promedio, esta influencia externa tenderá a cero, lo que hará que la media de votos tienda a cero.

Por otro lado, para valores de J grandes, la influencia del líder es despreciable, ya que los votantes se agrupan con otros con preferencias similares. En este caso, se espera que el líder se mantenga en el poder.

Influencia de J y μ

Para estudiar analíticamente cómo afecta la influencia del líder al voto para uno o varios distritos, se considera una aproximación de campo medio, es decir, se considera un único distrito y se aproxima la interacción de los vecinos cercanos por su voto medio \bar{s} . Se obtiene la siguiente tensión aproximada para cada votante:

$$T(s_i) \approx -\left(\frac{Jz}{2}\bar{s} + \mu\right)s_i = T_{\text{aprox}} \cdot s_i, \quad (13)$$

Como la tensión de cada votante está desacoplada, la probabilidad de que un votante s_i tenga una preferencia específica se expresa como:

$$p(s_i) = \frac{e^{-T_{\text{aprox}} s_i}}{e^{T_{\text{aprox}}} + e^{-T_{\text{aprox}}}} \quad (14)$$

y para mantener consistencia con la aproximación de campo medio, el promedio esperado del valor del voto (\bar{s}) calculado a partir de esta distribución debe coincidir con el promedio de votos:

$$\bar{s} = \sum_{s_i=\pm 1} p(s_i)s_i = \frac{e^{-T_{\text{aprox}}} - e^{T_{\text{aprox}}}}{e^{T_{\text{aprox}}} + e^{-T_{\text{aprox}}}} = \tanh(T_{\text{aprox}}) \quad (15)$$

De esta forma, se obtiene la ecuación de campo medio que describe el comportamiento promedio del modelo:

$$\bar{s} = \tanh\left(\frac{Jz}{2}\bar{s} + \mu\right). \quad (16)$$

Esta ecuación es trascendental, lo que implica que no tiene solución analítica. Para poder extraer información se buscan soluciones de forma gráfica para distintas combinaciones de los parámetros: $Jz \gg 1$, $Jz \ll 1$ y $\mu > 0$, $\mu < 0$.

Si $Jz \gg 1$ (una influencia social alta) obtenemos los resultados de la figura 8

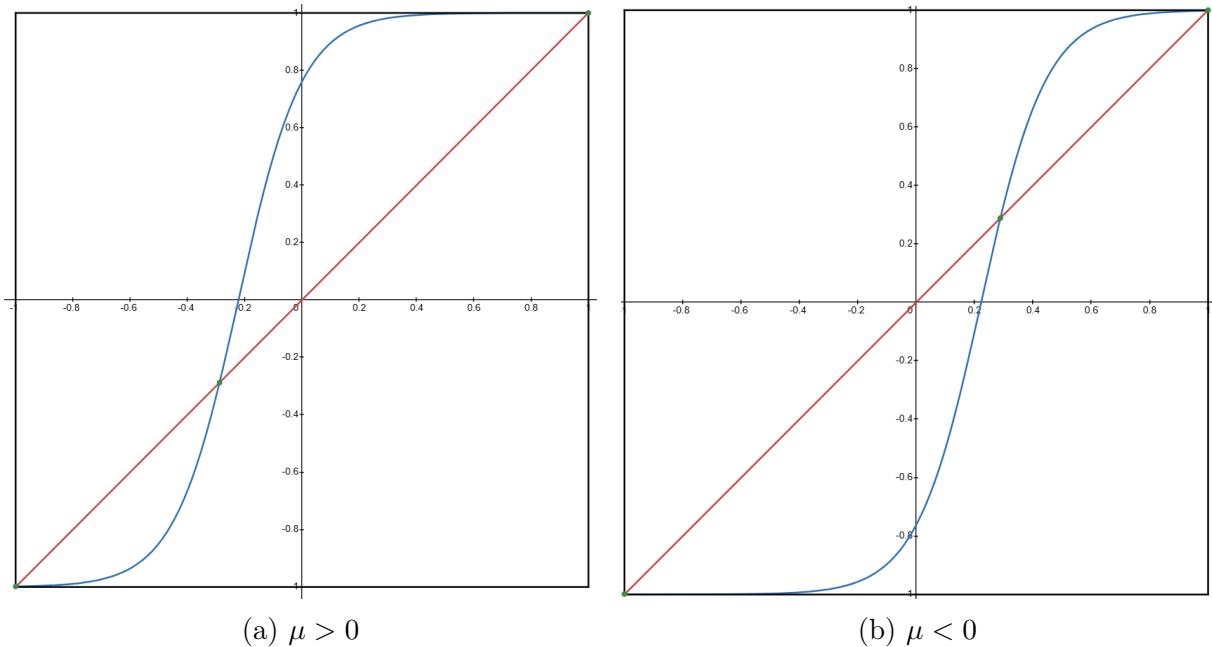


Figura 4: La curva roja/azul representa el lado izquierdo/derecho de la ecuación (16) cuando $Jz \gg 1$. Las intersecciones se marcan en verde

Si $Jz \gg 1$ (una influencia social baja) obtenemos los resultados de la figura 8

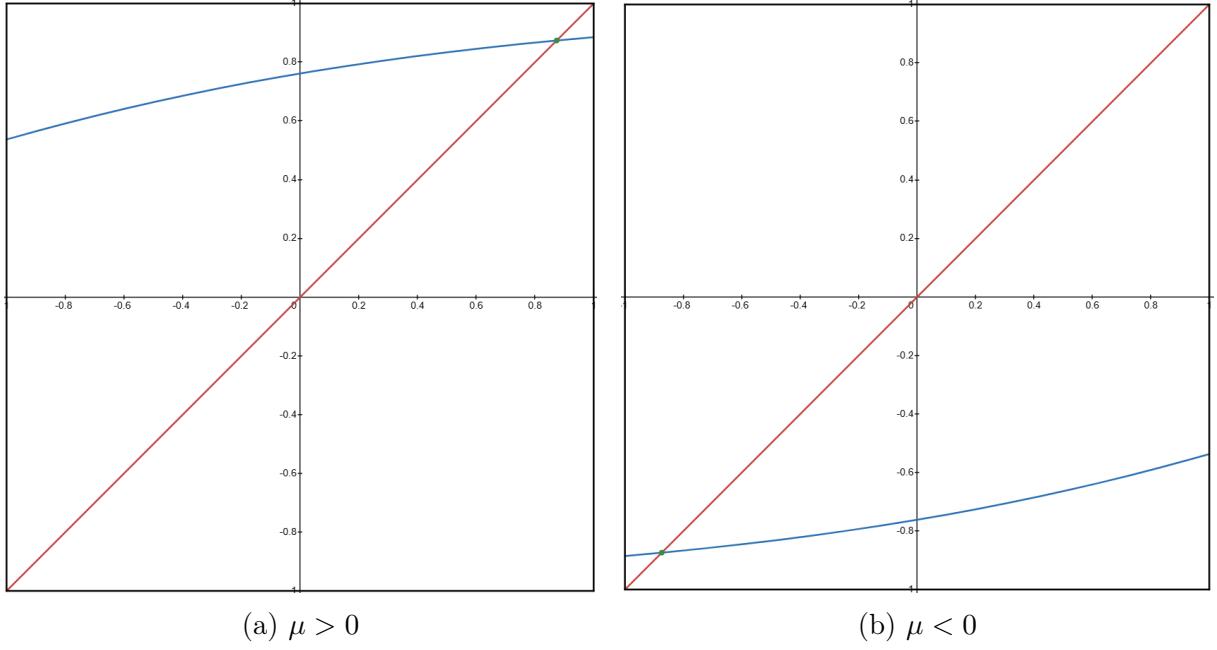


Figura 5: La curva roja/azul representa el lado izquierdo/derecho de la ecuación (16) cuando $Jz \ll 1$. Las intersecciones se marcan en verde

Se observa que el efecto de la influencia externa μ conduce a un \bar{s} del mismo signo que μ . Si además β toma valores pequeños, aumenta la preferencia por votar al mismo resultado de las elecciones anteriores, ya que de la ecuación 11 $\mu \approx \text{sign}(\bar{s}^{\text{prev}})$ y se acaba de ver que \bar{s} tendrá el mismo signo que μ .

La solución de esta ecuación revela que la influencia externa μ genera un \bar{s} con el mismo signo que μ . Por lo tanto, cuando $\mu > 0$, el promedio se desplaza hacia valores positivos; y cuando $\mu < 0$, hacia valores negativos. Esto puede observarse en las soluciones gráficas de la ecuación para diferentes valores de μ , Jz , y pequeñas perturbaciones en β . Por otro lado, a medida que el acoplamiento crece, las soluciones tienden a los valores extremos (± 1), reflejando una mayor polarización entre los votantes.

Influencia de β

Para analizar la influencia del parámetro β se asume el caso en el que todos los distritos electorales votan con el mismo promedio \bar{s} y por lo tanto $\frac{1}{N_c} \sum_{j=1}^{N_c} \bar{s}_j^{\text{prev}} = 1$. Esto implica que la influencia externa puede escribirse:

$$\mu = \text{sign}(\bar{s})(1 - \beta)$$

Para trabajar analíticamente se aproxima por una función suave:

$$\mu = \text{sign}(\bar{s})(1 - \beta) \approx \tanh(\alpha \bar{s})(1 - \beta), \quad \alpha \in \mathbb{R}^+, \quad (17)$$

con $\alpha \gg 1$.

Para obtener soluciones de la ecuación 16 no triviales ($\bar{s} \neq 0$), el sistema no puede ser estable alrededor de esta solución. Esto implica que la derivada de la parte derecha de la ecuación en el 0 sea mayor que la unidad:

$$\frac{Jz}{2} + \alpha(1 - \beta) > 1. \quad (18)$$

Si $0 < \beta < 1$ (siendo α lo suficientemente grande) el sistema siempre cumplirá la condición sin importar el valor de J . Por otro lado, si $J > 1$ nunca se cumple la condición. Por ello, no se espera encontrar un valor crítico de J en las simulaciones.

Siempre que $0 < \beta < 1$ y α sean lo suficientemente grandes, el modelo cumplirá esta condición independientemente del valor de la constante de acoplamiento J . Por el contrario, si $\beta > 1$, entonces, para α suficientemente grande, nunca se espera cumplir esta condición. Por lo tanto, no se espera observar un valor crítico de la constante de acoplamiento en las simulaciones numéricas, pero la constante de acoplamiento aún puede tener un efecto en las tasas de reelección.

Aunque no parece haber un parámetro crítico, se observa un efecto claro al variar el valor de la constante de acoplamiento J y los parámetros de reputación β en la aproximación de campo medio. En particular, un parámetro de reputación positiva $\beta > 0$ tiene un impacto negativo en los votantes, reduciendo así la probabilidad de que un líder sea reelegido, mientras que un $\beta < 0$ aumenta dicha probabilidad. También se observa que, a medida que la constante de acoplamiento aumenta, las soluciones no triviales de la ecuación 16 tienden a ± 1 . Por lo tanto, para que el modelo reproduzca los diferentes comportamientos de votación observados, es necesario variar ambos parámetros.

5.3.2. Resultados experimentales

Para realizar las simulaciones, se elige $N_c = 52$, ya que en España existen 52 circunscripciones electorales, y $N_v = 100$. Los agentes se colocan en enrejados de tamaño $N_v \cdot N_v$ y cada uno se verá influenciado por sus cuatro vecinos más próximos: arriba, abajo, izquierda y derecha con condiciones de frontera periódicas.

Para calcular los porcentajes de reelección de líderes, se simulan 30 elecciones y se repite el proceso 5 veces para obtener la media y reducir el error estadístico. Como se calcula la media de 5 simulaciones, se puede obtener el error estándar. El mayor error cometido: $E_{max} = 0.03$ que es aproximadamente un error de 3 %.

En primer lugar, se quiere conocer experimentalmente los efectos de los parámetros $J, \beta_{exito}, \beta_{noexito}$. Para ello, tal y como se puede observar en la figura 6, se fija el valor $\beta_{exito} = 0$ y se realizan las simulaciones para cada pareja de valores $J, \beta_{noexito}$ en el intervalo $[0, 1]$. Se observa que cuanto menor sea $\beta_{noexito}$, menor es la tasa de reelección; esto tiene sentido ya que una $\beta_{noexito}$ alta implica que la probabilidad de reelegir líderes más exitosos es más alta, mientras que la probabilidad de reelegir líderes menos exitosos es más baja.

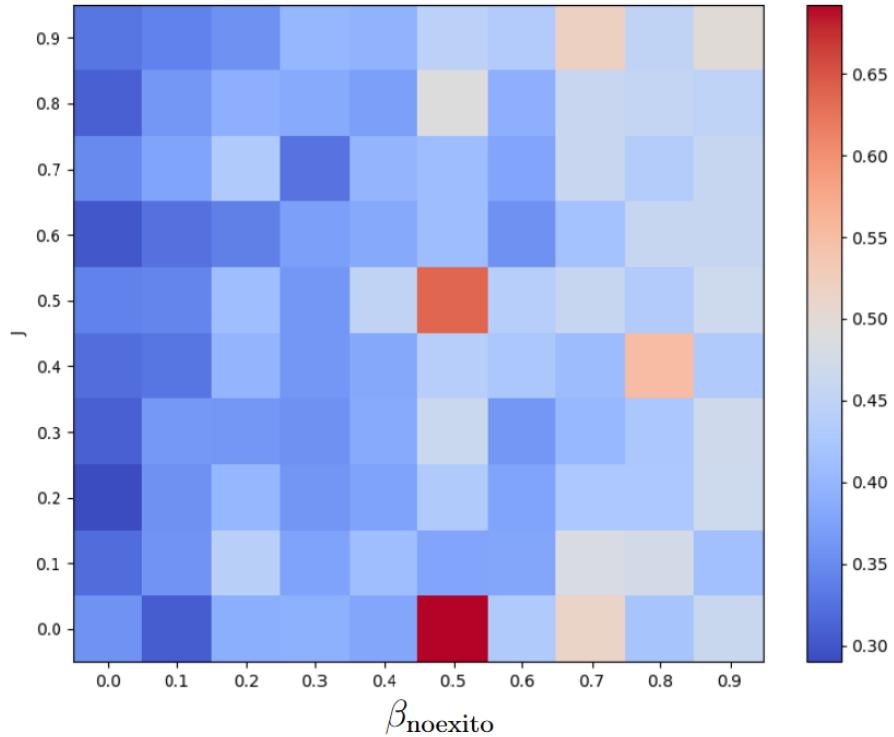


Figura 6: Tasas promedio de reelección de líderes, con $\beta_{exit} = 0$. Los colores indican la tasa de reelección de un líder.

Por otro lado, podemos realizar las mismas simulaciones anulando esta vez $\beta_{noexit} = 0$. Se obtiene la figura 7. Se observa que la tasa de reelección es muy baja para todos los valores de J, β_{noexit} . Por tanto, se deduce que β_{noexit} tiene un impacto mucho mayor en los votantes de lo que tiene β_{exit} , ya que cuando la primera se anula, la tasa de reelección cae drásticamente.

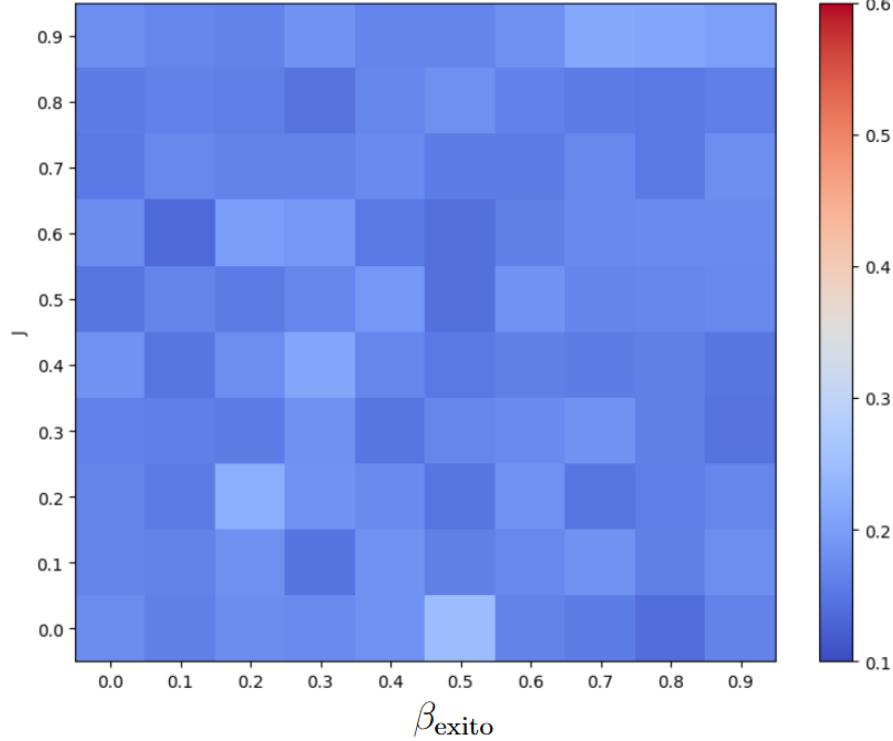


Figura 7: Tasas promedio de reelección de líderes, con $\beta_{\text{noexito}} = 0$. Los colores indican la tasa de reelección de un líder.

Por último, se prueba a fijar $J = \{0, 1\}$ y se observa el comportamiento del sistema en función de los otros dos parámetros $\beta_{\text{exito}}, \beta_{\text{noexito}}$. Se comprueba que si la constante de acoplamiento J se anula, los votos tienden a ser más aleatorios, lo que implica que el voto medio se situará alrededor del 0 y, por lo tanto, la tasa de reelección es muy baja para cualquier par de valores de β . Por otro lado, cuando $J = 1$ valores pequeños de β implican tasas de reelección más altas.

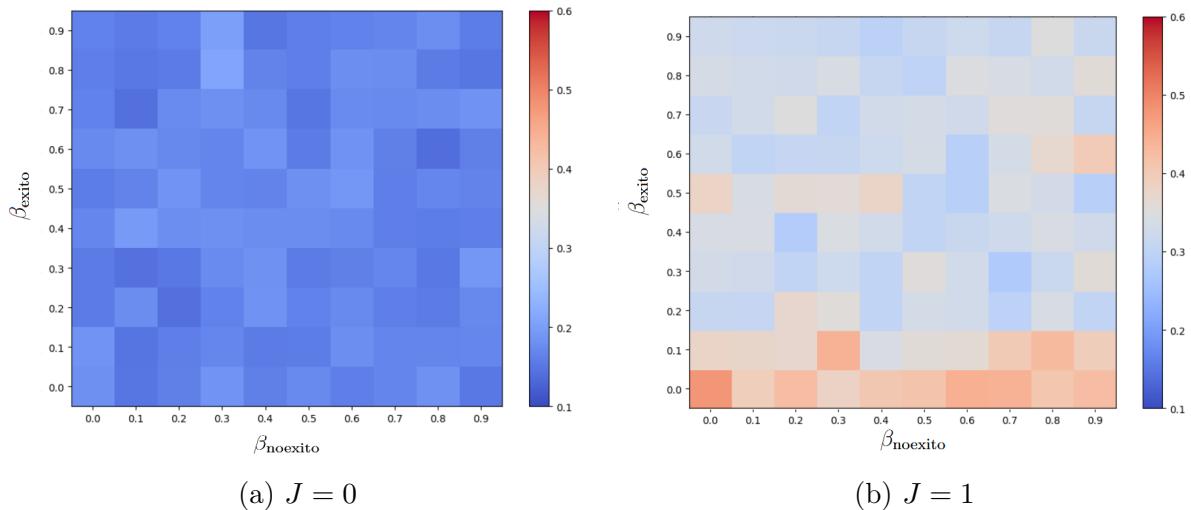


Figura 8: Tasas promedio de reelección de líderes, con $J = \{0, 1\}$. Los colores indican la tasa de reelección de un líder.

Una vez analizado el modelo analíticamente y computacionalmente, se procede a intentar ajustar los parámetros para las tasas de reelección obtenidas en la sección 5.2 para España, que corresponden a una tasa de reelección del 58 %. Utilizando las figuras obtenidas en esta sección, se ajusta el modelo mediante prueba y error, obteniendo los siguientes resultados: $J = 0.45$, $\beta_{exito} = 0.09$ y $\beta_{noexito} = 0.4$. Ahora, se dispone de un modelo físico que refleja el comportamiento de la sociedad española en la decisión de voto. De esta manera, es posible parametrizar la importancia de cada variable del modelo en este ámbito sociopolítico y potencialmente compararlo con otros.

6. Conclusiones

En este trabajo, se ha realizado un estudio sociofísico de la política en el ámbito español. En primer lugar, se ha comprobado si el campo de estudio era adecuado para aplicar modelos sociofísicos, se ha descubierto que la dinámica de la transferencia de opinión sigue patrones universales, que en España se comparten, aunque con ciertas diferencias debido a las particularidades sociopolíticas y del sistema electoral. Por lo que se ha concluido que el uso de técnicas sociofísicas está justificado en este ámbito.

Por otro lado, se ha propuesto la hipótesis de que la dinámica electoral era un sistema con memoria, de manera que para modelizar el comportamiento de los electores era necesario tener en cuenta los resultados del proceso electoral anterior. Para demostrar esta hipótesis, se ha planteado un modelo de Markov de 3 estados y se ha calculado el estado estacionario del sistema. De esta manera, se ha logrado obtener la probabilidad de reelección del presidente del gobierno en este estado estacionario, que representa el sistema cuando el número de elecciones tiende a infinito. Se ha comprobado que las probabilidades de reelección obtenidas del modelo son estadísticamente relevantes, por tanto, el modelo es capaz de reflejar el sistema real y eso implica que la hipótesis inicial de que el sistema tenía memoria no se puede descartar.

Por último, se ha planteado un modelo sociofísico para intentar reproducir esta dinámica electoral de manera analítica. Las interacciones entre agentes en el modelo se han representado mediante un modelo de Ising adaptado, donde el hamiltoniano del sistema tiene en cuenta la interacción entre los agentes debido a las diferencias de opinión entre un individuo y sus vecinos. Además, se incluye lo descubierto anteriormente añadiendo una interacción externa que depende de los resultados electorales previos, añadiendo este comportamiento con memoria. Se ha estudiado analíticamente el impacto de los distintos parámetros, encontrando distintos valores críticos. Para terminar, se han realizado distintas simulaciones, permitiendo ver computacionalmente el comportamiento del sistema en casos límite. Con esta información, se ha ajustado el modelo para cumplir con los datos estadísticos reales de las tasas de reelección, con los siguientes valores para cada parámetro: $J = 0.45$, $\beta_{exito} = 0.09$ y $\beta_{noexito} = 0.4$, esto permite simular un entorno con dinámicas electorales equivalentes a las observadas en el ámbito español.

Futuras investigaciones

Una vez adquirido este modelo específico, se plantean futuras investigaciones que podrían resultar de gran interés tanto en el ámbito académico como en el práctico. En primer lugar, se sugiere contrastar los resultados obtenidos en España respecto a otros modelos electorales para examinar cómo las diferencias sociopolíticas y las particularidades de los sistemas de representación influyen en las dinámicas de votación. Este estudio comparativo posibilitaría la exploración de patrones comunes y discrepancias según las características distintivas de cada contexto.

Además se sugiere utilizar el modelo creado para estudiar cómo diversos sucesos sociales como crisis económicas, movimientos sociales o modificaciones en la legislación podrían afectar los resultados electorales al introducir alteraciones externas en el sistema. Estas si-

mulaciones facilitarían examinar situaciones hipotéticas y valorar la resistencia del sistema electoral ante tales sucesos.

Finalmente sería interesante ampliar el modelo para abarcar dinámicas temporales más elaboradas; por ejemplo, incorporando ciclos electivos o variaciones en las normativas electorales del sistema político. De esta manera se podría investigar cómo la transformación de las preferencias políticas en el transcurso del tiempo - impulsada por factores socio-culturales o tecnológicos - podría afectar la estabilidad del sistema o dar lugar a nuevos comportamientos innovadores.

En resumen, este estudio ha evidenciado que las técnicas sociofísicas son un recurso válido para analizar la dinámica electoral en el contexto español; los hallazgos obtenidos abren nuevas vías de investigación que podrían enriquecer nuestra comprensión de los procesos sociopolíticos y su modelado matemático. Estas futuras investigaciones no solo podrían impulsar desarrollos teóricos en la intersección de la física y las ciencias sociales, sino también brindar herramientas prácticas para mejorar y fortalecer sistemas electorales más representativos y resilientes.

Referencias

- [1] S. Galam and Y. Gefen (Feigenblat), Y. Shapir, “Sociophysics: A new approach of sociological collective behaviour: I. mean-behaviour description of a strike,” *Journal of Mathematical Sociology*, vol. 9, no. 1, pp. 1–13, 1982.
- [2] S. GALAM, “Sociophysics: A review of galam models,” *International Journal of Modern Physics C*, vol. 19, no. 03, pp. 409–440, 2008.
- [3] S. Galam, “The trump phenomenon: An explanation from sociophysics,” *International Journal of Modern Physics B*, vol. 31, no. 10, p. 1742015, 2017.
- [4] R. N. Mantegna and H. E. Stanley, “Scaling behavior in the dynamics of an economic index,” *Nature*, vol. 376, pp. 46–49, 1995.
- [5] E. Ising, “Beitrag zur theorie des ferromagnetismus [contribution to the theory of ferromagnetism],” *Zeitschrift für Physik*, vol. 31, no. 1, pp. 253–258, 1925.
- [6] K. Sznajd-Weron and J. Sznajd, “Opinion evolution in closed community,” *International Journal of Modern Physics C*, vol. 11, no. 6, pp. 1157–1165, 2000.
- [7] P. S. Foundation, “Python programming language – official website,” 2024, accedido: 2025-01-01. [Online]. Available: <https://www.python.org>
- [8] T. N. Community, “Numpy: The fundamental package for scientific computing with python,” 2024, accedido: 2025-01-09. [Online]. Available: <https://numpy.org>
- [9] T. P. D. Team, “Pandas: Python data analysis library,” 2024, accedido: 2025-01-09. [Online]. Available: <https://pandas.pydata.org>
- [10] T. M. D. Team, “Matplotlib: Visualization with python,” 2024, accedido: 2025-01-09. [Online]. Available: <https://matplotlib.org>
- [11] S. de España, “Resultados electorales,” 2023, accedido: 13 de noviembre de 2024. [Online]. Available: <https://www.senado.es/web/composicionorganizacion/senadores/sistemaelectoralresultados/resultadoselectorales/index.html>
- [12] E. Ministerio del Interior, “Resultados electorales - elecciones celebradas,” 2023, accedido: 13 de noviembre de 2024. [Online]. Available: <https://infoelectoral.interior.gob.es/es/elecciones-celebradas/resultados-electorales/>
- [13] S. Fortunato and C. Castellano, “Scaling and universality in proportional elections,” *Phys. Rev. Lett.*, vol. 99, p. 138701, Sep 2007.
- [14] J. Montabes Pereira and C. Ortega Villodres, “El voto limitado en las elecciones al senado español: estrategias de nominación y rendimientos partidistas en las elecciones de marzo de 2000,” *Revista Española de Ciencia Política*, no. 7, pp. 103–130, Octubre 2002.
- [15] J. Rowden, D. J. Lloyd, and N. Gilbert, “A model of political voting behaviours across different countries,” *Physica A: Statistical Mechanics and its Applications*, vol. 413, pp. 609–625, 2014.

- [16] P. A. Beck and M. K. Jennings, “Family traditions, political periods, and the development of partisan orientations,” *The Journal of Politics*, vol. 53, no. 3, pp. 742–763, 1991.
- [17] G. M. Pomper, “The impact of the american voter on political science,” *Political Science Quarterly*, vol. 93, no. 4, pp. 617–628, 1978, winter 1978.
- [18] T. Capelos, “Feeling the issue: How citizens’ affective reactions and leadership perceptions shape policy evaluations,” *Journal of Political Marketing*, 2010.
- [19] S. Galam, “Application of statistical physics to politics,” *Physica A: Statistical Mechanics and its Applications*, vol. 274, no. 1, pp. 132–139, 1999.