

Estado de la publicación: El preprint no ha sido enviado para publicación

"Dime con quién andas y te diré quién eres": Un enfoque bayesiano para la inferencia de perfiles de liderazgo en sistemas sociales

Alejandro Buitrago-Ramírez, Alejandro Roldán-Correa

<https://doi.org/10.1590/SciELOPreprints.16557>

Enviado en: 2026-06-15

Postado en: 2026-06-16 (versión 1)

(AAAA-MM-DD)

“Dime con quién andas y te diré quién eres”:
Un enfoque bayesiano para la inferencia de perfiles de liderazgo
en sistemas sociales

*“Tell me who you hang out with, and I will tell you who you are”:
A Bayesian approach to inferring leadership profiles in social systems*

Alejandro Buitrago-Ramírez
Instituto de Matemáticas
Universidad de Antioquia, Colombia.
alejandrobuitrago@udea.edu.co
<https://orcid.org/0009-0004-2457-0718>

Alejandro Roldán-Correa
Instituto de Matemáticas
Universidad de Antioquia, Colombia.
alejandroroldan@udea.edu.co
<https://orcid.org/0000-0003-1817-2583>

Resumen

El proverbio popular “dime con quién andas y te diré quién eres” captura una acertada intuición estadística acerca de la homofilia en las estructuras sociales. Este artículo formaliza matemáticamente dicho refrán al proponer un modelo probabilístico para la clasificación de líderes sociales mediante el análisis de variables latentes. Utilizando un enfoque bayesiano, se demuestra cómo la composición de una muestra de seguidores permite inferir la probabilidad de que un líder pertenezca a una categoría ética o ideológica específica. El modelo ofrece un marco metodológico para comprender la actualización de creencias bajo incertidumbre y el impacto del ruido del sistema en la fiabilidad de las redes de seguidores.

Palabras clave: Inferencia Bayesiana, Clases Latentes, Homofilia, Teorema de Doob.

Abstract

The popular proverb “tell me who you hang out with, and I will tell you who you are” captures an accurate statistical intuition regarding homophily in social structures. This paper mathematically formalizes this adage by proposing a probabilistic model for the classification of social leaders through latent variable analysis. Using a Bayesian approach, we demonstrate how the composition of a sample of followers enables the inference of the probability that a leader belongs to a specific ethical or ideological category. The model provides a methodological framework for understanding belief updating under uncertainty and the impact of system noise on the reliability of follower networks.

Keywords: Bayesian inference, Latent classes, Homophily, Doob’s theorem.

1. Introducción

En la teoría de la elección social y las ciencias sociales cuantitativas, la calidad de un líder a menudo se trata como una variable no observable directamente. Sin embargo, los seguidores revelan información a través de su afiliación [1]. El presente trabajo propone una estructura matemática fundamentada en el Análisis de Clases Latentes [4] para decodificar este comportamiento, asumiendo que los individuos tienden a alinearse con líderes que reflejan su propio perfil ético o ideológico. Este fenómeno está extensamente documentado en la sociología bajo el principio de homofilia [6] y guarda relación con la dinámica de polarización estudiada en los modelos de votante dentro de la teoría de sistemas de partículas interactuantes [3, 5].

2. Especificación del Modelo

Para formalizar la estructura probabilística, definimos tres variables aleatorias:

1. **Variable Latente del Líder (Z):** Sea $Z \in \{0, 1\}$ el estado latente del líder, donde $Z = 1$ representa un líder “bueno” y $Z = 0$ un líder “malo”.¹
2. **Perfil del Individuo (X):** Cada individuo en la población se caracteriza por una variable $X \in \{0, 1\}$, donde $X = 1$ indica que el individuo es “bueno” y $X = 0$ que es “malo”.
3. **Indicadora de Apoyo (Y):** Sea $Y \in \{0, 1\}$ la variable que registra la afiliación del individuo, donde $Y = 1$ significa que la persona decide apoyar al líder (ya sea afiliándose a su red o mostrando respaldo público), y $Y = 0$ representa la ausencia de dicho apoyo.

La formación de la red de seguidores está determinada por cuatro parámetros probabilísticos que dividimos en dos categorías de comportamiento:

I. Alineación Congruente (Homofilia)

Representan la tendencia natural de los individuos a asociarse con líderes que comparten su propia naturaleza:

- $\beta = P(Y = 1 | X = 1, Z = 1)$: *El seguidor afín*. Probabilidad de que un individuo “bueno” identifique correctamente y apoye a un líder “bueno”.
- $\alpha = P(Y = 1 | X = 0, Z = 0)$: *La solidaridad en la corrupción*. Probabilidad de que un individuo “malo” simpatice y se alinee con un líder “malo”.

II. Alineación Errónea (Ruido del Sistema)

Las redes sociales y de liderazgo sufren perturbaciones por desinformación u oportunismo. Capturamos estos “falsos positivos” mediante:

- $\delta = P(Y = 1 | X = 1, Z = 0)$: *El seguidor engañado*. Probabilidad de que una persona “buena” sea manipulada, confundida por la propaganda o se equivoque, apoyando a un líder “malo”.
- $\gamma = P(Y = 1 | X = 0, Z = 1)$: *El seguidor oportunista*. Probabilidad de que una persona “mala” apoye a un líder “bueno”, ya sea por cálculo estratégico, presión social o para pasar desapercibido en la corriente mayoritaria.

3. Inferencia Bayesiana

Supongamos que observamos la base de seguidores de un líder específico. Nuestro objetivo es calcular la probabilidad a posteriori de que el líder sea bueno ($Z = 1$) dado que conocemos la composición de sus seguidores.

Sea $\pi_g = P(X = 1)$ la proporción a priori de individuos “buenos” en la población. Primero, calculamos la probabilidad de que un seguidor seleccionado al azar ($Y = 1$) pertenezca a la clase $X = 1$, condicionado a que $Z = 1$. Denotaremos esta probabilidad como q_1 . Aplicando el Teorema de Bayes obtenemos que:

$$q_1 = P(X = 1 | Y = 1, Z = 1) = \frac{P(Y = 1 | X = 1, Z = 1)P(X = 1 | Z = 1)}{P(Y = 1 | Z = 1)}$$

¹Las etiquetas “bueno” o “malo” no asumen una moralidad objetiva. Matemáticamente, operan como categorías relativas al sistema de valores del observador que define el experimento.

Un supuesto clave en nuestro análisis es la *independencia a priori* entre la composición de la sociedad y el perfil del líder. En otras palabras, asumimos que la proporción base de individuos “buenos” no cambia por el simple hecho de evaluar a un líder en particular, lo que matemáticamente se traduce en $P(X = 1 | Z = 1) = P(X = 1) = \pi_g$. Por otro lado, utilizando el Teorema de la Probabilidad Total se sigue que:

$$P(Y = 1 | Z = 1) = P(Y = 1 | X = 1, Z = 1)P(X = 1) + P(Y = 1 | X = 0, Z = 1)P(X = 0)$$

Sustituyendo obtenemos que la “pureza” esperada de los seguidores de un líder bueno es:

$$q_1 = \frac{\beta\pi_g}{\beta\pi_g + \gamma(1 - \pi_g)}$$

De manera análoga, para un líder malo ($Z = 0$), la probabilidad de que uno de sus seguidores sea bueno es:

$$q_0 = \frac{\delta\pi_g}{\delta\pi_g + \alpha(1 - \pi_g)}$$

Consideremos una muestra de N seguidores, con características observables X_1, \dots, X_N . Supongamos que la muestra está conformada por N_g individuos “buenos” y N_b “malos”. Asumiendo que estas características son condicionalmente independientes e idénticamente distribuidas (i.i.d.) dado el perfil del líder, la función de verosimilitud $L(Z)$ sigue una distribución binomial [2, Sec. 2.1]:

$$L(Z = 1) = \binom{N}{N_g} q_1^{N_g} (1 - q_1)^{N_b} \quad (1)$$

$$L(Z = 0) = \binom{N}{N_g} q_0^{N_g} (1 - q_0)^{N_b} \quad (2)$$

Sea $p_0 = P(Z = 1)$ la creencia a priori. La actualización de la creencia (probabilidad a posteriori) es:

$$P(Z = 1 | N_g, N_b) = \frac{p_0 \cdot L(Z = 1)}{p_0 \cdot L(Z = 1) + (1 - p_0) \cdot L(Z = 0)} \quad (3)$$

El comportamiento asintótico de la probabilidad a posteriori (3) es la propiedad fundamental de nuestro modelo. Específicamente, como se establece en el siguiente resultado, la acumulación de evidencia empírica es suficiente para corregir cualquier sesgo introducido por la creencia a priori (p_0).

Proposición 1. *Asumiendo que el sistema es identificable ($q_1 \neq q_0$) y que la creencia a priori no es dogmática ($0 < p_0 < 1$), la distribución a posteriori converge casi seguramente (c.s.) al verdadero estado latente del líder cuando el tamaño de la muestra $N \rightarrow \infty$. Específicamente:*

$$\text{Si el líder es “bueno”}: \quad \lim_{N \rightarrow \infty} P(Z = 1 | N_g, N_b) = 1 \quad \text{c.s.}$$

$$\text{Si el líder es “malo”}: \quad \lim_{N \rightarrow \infty} P(Z = 1 | N_g, N_b) = 0 \quad \text{c.s.}$$

La demostración de esta proposición es una aplicación directa del Teorema de Consistencia de Doob [9, Sec. 10.4], cuyos detalles se desarrollan en el Apéndice A. De manera complementaria, la Figura 1 ilustra este comportamiento, evidenciando cómo la probabilidad a posteriori supera rápidamente la incertidumbre inicial y converge hacia la certeza absoluta a medida que aumenta el tamaño de la muestra.

En el contexto del aprendizaje automático, la Ecuación (3) corresponde a la formulación exacta de un clasificador Naive Bayes para variables de Bernoulli [7, Sec. 3.5]. Aquí, el líder opera como la clase latente a inferir (Z) y los perfiles de los seguidores actúan como características observacionales condicionalmente independientes (X_i). Así, el algoritmo Naive Bayes describe el proceso de actualizar las creencias paso a paso frente a cada nueva observación, mientras que la Proposición 1 asegura que el proceso de aprendizaje no queda atrapado en sesgos iniciales, sino que converge hacia el verdadero perfil del líder al acumularse la evidencia empírica.

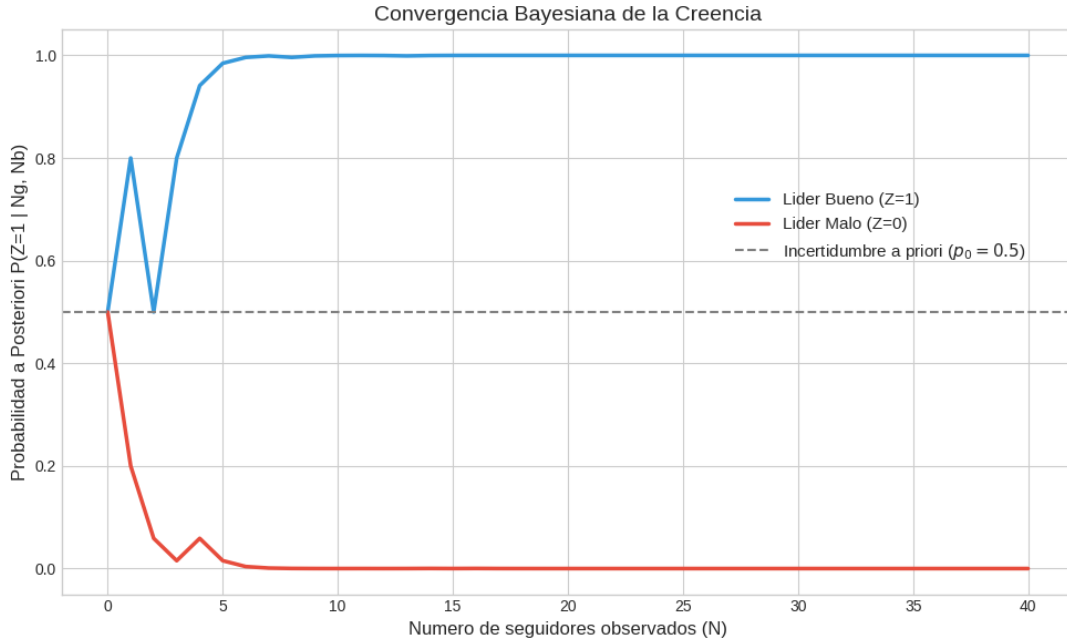


Figura 1: Convergencia asintótica de la probabilidad a posteriori. Simulación para $N = 40$ seguidores partiendo de máxima incertidumbre inicial ($p_0 = 0,5$). La dinámica asume una población simétrica ($\pi_g = 0,5$) con una fuerte tendencia homofílica ($\alpha = 0,8, \beta = 0,8$) y presencia de ruido en el sistema ($\gamma = 0,2, \delta = 0,2$).

4. Ejemplo Metodológico

Para ilustrar la utilidad de este marco teórico, consideremos una simulación práctica. Supongamos que un individuo evalúa a un líder sobre el cual tiene máxima incertidumbre inicial ($p_0 = 0,5$). Si asumimos una población simétrica ($\pi_g = 0,5$) y una dinámica de apoyo con fuerte tendencia homofílica ($\alpha = 0,8, \beta = 0,8, \gamma = 0,2, \delta = 0,2$), obtenemos que $q_1 = 0,8$ y $q_0 = 0,2$. La estructura oculta detrás de estas interacciones se ilustra en la Figura 2. Al realizar un muestreo en su entorno, el observador identifica a $N = 20$ seguidores, de los cuales clasifica a $N_g = 16$ como “buenos” y a los $N_b = 4$ restantes como “malos”.

Sustituyendo en la Ecuación (3), la verosimilitud se maximiza en $Z = 1$, obteniéndose una probabilidad a posteriori $P(Z = 1 | 16, 4) \approx 0,999$. Matemáticamente, esto indica que los datos favorecen la hipótesis $Z = 1$.

Es importante aclarar que este ejemplo supone una población dividida en partes iguales y con la misma tendencia a agruparse. En la realidad, las subpoblaciones suelen presentar asimetrías ($\pi_g \neq 0,5$ o $\alpha \neq \beta$). Si, por ejemplo, un grupo es minoritario pero posee una lealtad ideológica extrema, el modelo es matemáticamente sensible a este desbalance, requiriendo un tamaño de muestra observacional N sustancialmente mayor para que la inferencia asintótica converja.

5. Discusión y Limitaciones del Modelo

La validez de la Proposición 1 requiere condiciones de muestreo ideales; sin embargo, la aplicación de este modelo en sistemas sociales reales introduce ruidos no triviales que limitan su alcance directo. A continuación, discutimos tres factores que perturban la inferencia:

- **Subjetividad en la Clasificación:** La etiqueta X_i no es un observable físico estricto, sino una estimación condicionada al esquema de valores del observador. El resultado del modelo

Red Bipartita: Homofilia y Ruido en la Formación de Seguidores

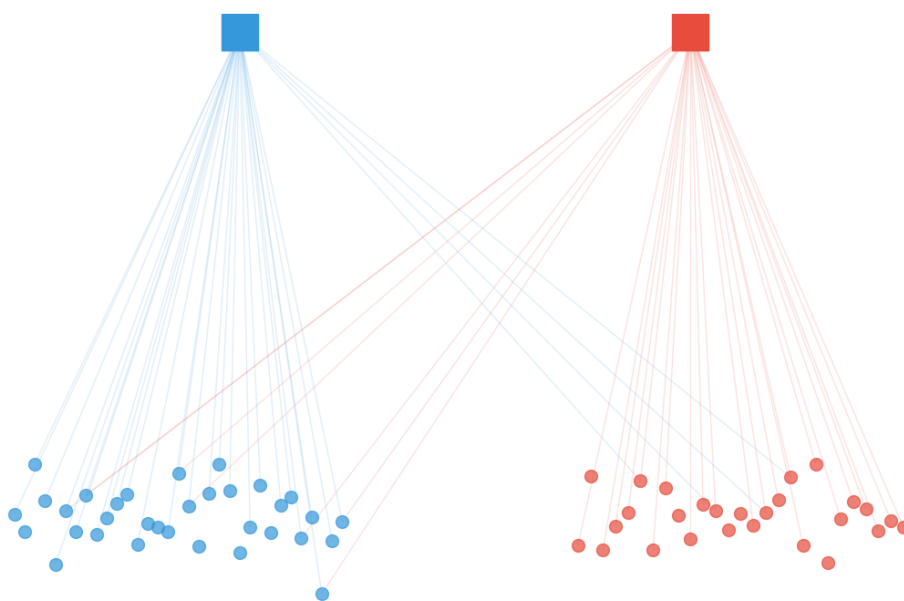


Figura 2: Estructura bipartita generada para una población de 60 individuos. La simulación asume una población simétrica ($\pi_g = 0,5$) y una dinámica fuerte de homofilia ($\alpha = 0,8, \beta = 0,8, \gamma = 0,2$ y $\delta = 0,2$).

no confirma la bondad universal del líder, sino que certifica la consistencia estadística de la bondad *relativa* al sistema de referencia de quien evalúa.

- **Contagio Social e Independencia Condicional:** La derivación de las Ecuaciones (1) y (2) asume que las características observables de los seguidores son independientes dado el líder (supuesto i.i.d.). En redes reales, el contagio social invalida en cierta medida esta premisa debido a que los seguidores de un líder suelen influenciarse mutuamente. Así, si las posturas están correlacionadas, la evidencia empírica se vuelve redundante. Contabilizar a N seguidores que se influncian fuertemente entre sí aporta menos información estadística que observar a N individuos con criterios aislados y autónomos.
- **Cámara de Eco y Muestreo No Uniforme:** La aproximación asume que la muestra proviene de una distribución representativa de la población global. Al observar el entorno inmediato, el muestreo sufre un severo sesgo de selección causado por la conectividad preferencial. Esto retroalimenta las creencias del observador (el efecto *Echo Chamber* [8]) y produce una sobreestimación en la varianza de la verosimilitud del modelo.

6. Conclusiones

El modelo planteado permite concluir que la red social de un líder no es una consecuencia fortuita de su discurso, sino una manifestación estadística de su estado latente. Comprendiendo las perturbaciones del muestreo en redes homofílicas, y bajo la premisa de que los niveles de ruido del sistema (γ, δ) no superen las fuerzas de afinidad natural (α, β), la inferencia bayesiana soporta la sabiduría popular: *si tienes dudas sobre un líder, basta observar a sus seguidores*.

Más allá de su valor metodológico, este resultado invita a una reflexión frente a coyunturas marcadas por la polarización. En escenarios donde el debate público se satura, el discurso de

los líderes tiende a perder su capacidad diferenciadora y las verdaderas intenciones se ocultan tras estrategias de comunicación. Las palabras que profesa un líder pueden ser cuidadosamente moldeadas para persuadir o engañar; sin embargo, la composición de su base de apoyo constituye un fenómeno estadístico mucho más difícil de falsificar. Por lo tanto, es necesario no limitarse al análisis individual de un líder, sino evaluar con el mismo rigor el entorno social y ético que lo respalda.

Finalmente, destacamos la simetría del modelo propuesto. Aunque este trabajo se ha centrado en inferir el estado latente de la figura central (Z) a partir de sus seguidores observables (X_i), el marco probabilístico es bidireccional. Conservando los mismos parámetros de transición homofílica y de ruido ($\alpha, \beta, \gamma, \delta$), el modelo permite resolver el problema inverso: inferir la naturaleza latente de un individuo particular (X) a partir del perfil conocido de los múltiples líderes a los que decide apoyar (Z_k). Esta dualidad también fundamenta el principio inverso: “*dime a quién sigues y te diré quién eres*”.

Referencias

- [1] Clinton, J., Jackman, S., & Rivers, D. (2004). The Statistical Analysis of Roll Call Data. *American Political Science Review*, 98(2), 355-370.
- [2] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd ed.). CRC press.
- [3] Lanchier, N. (2017). *Stochastic Interacting Systems in Life and Social Sciences*. Springer.
- [4] Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent Structure Analysis*. Houghton Mifflin.
- [5] Liggett, T. M. (1999). *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Springer-Verlag.
- [6] McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), 415-444.
- [7] Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. MIT Press.
- [8] Sunstein, C. R. (2001). *Echo Chambers: Bush v. Gore, Impeachment, and Beyond*. Princeton University Press.
- [9] van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press.

Contribución de los autores

Los autores declaran haber contribuido de manera equitativa en todas las etapas de la investigación, incluyendo la conceptualización, la metodología, el análisis formal y la redacción del manuscrito.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Disponibilidad de los datos de investigación

El código de las simulaciones se encuentra disponible en el Apéndice B del presente documento.

A. Teorema de Doob y Demostración de la Proposición 1

Para probar la Proposición 1, usamos el Teorema de Consistencia de Doob (1949; véase [9, Cap. 10]). Este resultado de la estadística asintótica garantiza que, bajo condiciones generales, la distribución a posteriori siempre converge hacia la verdad a medida que acumulamos datos. A continuación, enunciamos el teorema general y verificamos sus condiciones para nuestro modelo.

Teorema (Doob, 1949). *Sea Θ el espacio de parámetros y \mathcal{X} el espacio muestral. Sea $\theta \mapsto P_\theta$ una aplicación inyectiva que asigna a cada parámetro una medida de probabilidad en \mathcal{X} . Si Π es una distribución a priori sobre Θ , entonces para casi todo $\theta_0 \in \Theta$ con respecto a la medida Π , la distribución a posteriori $\Pi(\cdot \mid X_1, \dots, X_N)$ converge a una distribución degenerada en el verdadero parámetro θ_0 casi seguramente cuando $N \rightarrow \infty$.*

Demostración de la Proposición 1.

Para aplicar el Teorema de Doob a nuestro modelo, verificamos el cumplimiento de las siguientes condiciones:

1. **Estructura de los espacios:** Tanto el espacio del parámetro (el estado oculto del líder, $\Theta = \{0, 1\}$) como el espacio de las observaciones (el perfil de cada seguidor, $\mathcal{X} = \{0, 1\}$) son conjuntos finitos y discretos. Esto garantiza que las operaciones de probabilidad condicional estén bien definidas.
2. **Identificabilidad del modelo (Inyectividad):** Cada estado del líder debe generar una distribución de seguidores claramente distinguible. En nuestro modelo, un líder bueno y un líder malo producen probabilidades distintas de observar seguidores buenos (q_1 y q_0 , respectivamente). Al exigir la condición de identificabilidad ($q_1 \neq q_0$), garantizamos que los datos contengan información útil. Si las probabilidades fueran iguales ($q_1 = q_0$), la aplicación no sería inyectiva, por lo que ambos líderes serían estadísticamente idénticos y el sistema no podría aprender de las observaciones.
3. **Creencia inicial no dogmática:** El teorema requiere que la distribución a priori Π no descarte el verdadero estado de la naturaleza (θ_0). Al establecer una creencia inicial no dogmática ($0 < p_0 < 1$), aseguramos que tanto el estado bueno ($\theta = 1$) como el estado malo ($\theta = 0$) tengan una probabilidad inicial estrictamente mayor que cero. Por lo tanto, el teorema es válido sin importar cuál sea la verdadera naturaleza del líder.

Al cumplirse todas las condiciones, el Teorema de Doob asegura que la distribución a posteriori converge casi seguramente a una distribución degenerada en el verdadero estado del líder (θ_0). Esto es, con probabilidad 1, tenemos que

$$\lim_{N \rightarrow \infty} P(Z = 1 \mid X_1, \dots, X_N) = \begin{cases} 1, & \text{si el líder es bueno } (\theta_0 = 1), \\ 0, & \text{si el líder es malo } (\theta_0 = 0). \end{cases}$$

Esto demuestra formalmente el comportamiento asintótico presentado en la Proposición 1.

B. Código en Python: Simulación y Gráficas

El siguiente script de Python fue utilizado para generar las simulaciones de la dinámica bayesiana y la estructura de interacción bipartita, usando la librería `networkx` para modelar la homofilia estocástica.

Nota sobre la implementación computacional: Computar la Ecuación (3) con muestras grandes implica multiplicar repetidamente números muy pequeños, lo que puede causar que el

computador redondee a cero (un error de precisión numérica conocido como *underflow*). Para evitar esto y garantizar la estabilidad de la simulación, el script realiza la inferencia calculando el logaritmo de la razón de probabilidades (*log-odds*), es decir, el logaritmo del cociente entre $P(Z = 1 | N_g, N_b)$ y $P(Z = 0 | N_g, N_b)$. Esto permite sumar los valores en lugar de multiplicarlos. Finalmente, el resultado obtenido se proyecta de vuelta al espacio de probabilidad estándar $[0, 1]$ mediante la función logística (sigmoide).

```

1 import numpy as np
2 import matplotlib.pyplot as plt
3 import networkx as nx
4
5 # Configuraciones de estilo
6 plt.style.use('seaborn-v0_8-whitegrid')
7 colors = ['#e74c3c', '#3498db'] # Rojo (Malo), Azul (Bueno)
8
9 # -----
10 # 1. SIMULACION DE CONVERGENCIA ASINTOTICA (Teorema de Doob)
11 # -----
12 np.random.seed(42)
13 p0 = 0.5
14 pi_g = 0.5
15 beta, alpha = 0.8, 0.8
16 gamma, delta = 0.2, 0.2
17 q1 = (beta * pi_g) / (beta * pi_g + gamma * (1 - pi_g))
18 q0 = (delta * pi_g) / (delta * pi_g + alpha * (1 - pi_g))
19 N_max = 40
20
21 # Simulamos los seguidores reales de un LIDER BUENO (Z=1)
22 followers = np.random.binomial(1, q1, N_max)
23 N_g_cum = np.cumsum(followers)
24 N_b_cum = np.arange(1, N_max + 1) - N_g_cum
25
26 # Inicializamos las listas con la creencia a priori en N=0
27 posteriors_good = [p0]
28 posteriors_bad = [p0]
29
30 # Calcular la posterior iterativamente
31 for ng, nb in zip(N_g_cum, N_b_cum):
32     log_odds = np.log(p0 / (1 - p0)) + ng * np.log(q1 / q0) + nb * np.log((1 -
33     q1) / (1 - q0))
34     post_1 = 1 / (1 + np.exp(-log_odds))
35     posteriors_good.append(post_1)
36
37 # Simulamos lo mismo pero para un LIDER MALO (Z=0)
38 followers_bad_leader = np.random.binomial(1, q0, N_max)
39 N_g_cum_b = np.cumsum(followers_bad_leader)
40 N_b_cum_b = np.arange(1, N_max + 1) - N_g_cum_b
41
42 for ng, nb in zip(N_g_cum_b, N_b_cum_b):
43     log_odds = np.log(p0 / (1 - p0)) + ng * np.log(q1 / q0) + nb * np.log((1 -
44     q1) / (1 - q0))
45     post_1 = 1 / (1 + np.exp(-log_odds))
46     posteriors_bad.append(post_1)
47
48 plt.figure(figsize=(10, 6))
49 plt.plot(posteriors_good, label='Lider Bueno (Z=1)', color=colors[1], lw=2.5)
50 plt.plot(posteriors_bad, label='Lider Malo (Z=0)', color=colors[0], lw=2.5)
51 plt.axhline(0.5, color='gray', linestyle='--', label='Incertidumbre a priori (
    $p_0=0.5$)')
52 plt.title('Convergencia Bayesiana de la Creencia', fontsize=14)
53 plt.xlabel('Numero de seguidores observados (N)', fontsize=12)

```

```

52 plt.ylabel('Probabilidad a Posteriori P(Z=1 | Ng, Nb)', fontsize=12)
53 plt.legend(fontsize=11, loc='center left', bbox_to_anchor=(0.65, 0.65))
54 plt.tight_layout()
55 plt.savefig('convergencia_bayesiana.png', dpi=300)
56 plt.show()
57
58 # -----
59 # 2. MODELADO DE RED BIPARTITA (Homofilia)
60 # -----
61 B = nx.Graph()
62
63 # Nodos Lideres
64 B.add_node("L_Bueno", bipartite=0, type='good')
65 B.add_node("L_Malo", bipartite=0, type='bad')
66
67 # Nodos Seguidores (N=60)
68 n_followers = 60
69 follower_types = np.random.choice(['good', 'bad'], size=n_followers, p=[pi_g,
    1-pi_g])
70
71 for i in range(n_followers):
72     follower_id = f"S_{i}"
73     B.add_node(follower_id, bipartite=1, type=follower_types[i])
74
75     # Asignar aristas segun probabilidades alfa, beta, gamma, delta
76     if follower_types[i] == 'good':
77         if np.random.rand() < beta: B.add_edge(follower_id, "L_Bueno")
78         if np.random.rand() < delta: B.add_edge(follower_id, "L_Malo")
79     else:
80         if np.random.rand() < gamma: B.add_edge(follower_id, "L_Bueno")
81         if np.random.rand() < alpha: B.add_edge(follower_id, "L_Malo")
82
83 # Dibujar
84 plt.figure(figsize=(12, 8))
85 pos = dict()
86 pos["L_Bueno"] = (-1, 10)
87 pos["L_Malo"] = (1, 10)
88
89 followers_good = [n for n, d in B.nodes(data=True) if d.get('bipartite') == 1
    and d.get('type') == 'good']
90 followers_bad = [n for n, d in B.nodes(data=True) if d.get('bipartite') == 1
    and d.get('type') == 'bad']
91
92 # Distribuir seguidores en arco
93 for i, node in enumerate(followers_good):
94     pos[node] = (-2 + 1.5*(i/len(followers_good)), np.random.normal(0, 0.5))
95 for i, node in enumerate(followers_bad):
96     pos[node] = (0.5 + 1.5*(i/len(followers_bad)), np.random.normal(0, 0.5))
97
98 # Dibujar Nodos
99 nx.draw_networkx_nodes(B, pos, nodelist=["L_Bueno"], node_color=colors[1],
    node_size=800, node_shape='s')
100 nx.draw_networkx_nodes(B, pos, nodelist=["L_Malo"], node_color=colors[0],
    node_size=800, node_shape='s')
101 nx.draw_networkx_nodes(B, pos, nodelist=followers_good, node_color=colors[1],
    node_size=100, alpha=0.7)
102 nx.draw_networkx_nodes(B, pos, nodelist=followers_bad, node_color=colors[0],
    node_size=100, alpha=0.7)
103
104 # Dibujar Aristas
105 edges_good = [(u, v) for u, v in B.edges() if v == "L_Bueno" or u == "L_Bueno"]
106 edges_bad = [(u, v) for u, v in B.edges() if v == "L_Malo" or u == "L_Malo"]
107

```

```
108 nx.draw_networkx_edges(B, pos, edgelist=edges_good, alpha=0.15, edge_color=
    colors[1])
109 nx.draw_networkx_edges(B, pos, edgelist=edges_bad, alpha=0.15, edge_color=
    colors[0])
110
111 plt.title('Red Bipartita: Homofilia y Ruido en la Formacion de Seguidores',
    fontsize=16)
112 plt.axis('off')
113 plt.tight_layout()
114 plt.savefig('red_bipartita.png', dpi=300)
115 plt.show()
```

Este preprint fue presentado bajo las siguientes condiciones:

- Los autores declaran que se obtuvieron los términos necesarios del consentimiento libre e informado de los participantes o pacientes en la investigación y se describen en el manuscrito, cuando corresponde.
- Los autores declaran que la preparación del manuscrito siguió las normas éticas de comunicación científica.
- Los autores declaran que son conscientes de que son los únicos responsables del contenido del preprint y que el depósito en SciELO Preprints no significa ningún compromiso por parte de SciELO, excepto su preservación y difusión.
- Los autores declaran que los datos, las aplicaciones y otros contenidos subyacentes al manuscrito están referenciados.
- El manuscrito depositado está en formato PDF.
- Los autores declaran que la investigación que dio origen al manuscrito siguió buenas prácticas éticas y que las aprobaciones necesarias de los comités de ética de investigación, cuando corresponda, se describen en el manuscrito.
- Los autores declaran que una vez que un manuscrito es postado en el servidor SciELO Preprints, sólo puede ser retirado mediante solicitud a la Secretaría Editorial deSciELO Preprints, que publicará un aviso de retracción en su lugar.
- Los autores aceptan que el manuscrito aprobado esté disponible bajo licencia [Creative Commons CC-BY](#).
- El autor que presenta el manuscrito declara que las contribuciones de todos los autores y la declaración de conflicto de intereses se incluyen explícitamente y en secciones específicas del manuscrito.
- Los autores declaran que el manuscrito no fue depositado y/o previamente puesto a disposición en otro servidor de preprints o publicado en una revista.
- Si el manuscrito está siendo evaluado o siendo preparando para su publicación pero aún no ha sido publicado por una revista, los autores declaran que han recibido autorización de la revista para hacer este depósito.
- El autor que envía el manuscrito declara que todos los autores del mismo están de acuerdo con el envío a SciELO Preprints.