

## Tesis Doctoral

# Fenómenos de propagación en redes complejas

Medus, Andrés Daniel

2015-06-12

Este documento forma parte de la colección de tesis doctorales y de maestría de la Biblioteca Central Dr. Luis Federico Leloir, disponible en [digital.bl.fcen.uba.ar](http://digital.bl.fcen.uba.ar). Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the doctoral theses collection of the Central Library Dr. Luis Federico Leloir, available in [digital.bl.fcen.uba.ar](http://digital.bl.fcen.uba.ar). It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

Medus, Andrés Daniel. (2015-06-12). Fenómenos de propagación en redes complejas. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.

Cita tipo Chicago:

Medus, Andrés Daniel. "Fenómenos de propagación en redes complejas". Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 2015-06-12.

**EXACTAS** UBA

Facultad de Ciencias Exactas y Naturales



**UBA**

Universidad de Buenos Aires



UNIVERSIDAD DE BUENOS AIRES

Facultad de Ciencias Exactas y Naturales

Departamento de Física

*Fenómenos de propagación en redes complejas*

Tesis presentada para optar al título de  
Doctor de la Universidad de Buenos Aires en el área Ciencias Físicas

por

**Lic. Andrés Daniel Medus**

Director de Tesis: Dr. Claudio Oscar Dorso

Consejero de Estudios: Dr. Pablo Daniel Mininni

Lugar de Trabajo: Departamento de Física - FCEyN - UBA

Buenos Aires, 2015

Fecha de defensa: 12/06/2015



---

## Resumen

En esta tesis estudiamos dos ejemplos de fenómenos de propagación sobre redes sociales, ambos vinculados con salud pública. En el primero de ellos, analizamos el efecto de la diseminación de la opinión respecto de la vacunación sobre una red sencilla integrada por grupos familiares. En el espíritu del modelo de Axelrod para la diseminación cultural en una sociedad, consideramos la conducta de vacunación como una de las características que definen el acervo cultural de los agentes sociales. Estudiamos las consecuencias del accionar de pequeños grupos anti-vacunación en la diseminación de la opinión contraria a la vacunación y, con ello, en la cobertura final de una vacuna. En particular, nos interesamos por el caso de la vacuna Triple Vírica (**TC**) contra el sarampión, la rubeola y las paperas, debido al reciente resurgimiento del sarampión en países con amplia disponibilidad de vacunas como EEUU, Inglaterra y Gales, entre otros. Mostramos que el mecanismo de diseminación cultural aquí introducido promueve la formación de dominios o clusters de no-vacunadores, susceptibles de contraer sarampión, incrementando la probabilidad de brotes, aún en aquellos casos donde se cumple con la meta del 95 % de cobertura de la vacuna **TC**, propuesta por la Organización Mundial de la Salud (OMS).

Por otra parte, consideramos una variante del modelo de red social donde las opiniones pueden propagarse tanto a través de enlaces personales como no-personales, de modo que estos últimos no permiten el contagio de la enfermedad, aunque sí la diseminación de opiniones. Asimismo, estudiamos el efecto de considerar redes donde la diseminación de opiniones es acompañada por un proceso de adaptación estructural a través del reconexión de enlaces.

El segundo fenómeno analizado corresponde a la propagación del sarampión sobre una red social. Para ello, desarrollamos un modelo generador de redes sociales con características compatibles con las reportadas para redes obtenidas a través de experimentos sociales, o mediante el análisis de redes online como Facebook. Mostramos que las características estructurales de la red afectan la probabilidad de persistencia del sarampión sobre la misma.

**Palabras Clave:** redes sociales, diseminación de opiniones, vacunación, modelos basados en agentes, modelos de crecimiento de redes complejas, epidemias

---

---

## Spreading phenomena in complex networks

### Abstract

In this thesis we study two examples of spreading phenomena on social networks, both related with public health. In the first one, we analyze the effect of the spread of opinion regarding vaccination on family groups composing a simple social network. We consider the vaccination behavior of family groups as one of the defining characteristics of their cultural heritage, in the spirit of Axelrod's model for the study of cultural dissemination in human societies. We study the consequences of the actions of small anti-vaccination groups in the dissemination of opinions against vaccination, thus on vaccine coverage. In particular, we are interested in the case of the MMR vaccine against measles, rubella and mumps, due to the recent measles resurgence in countries with widespread availability of vaccines such as USA, England and Wales, among others. We show that the mechanism of cultural dissemination introduced here promotes the formation of domains or clusters of non-vaccinators, then susceptible to measles, increasing the outbreaks probability, even in those cases where the goal of 95% MMR vaccine coverage recommended by the World Health Organization (WHO) is fulfilled.

Moreover, we consider a variant of our social network model where opinions can spread through both personal and non-personal links, so that the latter allow for the dissemination of opinions, but does not allow the spread of the disease. We also study the effect of considering social networks where the spread of opinions is accompanied by a structural adaptation process through the rewiring of links.

The second phenomenon discussed here corresponds to the spread of measles on a social network. For this aim, we developed a model for social networks growth with features compatible with those reported in literature for networks obtained through social experiments, or from online social networks such as Facebook. We show that topological characteristics of the network affect measles persistence on the top of them.

**Keywords:** social networks, spread of opinions, vaccination, agent-based modeling, complex networks growth models, epidemics

---

---

## Agradecimientos

Pasaron varios años, mucho aprendizaje, cosas que salieron bien, otras que no tanto, pero ha llegado el final de este camino. Por ello debo agradecer a varias personas que hicieron de esta etapa una experiencia transformadora.

En primer lugar, agradezco a Claudio por haberme abierto las puertas de su grupo, por orientarme y permitirme trabajar en un ambiente estimulante, siempre pendiente de los proyectos, con entusiasmo y energía envidiables. Sospecho que al leer estas líneas estará sorprendido ya que nunca, en todos estos años de tratarlo, lo llamé por su nombre de pila. Ha sido una gran experiencia compartir estos años de intenso trabajo con él. Gracias por todo Claudio.

Agradezco al Departamento de Física y a todos sus miembros, que trabajan incansablemente para que las cosas funcionen. Más allá de algunos desacuerdos, hay cosas en las que todos debemos coincidir. El objetivo de brindar una educación pública, gratuita y de excelencia, es sin dudas una de ellas. Hago extensivo este agradecimiento a la Universidad de Buenos Aires por brindarme la posibilidad de estudiar en forma gratuita y financiar mi doctorado.

En este camino tuve la suerte de cruzarme con algunas personas que facilitaron mi existencia académica. Una de ellas es Marcelo Otero, un gran compañero de oficina y mejor amigo, tan brillante como humilde, además de excelente persona. También quiero agradecer a Daniel Barmak por los gratos momentos, las conversaciones tecnológicas y las infusiones exóticas compartidas (mate cocido sabor frutilla!?). A Paula Bergero, otra excelente compañera de oficina y amiga. A Inés, una amiga que siempre estuvo en los momentos difíciles. A los compañeros de la otra oficina: Pedrito, Guille, Juani y Pablo. A Pablo Balenzuela por el apoyo. A Adrián, el RRPP del DF.

Agradezco a mis padres y hermana por el apoyo. A mis tías. A mi gran amiga Vero, por acompañarme con su presencia tanto real como virtual. A Marianito y Julián, por estar ahí aunque no nos veamos muy seguido.

Para el final, lo más importante: esta tesis está dedicada a Marcela, mi compañera. Gracias amor por estar siempre para apoyarme y sostenerme incondicionalmente, por darme paz, y, aún más importante, por enseñarme día a día el difícil arte de vivir.



# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Diseminación de la vacunación voluntaria . . . . .	2
1.1.1. Vacunación y redes sociales: contactos personales y no-personales . . . . .	5
1.1.2. Conducta de vacunación como un rasgo cultural: Modelo de Axelrod . . . . .	7
1.2. Propiedades estáticas y dinámicas de las redes sociales . . . . .	10
1.3. Persistencia del sarampión . . . . .	12
1.3.1. Sensibilidad de la persistencia con la distribución de tiempos	16
1.3.2. Efecto de la estructura de red social en la persistencia . . .	17
1.4. Estructura de la tesis . . . . .	18
<b>2. Breve introducción a las redes complejas</b>	<b>19</b>
2.1. Redes estáticas: propiedades topológicas . . . . .	19
2.1.1. Distribución de grado . . . . .	20
2.1.2. Distancia media entre nodos . . . . .	21
2.1.3. Coeficiente de clusterización . . . . .	23
2.1.4. Grado medio de primeros vecinos . . . . .	24
2.1.5. Estructura de comunas . . . . .	26
2.2. Clasificación topológica de redes complejas . . . . .	28
2.2.1. Redes aleatorias . . . . .	28
2.2.2. Redes de Mundo Pequeño . . . . .	30
2.2.3. Redes libres de escala . . . . .	31
<b>3. Diseminación de opiniones y vacunación: modelo de red estática</b>	<b>33</b>
3.1. Modelo de red social . . . . .	34
3.2. Modelo de diseminación cultural y vacunación. . . . .	35
3.2.1. Diseminación de la opinión respecto de la vacunación . . .	36
3.2.2. Movimientos anti-vacunación . . . . .	38
3.3. Descripción algorítmica del modelo . . . . .	38
3.4. Transición vacunadores a no-vacunadores . . . . .	40
3.5. Análisis de brotes de sarampión . . . . .	42
3.6. Análisis de clusters . . . . .	43
3.6.1. Incidencia sobre el cluster máximo . . . . .	45



3.6.2. Clusters máximos: Propiedades topológicas . . . . .	46
3.7. Alta cobertura de la vacuna . . . . .	47
3.8. Conclusiones preliminares . . . . .	49
<b>4. Diseminación de opiniones y vacunación: modelo generalizado</b>	<b>51</b>
4.1. Contactos personales y de opinión . . . . .	52
4.2. Enlaces adaptativos: mecanismo de reconexionado . . . . .	53
4.3. Descripción algorítmica del modelo generalizado . . . . .	55
4.4. Red estática: efectos de la heterogeneidad de enlaces . . . . .	56
4.5. Red adaptativa . . . . .	58
4.5.1. Formación de clusters . . . . .	59
4.5.2. Clusters máximos . . . . .	60
4.6. Acción de las campañas de vacunación . . . . .	62
4.6.1. Modelo generalizado con campo externo . . . . .	63
4.6.2. Barrido en $q$ y $\phi$ . . . . .	64
4.6.3. Acción de $\phi$ sobre los modelos con red estática y adaptativa	67
4.6.4. Incidencia de $\phi$ en la topología de los clusters máximos. . .	68
4.7. Modelo generalizado con alta cobertura . . . . .	70
4.8. Conclusiones preliminares . . . . .	71
<b>5. Generación de redes sociales con base empírica</b>	<b>73</b>
5.1. Actividad social como generador de heterogeneidades estructurales	74
5.2. Modelo estocástico generalizado . . . . .	76
5.2.1. Efectos de memoria no-markovianos: cierre triádico . . . . .	77
5.2.2. Formulación analítica del modelo . . . . .	78
5.2.3. Población constante ( $\gamma = 0$ ) . . . . .	81
5.2.4. Población creciente . . . . .	84
5.3. Correlaciones de grado de alto orden . . . . .	86
5.3.1. Grado medio de primeros vecinos . . . . .	86
5.4. Coeficiente de clusterización . . . . .	90
5.5. Casos de estudio . . . . .	91
5.5.1. Propiedades topológicas . . . . .	93
5.5.2. Mecanismo <b>TC</b> en redes reales . . . . .	95
<b>6. Persistencia del sarampión sobre redes sociales</b>	<b>97</b>
6.1. El modelo . . . . .	98
6.1.1. La red social . . . . .	99
6.1.2. Tasas de transición . . . . .	99

---

6.1.3. Paso temporal . . . . .	101
6.1.4. Aproximación multinomial . . . . .	102
6.2. Análisis de la persistencia . . . . .	102
6.2.1. Extinciones post-epidémicas con importación permanente de infectados . . . . .	105
<b>7. Conclusiones Finales y Perspectivas</b>	<b>109</b>
<b>A. Modelo epidemiológico</b>	<b>115</b>
A.1. Modelo compartimental estocástico . . . . .	115
<b>B. Breve introducción al modelo de variables ocultas</b>	<b>119</b>
B.1. Modelo de variables ocultas . . . . .	119
<b>Bibliografía</b>	<b>123</b>



CAPÍTULO 1

# Introducción

---

## Índice

---

<b>1.1. Diseminación de la vacunación voluntaria . . . . .</b>	<b>2</b>
1.1.1. Vacunación y redes sociales: contactos personales y no-personales . . . . .	5
1.1.2. Conducta de vacunación como un rasgo cultural: Modelo de Axelrod . . . . .	7
<b>1.2. Propiedades estáticas y dinámicas de las redes sociales .</b>	<b>10</b>
<b>1.3. Persistencia del sarampión . . . . .</b>	<b>12</b>
1.3.1. Sensibilidad de la persistencia con la distribución de tiempos	16
1.3.2. Efecto de la estructura de red social en la persistencia . .	17
<b>1.4. Estructura de la tesis . . . . .</b>	<b>18</b>

---

Las sociedades humanas constituyen ejemplos notables de sistemas complejos. Estos grandes entramados de individuos vinculados dan lugar a comportamientos emergentes, muchas veces imprevistos al considerar los sencillos mecanismos microscópicos involucrados. Esto se debe a que tanto el patrón como la naturaleza de las interacciones intervinientes entre un gran número de agentes sociales, desempeñan un rol fundamental en el desarrollo de la complejidad.

El análisis cuantitativo de los fenómenos de propagación de enfermedades, rumores o rasgos culturales entre individuos, ha sido tradicionalmente abordado desde los llamados modelos de *campo medio*. Estos posibilitan una gran simplificación en el modelado matemático al soslayar algunas de las heterogeneidades que naturalmente presentan los sistemas sociales.

Entre las heterogeneidades fundamentales, se encuentra la fuerte estructura subyacente de vínculos entre sujetos. Dicha estructura suele representarse a través de una *red compleja*, donde sus *nodos* identifican a los agentes sociales, mientras que los *enlaces* que los conectan pueden representar vínculos laborales o familiares, amistades, proximidad espacial o algún grado de similitud entre los

mismos. En ciertas circunstancias, dichas heterogeneidades impactan fuertemente tanto en la evolución como en el resultado final del fenómeno de propagación analizado, tornándose inapropiado su tratamiento a partir de modelos de campo medio.

Analizar la microescala del entramado social es actualmente una tarea posible gracias al advenimiento de las nuevas tecnologías de la información. El análisis de la participación en redes sociales online [Lazer 2009, Centola 2010, Ugander 2011, Szell 2010a], el patrón de llamadas por telefonía celular [Gonzalez 2008, Song 2010], el Sistema de Posicionamiento Global (GPS) y las tecnologías de comunicación inalámbrica de campo cercano disponibles en los nuevos teléfonos celulares [Eagle 2006, Eagle 2009, Cho 2011, Crandall 2010], todos ellos contribuyen a desentrañar las leyes microscópicas que dan forma a la organización de las sociedades humanas. Asimismo, han comenzado a desarrollarse diferentes experimentos tendientes a estudiar la evolución de los contactos sociales sobre poblaciones pequeñas de individuos en espacios cerrados [Cattuto 2010, Isella 2011]. Debido a su enorme volumen, estos datos son clasificados dentro de la categoría denominada “Big Data” y la física estadística provee técnicas adecuadas para su estudio.

El objetivo de esta tesis es analizar dos problemas vinculados con salud pública, donde la estructura de la red social condiciona el fenómeno de propagación particular que tiene lugar sobre ella. Adicionalmente, formularemos un modelo de crecimiento de redes que concilia el patrón de evolución observado para los contactos sociales en distintos experimentos, con las propiedades estáticas de las redes construidas a partir de los enlaces acumulados en el tiempo.

## 1.1. Diseminación de la vacunación voluntaria

La vacunación es el medio más eficaz para combatir la propagación de enfermedades infecciosas en seres humanos, solo detrás de la distribución de agua potable [Plotkin 2004]. Su acción ha permitido la erradicación de la viruela y el control global de otras enfermedades infecciosas como el sarampión, la rubeola, las paperas, la poliomielitis y el tétanos [Andre 2008, Maurice 2009]. Un ejemplo notable que da cuenta del efecto de la vacunación se observa en la Figura 1.1, donde se representa la incidencia del sarampión (acumulada anualmente) en el Reino Unido antes y después de la administración de la primera vacuna simple en 1968, así como también de la introducción de la vacuna Triple Vírica (TV) contra el sarampión, las paperas y la rubeola en 1988.

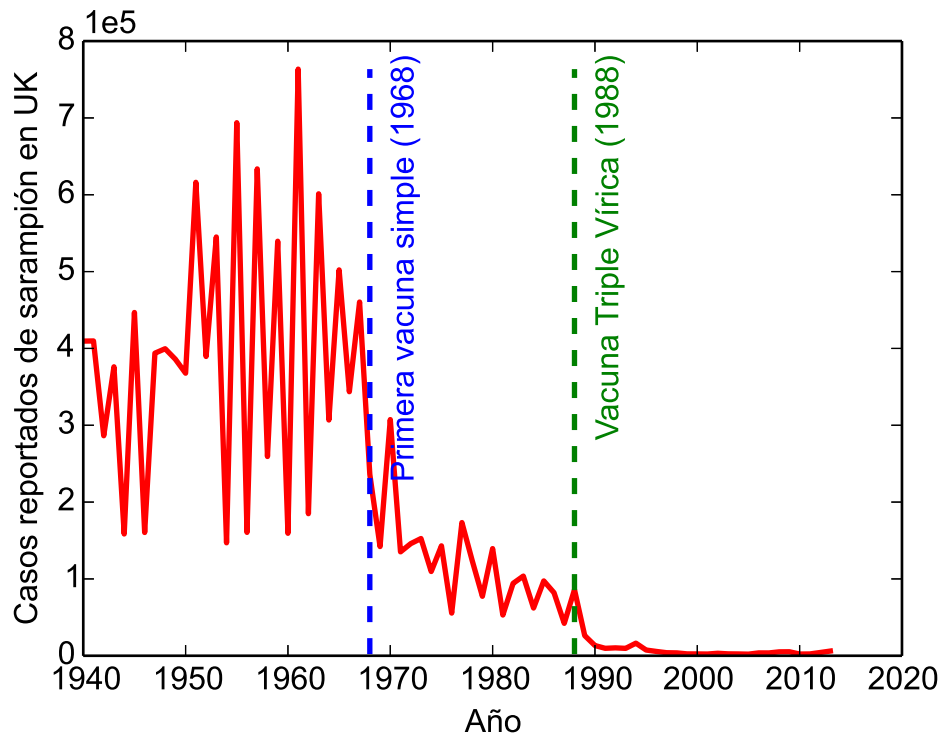


Figura 1.1: Casos de sarampión reportados anualmente en Reino Unido entre 1940-2013. En línea de trazos se señalan los años correspondientes a la introducción de las vacunas simple y Triple Vírica [HPA] (información pública disponible en el sitio web de Public Health England <http://www.hpa.org.uk>).

A pesar de su demostrada eficacia, la cobertura de la vacuna TV ha disminuido en algunos países desarrollados, con amplia disponibilidad de la misma, dando lugar a recientes brotes de sarampión como los acaecidos en Gales (2013, [Wal 2013]) y los Estados Unidos (2014/15, [Gastañaduy 2014]). Se estima que tal disminución en la cobertura debe su origen a la difusión de un estudio, rápidamente refutado, que establecía una correlación entre la administración de la vacuna y casos de autismo declarado en niños pequeños [Wakefield 1998, Taylor 1999]. Esta controversia constituye un ejemplo concreto que pone de manifiesto la importancia de la diseminación de opiniones en la percepción que los individuos desarrollan acerca de una vacuna [Burgess 2006, Brown 2012].

En aquellos casos donde la vacunación es opcional, o al menos donde existen exenciones por razones de conciencia (por ejemplo en algunos estados de los EEUU), son los padres quienes deciden respecto de la vacunación de sus hijos. No lo hacen como individuos aislados, sino como parte de una estructura social que incide en sus opiniones y sobre la cual, a su vez, pueden incidir. Es por

ello que recientemente ha comenzado a modelarse la diseminación de la *conducta respecto de la vacunación* como un proceso de *imitación social* en el contexto de modelos de campo medio [Bauch 2005]. De este modo, los individuos deciden su conducta ante la vacunación imitando el comportamiento medio de su entorno. Adicionalmente, otros modelos asumen individuos racionales, capaces de evaluar el riesgo latente de contagio de la enfermedad y de contrastarlo con los posibles efectos adversos de la vacuna, en el contexto de teoría de juegos [Bauch 2003, Bauch 2004, d'Onofrio 2011].

Alcanzar una amplia cobertura de la vacuna, por un lado, disminuye la probabilidad de brotes epidémicos y, por el otro, protege a aquellos individuos que por algún motivo ajeno a su voluntad no han podido vacunarse (por ejemplo, niños menores de 1 año, inmunodeprimidos, individuos alérgicos a los compuestos de la vacuna, etc.) a través del efecto de *inmunidad de grupo*. En esencia, este último asume que el riesgo de contagio de un individuo susceptible disminuye con la inmunización de su entorno. En este sentido, la Organización Mundial de la Salud (OMS) recomienda a los países socios el 95 % de cobertura de la vacuna TV, no solo con el objetivo de alcanzar inmunidad de grupo, sino para lograr la erradicación definitiva del sarampión [WHO 2012].

Es evidente que un decrecimiento en la cobertura de la vacuna implica un incremento de la probabilidad de brote epidémico de sarampión. No obstante, dichos brotes también se han suscitado en países con alta cobertura promedio. En 1999 Holanda contaba con un 96 % de cobertura total de TV en niños mayores de 14 meses, sin embargo, sufrió un brote de sarampión que alcanzó 1750 casos declarados [Sheldon 2000]. En 2005, alrededor de 500 personas asistieron a una reunión en una localidad de Indiana, EEUU, un estado donde la cobertura de la vacuna TV era en ese entonces del 92 % en niños preescolares y del 98 % en niños de sexto grado. Una de las asistentes a la reunión era una muchacha norteamericana de 17 años que regresaba de una misión religiosa en Rumania, donde contrajo sarampión. De los asistentes a la reunión, aproximadamente 50 (10 % del total) carecían de inmunidad contra el sarampión, de los cuales 16 contrajeron la enfermedad. Durante las seis semanas posteriores, se registraron 34 casos totales confirmados, de los cuales el 94 % no estaban vacunados [Parker 2006]. Brotes semejantes se registraron en San Diego, California, EEUU, en 2008 [Sugerman 2010], en Cataluña, España, en 2006 [Domínguez 2008] y, nuevamente, en Holanda también durante 2008 [Hahne 2010]. A estos se suman los ya mencionados y más recientes brotes en EEUU (2014) y Gales (2013). En todos ellos se erige un denominador común: los individuos no-vacunados no apa-

recen distribuidos al azar en la población sino que, por el contrario, se encuentran agrupados en comunidades por afinidad cultural.

### 1.1.1. Vacunación y redes sociales: contactos personales y no-personales

Modelar la agrupación de individuos por rasgos culturales requiere apartarse de la aproximación de población totalmente mezclada característica de los modelos de campo medio. Esto puede lograrse representando el sustrato social a través de una red compleja, donde los agentes sociales y sus interacciones corresponden a los nodos y enlaces de la red, respectivamente. Tal fue el enfoque implementado en trabajos recientes que proponen modelos de diseminación de opiniones por imitación social, considerando una opinión binaria relativa a la vacunación. El trabajo de Salathé et al [Salathé 2008] muestra que los agentes sociales que comparten opinión contraria a la vacunación, se agrupan en clusters debido al mecanismo de imitación social. De este modo, el efecto de inmunidad de grupo resulta fuertemente afectado.

La Figura 1.2 muestra la probabilidad de brote epidémico de sarampión para distintos porcentajes de cobertura de la vacuna, comparando el caso que incorpora el mecanismo de formación de opinión con aquel donde los individuos no-vacunados son distribuidos al azar sobre la red. Puede verse que la clusterización de no-vacunados incrementa notablemente la probabilidad de brote, efecto que se ve magnificado para coberturas altas. Sin embargo, en [Salathé 2008] la masa de no-vacunados no se obtiene dinámicamente del proceso sino que es impuesta en el modelo y se conserva a lo largo del proceso de formación de opinión.

Otros modelos propuestos, también basados en redes sociales, asumen agentes racionales que deciden su conducta respecto a la vacunación en el contexto de teoría de juegos [Mbah 2012, Xia 2013, Cardillo 2013]. Cada agente posee información global de la incidencia de la enfermedad en el sistema y conoce los riesgos de sufrir efectos adversos debidos a la vacuna, luego toma una decisión racional a partir de cierta función de pago. Un aspecto a destacar de todos los modelos hasta aquí citados es que asumen que la enfermedad y la información son transmitidas a través de la misma red.

En el caso del modelo propuesto en [Mbah 2012], se considera una población de agentes racionales dispuestos sobre una red compleja que deciden su postura de vacunación adoptando una estrategia óptima en función de la elegida por su entorno. De este modo, cada sujeto evalúa el costo de la vacunación contrasta-



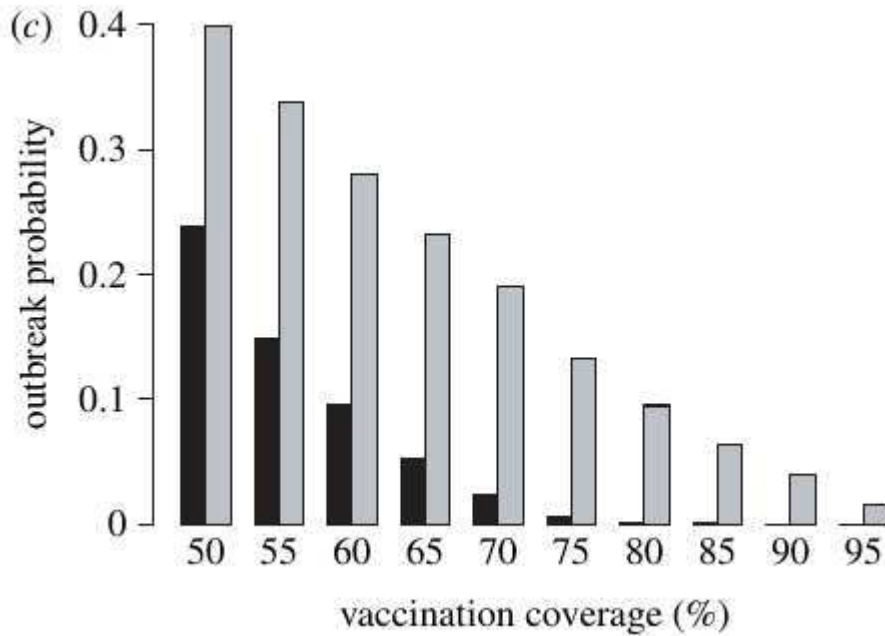


Figura 1.2: Probabilidad de brote de sarampión en función de la cobertura de la vacuna. Se comparan los casos con mecanismo de diseminación de opiniones (barras grises) y en ausencia del mismo (barras negras). Figura extraída de [Salathé 2008].

do con el riesgo percibido de infección, en función de las estrategias adoptadas por su entorno y de la incidencia de la enfermedad. Sin embargo, este modelo también considera un mecanismo de imitación social, muchas veces irracional, impuesto a una fracción de la población con el fin de dar lugar a la agrupación de individuos no-vacunados observada en los brotes ya citados. Luego, el modelo de [Mbah 2012] asume una población de sujetos que actúan racionalmente en función de su conocimiento local, paradójicamente mezclada con otra que implementa irracionalmente la imitación social para adoptar una conducta.

En esta tesis proponemos un enfoque alternativo que incorpora dos aspectos del problema hasta el momento no abordados: **i.** diseminación cultural de la conducta de vacunación y **ii.** enlaces de naturaleza personal y no-personal.

### 1.1.2. Conducta de vacunación como un rasgo cultural: Modelo de Axelrod

Las relaciones sociales entre individuos se establecen fundamentalmente a partir de un grado de afinidad por semejanza, el cual se incrementa durante el transcurso de la relación. Dicha afinidad por semejanza recibe el nombre de *homofilia* en las ciencias sociales [Lazarsfeld 1954, McPherson 2001]. Los modelos descriptos anteriormente asumen homogeneidad en todos los vínculos sociales, de modo que todos los contactos resultan igualmente influyentes. El modelo que aquí proponemos incorpora el grado de homofilia entre los individuos a través de un proceso de diseminación cultural, inspirado en el propuesto por Axelrod [Axelrod 1997] para estudiar la diseminación cultural en una sociedad representada a través de una red.

El trabajo original de Axelrod introduce un modelo basado en agentes sociales (ABM) dispuestos sobre una red cuadrada con  $N = D \times D$  nodos. Cada agente social es descrito a partir de un número de *características culturales*  $F \in \mathbb{N}$ , que representan, por ejemplo, sus opiniones respecto de política, economía, deportes, etc. A su vez, cada una de las características puede manifestarse según un número de *rasgos*  $q \in \mathbb{N}$ , por lo que toman valores en el conjunto  $\mathbb{I} = \{1, 2, 3, \dots, q\}$ . De este modo, el acervo cultural del agente  $i$  estará definido por el vector dependiente del tiempo  $\mathbf{V}^i(t) = (V_1^i(t), V_2^i(t), \dots, V_F^i(t))$ , con  $\mathbf{V}^i(t) \in \mathbb{I}^F$ . Claramente, existen  $q^F$  posibles configuraciones para el vector  $\mathbf{V}^i(t)$ .

En el contexto de este formalismo, el grado de homofilia  $h_t(i, j)$  entre dos agentes  $i$  y  $j$  se cuantifica como el número de características culturales en las que ambos comparten rasgo, por lo que  $0 \leq h_t(i, j) \leq F$ . Por ejemplo, supongamos  $\mathbf{V}^i(t) = (30, 41, 23, 45, 89)$  y  $\mathbf{V}^j(t) = (87, 41, 61, 45, 89)$ , entonces  $h_t(i, j) = 3$ . Cabe destacar que originalmente Axelrod utilizó como sustrato una red cuadrada simple con un número total de agentes  $N = D \times D$  (con  $D = 5, 10$  y  $100$ ), sin condiciones periódicas de contorno.

Finalmente, la dinámica vendrá descrita por el siguiente procedimiento algorítmico:

1. Se elige un agente  $i$  y uno de sus vecinos  $j$ , ambos uniformemente al azar.
2. **Imitación social por homofilia:** Con probabilidad

$$P(i \rightarrow j, t) = \frac{h_t(i, j)}{F} \quad (1.1)$$

el agente  $i$  interactúa con el  $j$ . En caso de concretarse la interacción, el

agente  $i$  imitará una característica  $n$  ( $1 \leq n \leq F$ ) del agente  $j$ , elegida uniformemente al azar, en la que previamente no coincidían. Es decir, si  $V_n^i(t-1) \neq V_n^j(t-1)$ , entonces

$$V_n^i(t) = \begin{cases} V_n^j(t-1) & \text{con probabilidad } P(i \rightarrow j, t) \\ V_n^i(t-1) & \text{con probabilidad } 1 - P(i \rightarrow j, t) \end{cases} . \quad (1.2)$$

3. Se incrementa el paso temporal del algoritmo  $t = t + 1$  (escala temporal arbitraria). El algoritmo se detiene cuando el sistema alcanza el equilibrio, es decir, cuando no existen transiciones posibles. En caso contrario, retorna al ítem 1.

Las componentes de los vectores  $\mathbf{V}^i(t)$  son inicialmente elegidas al azar. La dinámica del modelo conduce a una transición de fase fuera del equilibrio entre una fase multicultural y otra monocultural. La fase multicultural se caracteriza por presentar agentes aislados que difieren en sus vectores de características culturales. Mientras tanto, la fase monocultural corresponde a homogeneidad cultural, es decir, todos los agentes comparten el mismo  $\mathbf{V}^i(t)$ . El parámetro de orden de la transición se define como  $S_{max}/N$ , siendo  $S_{max}$  el tamaño del máximo dominio de convergencia cultural y  $N$  la cantidad total de agentes. En la Figura 1.3 se presenta el comportamiento característico de la transición en términos de la multiplicidad de rasgos culturales.

En [Castellano 2000] los autores muestran que la transición de fase de la Figura 1.3 para red cuadrada es continua si  $F = 2$  y de primer orden cuando  $F > 2$ . Por otra parte, las características de la transición también dependen de las propiedades topológicas de la red [Klemm 2003]. En tanto sistema que presenta una transición de fase fuera del equilibrio, el modelo de Axelrod aún depara muchos interrogantes activamente estudiados en el contexto de la sociología.

En esta tesis, proponemos un modelo basado en el de Axelrod para explicar la diseminación de la conducta de vacunación en una sociedad. Dicha conducta actuará como una característica adicional con rasgo cultural binario (vacunador o no-vacunador), que se diseminará en la red social. Las características culturales restantes serán las responsables de desarrollar homofilia entre los sujetos, proporcionando un grado de heterogeneidad adicional al modelo que actúa como catalizador del proceso de diseminación de la conducta de vacunación.

Como hemos mencionado en los párrafos precedentes, el modelo de Axelrod explica el surgimiento de consenso en una red social a través de la diseminación cultural, empleando para ello un mecanismo de imitación social. Nosotros aquí

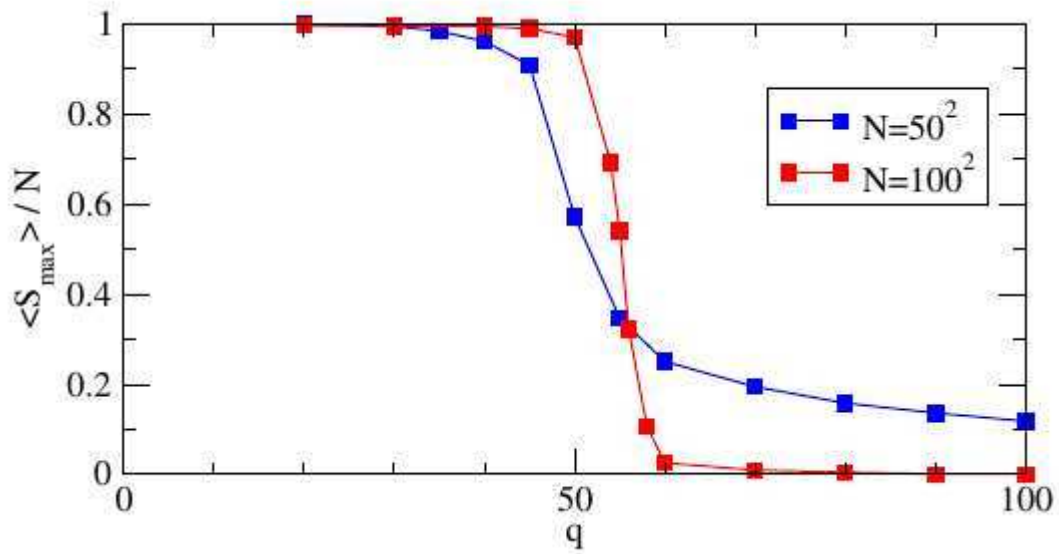


Figura 1.3: Transición de fase consenso-disenso para el modelo de Axelrod sobre una red cuadrada con  $N = 50 \times 50$  y  $N = 100 \times 100$ , en ambos casos con  $F = 10$  [Miguel 2005].

buscamos explicar el fenómeno inverso: cómo modelar el surgimiento de clusters de no-vacunados en circunstancias de alta cobertura inicial de la vacuna? Para ello incorporamos la acción de movimientos de no-vacunadores obstinados, es decir, sujetos cuya opinión respecto de la vacunación está fija y no cambia por el proceso de imitación social.

Los *movimientos anti-vacunación* generalmente carecen de argumentos científicos sólidos, aunque muy esporádicamente reciben el aval de algún miembro de la comunidad científica, como el Dr. Wakefield en la ya mencionada controversia de la vacuna TV [Wakefield 1998]. Sin embargo, es poco frecuente que los individuos no especialistas guíen su entendimiento a través del método científico, muy por el contrario, toman sus decisiones influenciados por la información que les llega desde su entorno personal o a través de medios alternativos. Esto se ve magistralmente reflejado en el relato en primera persona que la escritora norteamericana Eula Biss hace acerca de sus cavilaciones al momento de tomar la decisión de vacunar a su hijo, recientemente publicado y best seller en EEUU [Biss 2014]. Por todo ello, asumiremos que los agentes obstinados anti-vacunación pueden esparcir su postura por imitación social pesada por homofilia, independientemente de argumentaciones acerca de la racionalidad del comportamiento de los agentes sociales.

Los organismos de salud pública actúan a modo de antagonistas de los grupos anti-vacunación. Para ello hacen uso de campañas masivas de vacunación, que incluyen la amplia disponibilidad de vacunas. Sin embargo, ya hemos explicitado que la disponibilidad de vacunas no es el aspecto más relevante, aunque sí necesario, para su inserción en los países desarrollados. De modo que en esta tesis nos centraremos en la acción de campañas masivas de información positiva acerca de la vacunación por parte de los organismos de salud pública. Dichas campañas actuarán como un campo externo que incide en la postura de vacunación de los agentes sociales. Nuestro objetivo es modelar numéricamente el efecto mitigador de dichas campañas.

No pretendemos con este estudio aportar respuestas definitivas al problema de la diseminación del comportamiento social en lo que respecta a la conducta de vacunación. Nuestro objetivo es aportar nuevas herramientas de modelado que motiven nuevos interrogantes y posteriores experimentos tendientes a desvelarlos.

## 1.2. Propiedades estáticas y dinámicas de las redes sociales

Conocer las características topológicas particulares de las redes sociales es un aspecto de gran relevancia para el posterior modelado de fenómenos de propagación sobre ellas. Los trabajos fundacionales en el área dieron cuenta de algunas características distintivas de las redes sociales, tales como la propiedad de *mundo pequeño* [Milgram 1967, Watts 1998], distribuciones de grado  $P(k)$  del tipo leyes de potencias  $P(k) \sim k^{-\gamma}$  con  $\gamma \geq 2$  [Barabasi 1999b], fuerte estructura de comunidades [Newman 2002b, Newman 2004, Medus 2005], entre otras. Dichas propiedades caracterizan a las redes como objetos estáticos, resultado de la acumulación de nodos y enlaces a través de procesos dinámicos microscópicos. Desentrañar las características particulares de estos procesos ha desvelado a los expertos durante los últimos 15 años.

El modelo de crecimiento de redes más extendido hasta la fecha ha sido el de *enlace preferencial* propuesto por Barabasi y Albert [Barabasi 1999b] para explicar el origen de las distribuciones de grado del tipo ley de potencias. Su principio fundamental establece que la probabilidad que un nodo tiene de recibir nuevos enlaces es proporcional a la cantidad de enlaces que ya posee <sup>1</sup>.

---

<sup>1</sup>Este mecanismo ha sido aplicado con anterioridad en diversos contextos [Yule 1925, Simon 1955], incluso tiene un antecedente bíblico en la Parábola de los Talentos (Mateo 25:29), por lo cual también recibe el nombre de efecto Mateo.

Si bien el resultado final es el observado en redes sociales empíricas, el origen de tal mecanismo aún despierta controversias en la comunidad científica [Papadopoulos 2012, Perra 2012].

La gran cantidad de datos disponibles a partir del advenimiento de las nuevas tecnologías de la información ha permitido acceder a la evolución microscópica de las redes sociales con una resolución temporal sin precedentes [Lazer 2009]. En particular, algunos autores han empleado las comunicaciones por telefonía celular para establecer contactos entre individuos [Eagle 2006, Eagle 2009, Gonzalez 2008, Song 2010], otros han analizado las propiedades de redes de amigos online como Facebook [Ugander 2011, Bond 2012, Viswanath 2009], como así también la evolución e implicancias de la red de seguidores en Twitter [Lerman 2010, Yang 2010, Gonçalves 2011, Conover 2011]. No solo han sido registrados y analizados los contactos online sino también los contactos directos cara-a-cara entre los asistentes a reuniones sociales por medio de sensores de proximidad [Cattuto 2010, Isella 2011]. De modo que, a la luz de estos nuevos datos, ha sido posible no solo acceder a la evolución microscópica de la red, sino también a la dinámica propia de los contactos sociales.

Los experimentos ya descriptos dieron cuenta de la dinámica de contactos dando lugar a una nueva subárea dedicada al estudio de las denominadas *redes variables en el tiempo*. Uno de los modelos sencillos tendientes a reproducir la dinámica de contactos sociales asume agentes a modo de partículas activas realizando caminatas al azar con un potencial atractivo [Starnini 2013a]. Por otra parte, Nicola Perra et al [Perra 2012] propusieron un modelo de redes variables en el tiempo para el cual cada nodo es provisto de una característica intrínseca denominada *actividad social*. En cada paso temporal, los nodos se activan o no dependiendo de su actividad social y, al activarse, cada uno establece enlaces con otros  $m$  nodos elegidos al azar. Luego de cada paso temporal, todos los enlaces son borrados y el proceso se reanuda en el siguiente paso. Los autores demostraron que si la distribución de actividad social sigue una ley de potencias, entonces también lo hará la distribución de grados de la red acumulada en el tiempo.

La ciencia de las redes complejas comenzó a subdividirse entre aquellos que analizan propiedades topológicas estáticas y los que estudian la dinámica de los enlaces. Los enlaces surgen y son borrados en las redes variables en el tiempo, mientras que se asumen inalterables en la representación estática. De hecho, esta última frecuentemente es el resultado de acumular en el tiempo los enlaces que dinámicamente surgen en la representación variable en el tiempo. Sin embargo, al acumular en el tiempo los enlaces surgidos de los modelos de redes variables

ya citados, se obtienen redes que no poseen las fuertes correlaciones locales ampliamente observadas en muchas redes sociales estáticas. De modo que otro de los objetivos de esta tesis es el de proponer un modelo que permita generar redes con dinámica de enlaces, cuya versión acumulada durante una ventana temporal tenga las propiedades topológicas de las redes sociales estáticas.

### 1.3. Persistencia del sarampión

El último problema de propagación en redes sociales que trataremos en esta tesis también está relacionado con el sarampión, aunque en este caso con su probabilidad de *persistencia*. El sarampión es una enfermedad infecciosa extremadamente contagiosa. Se contrae fundamentalmente a través de las micro-gotas de Flügge que los sujetos infectados emiten al hablar o estornudar y afecta principalmente a niños en edad escolar. El período latente se prolonga por aproximadamente 8 días, en promedio, donde el enfermo desarrolla la enfermedad pero aún no la contagia, mientras que el período infeccioso tiene una duración media de 5 días en los que el enfermo es contagioso [Bailey 1954, Bailey 1956, Anderson 1991, Keeling 1997b, Conlan 2009]. Al generar inmunidad, es extremadamente poco probable que una persona contraiga sarampión dos veces en su vida.

Entre los años 1940 y 1968 el sarampión azotó gran parte del mundo, aunque existen registros particularmente exhaustivos del Reino Unido y EEUU, por ello son los casos más analizados en la literatura. En la Figura 1.4-a se muestra la cantidad total de nuevos casos reportados en Inglaterra y Gales entre 1948-1966, previos a la campaña nacional de inmunización de 1968, acumulada en intervalos de una semana. Adicionalmente, la Figura 1.4-b presenta los casos reportados para Birmingham entre 1948-1966, también acumulados semanalmente. Mientras que los gráficos de la Figura 1.4-c y 1.4-d presentan los registros acumulados mensualmente para el caso de la ciudad de Nueva York entre 1928-1964 y Copenhagen entre 1928-1968, respectivamente. El comportamiento periódico se manifiesta en el espectro de potencias como picos marcados para período anual y bianual, para los casos de Inglaterra-Gales y Birmingham (Figura 1.4-e), agregándose una componente fuerte trianual para los casos de Nueva York y Copenhagen (Figura 1.4-f).

La recurrencia del sarampión, tratada en el párrafo precedente, se encuentra vinculada con el fenómeno de persistencia. Al analizar las series temporales de distintas ciudades del Reino Unido, el matemático Maurice S. Bartlett obser-

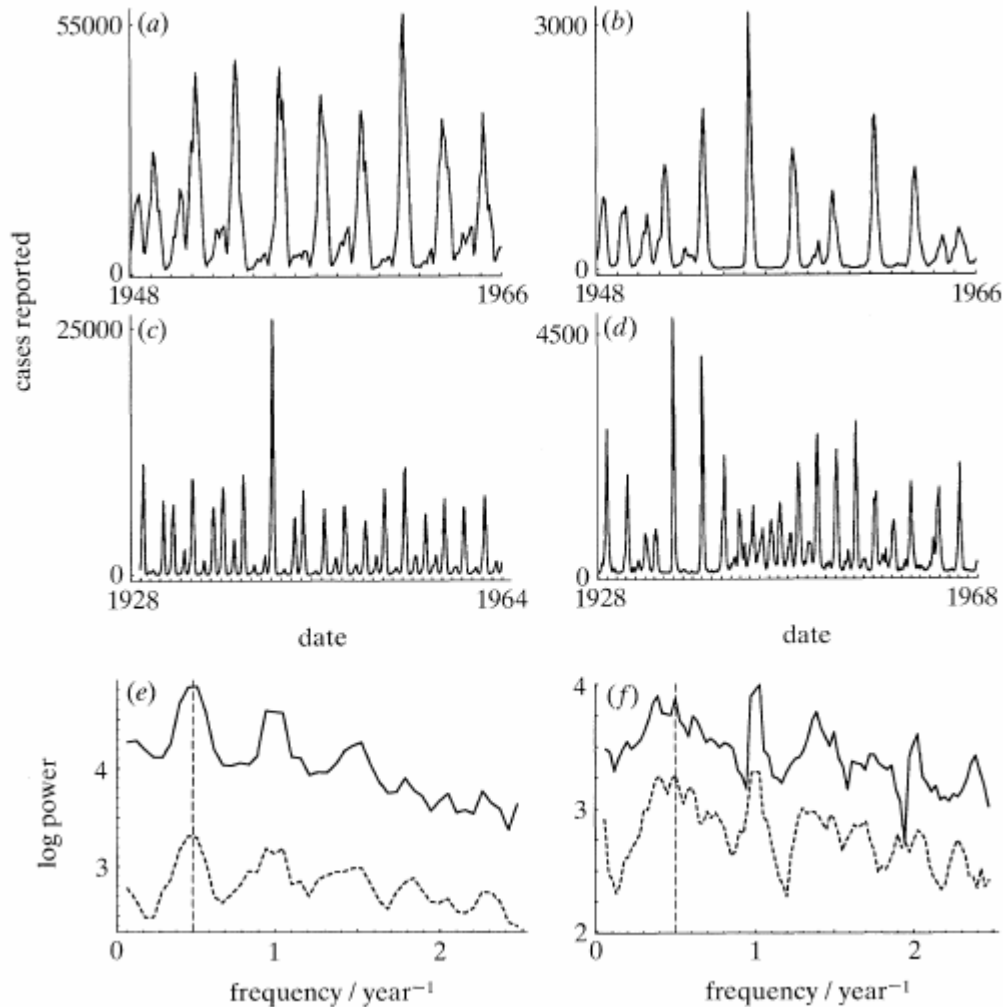


Figura 1.4: Casos reportados de sarampión acumulados semanalmente para (a) Inglaterra-Gales en 1948-1966 y (b) Birmingham en 1948-1966, y acumulados mensualmente para (c) Nueva York en el período 1928-1964 y (d) Copenhague en 1928-1968. En todos los casos se observa una fuerte periodicidad, también manifiesta en los espectros de potencia para (e) Inglaterra-Gales (línea sólida) y Birmingham (línea de trazos). (f) Espectros de potencia para la serie temporal de Nueva York (línea sólida) y Copenhague (línea de trazos). Figura extraída de [Bolker 1993].

vó que la enfermedad resultaba endémica oscilatoria en las ciudades grandes, mientras que mostraba varios períodos, de al menos 3 semanas de duración, sin casos reportados en las ciudades menos pobladas. Este hallazgo también puede apreciarse en la Figura 1.6, donde se presenta el tiempo medio entre brotes de sarampión en función de  $100/\sqrt{N}$ , siendo  $N$  la población de las ciudades de In-



*Nineteen Towns in England and Wales Ranked According to Population (Average from 1940)\**

Town	Population (Thousands)	Town	Population (Thousands)
(1) Birmingham . . . . .	1,046.0	(11) Carmarthen . . . . .	11.7
(2) Manchester . . . . .	658.0	(12) Penrith . . . . .	10.5
(3) Bristol . . . . .	415.2	(13) Ffestiniog . . . . .	7.14
(4) Hull . . . . .	269.4	(14) Brecon . . . . .	5.62
(5) Plymouth . . . . .	180.0	(15) Okehampton . . . . .	4.02
(6) Norwich . . . . .	113.2	(16) Cardigan . . . . .	3.51
(7) Barrow-in-Furness . . . . .	66.4	(17) South Molton . . . . .	3.07
(8) Carlisle . . . . .	65.0	(18) Llanrwst . . . . .	2.59
(9) Bridgwater . . . . .	21.8	(19) Appleby . . . . .	1.74
(10) Newbury . . . . .	18.1		

Figura 1.5: Población de las ciudades relevadas.

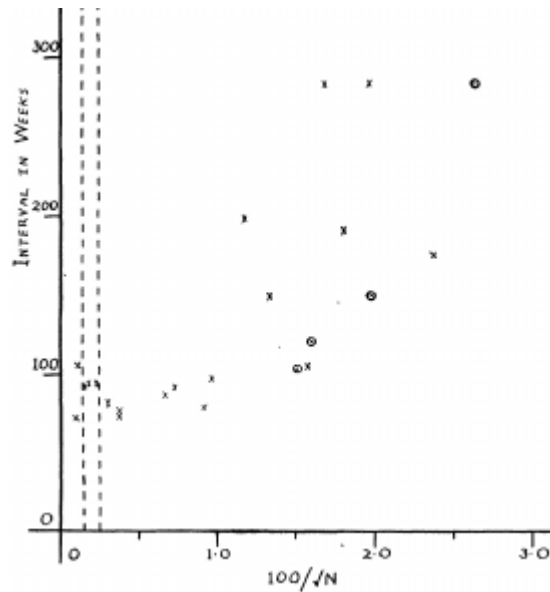


Figura 1.6: Tiempo medio entre brotes de sarampión en función de  $N^{-1/2}$  [Bartlett 1957] para las ciudades de la tabla 1.5.

glaterra y Gales indicadas en la tabla de la Figura 1.5 [Bartlett 1957]. Entonces, diremos que la enfermedad es persistente cuando su probabilidad de extinción por un período mayor a 3 semanas es nula.

Con el objeto de explicar el fenómeno de desvanecimiento o fade-out, Bartlett introduce el concepto de *tamaño crítico de comuna*  $N_C$ , a partir del cual el sarampión se torna endémico oscilatorio. Para ello, propone un modelo epidémico compartimental estocástico y discreto del tipo SIR (no considera la condición de expuesto), asumiendo que los únicos estados posibles de los agentes son susceptible (S), infectivo (I) e inmune o refractario (R) [Bartlett 1957]. La elección de un modelo estocástico discreto por sobre uno determinista se debe a que los modelos deterministas continuos admiten la posibilidad de fracciones de individuos infectados, en algunas ocasiones evitando la extinción del brote. Por otra parte,

	1	2	3	10	11	12	
A	4	5	6	13	14	15	B
	7	8	9	16	17	18	
	19	20	21	28	29	30	
C	22	23	24	31	32	33	D
	25	26	27	34	35	36	

Figura 1.7: Retículo cuadrado de  $6 \times 6$  utilizado en las simulaciones estocásticas de Bartlett [Bartlett 1957].

las fluctuaciones demográficas del tipo  $\sqrt{N}$ , propias de los modelos estocásticos, constituyen una importante fuente de extinción que tampoco surge naturalmente en los modelos deterministas. Adicionalmente, Bartlett consideró una estructura metapoblacional al incorporar un grillado cuadrado con migraciones de sujetos infectivos entre celdas adyacentes (ver Figura 1.7).

En el modelo de Bartlett el esquema temporal es discreto con intervalo de una semana, acomodándose a la frecuencia de muestreo de los registros históricos. Además de la tasa de infección  $\lambda$ , la tasa de recuperación  $\gamma$  y la de migración de sujetos infectivos entre celdas  $\varepsilon$ , el modelo requiere la incorporación permanente de nuevos susceptibles a través de una tasa demográfica  $\nu$ . Para el caso particular de la ciudad de Manchester, UK, Bartlett estima un tamaño crítico de comuna de aproximadamente 250000 habitantes, valor que ajusta adecuadamente los registros históricos para las diferentes ciudades. Consideraciones posteriores determinaron que el tamaño crítico de comuna se encuentra entre 250000-500000 habitantes.

Con el transcurrir del tiempo, tanto los modelos deterministas como los estocásticos incrementaron el nivel de detalle en la descripción considerando, por ejemplo, las fluctuaciones estacionales en el patrón de contactos debidas al ciclo lectivo [London 1973, Bolker 1993]. El modelo determinista SEIR forzado por estacionalidad más sencillo que permite obtener recurrencia periódica puede

escribirse como,

$$\begin{aligned} ds/dt &= \mu(1-s) - \beta(t)si; \\ de/dt &= \beta(t)si - (\mu + \sigma)e; \\ di/dt &= \sigma e - (\mu + \gamma)i; \end{aligned} \tag{1.3}$$

donde  $\beta(t)$  corresponde a la tasa de contacto,  $\mu$  es la tasa de natalidad/muerte (de este modo se realimenta con nuevos susceptibles),  $\sigma$  es la tasa de transición de expuesto a infectivo,  $\gamma$  la de recuperación y  $s + e + i + r = 1$  (expresado en densidades). La estacionalidad del ciclo lectivo afecta la tasa de contacto, por lo que se propone  $\beta(t) = \beta_0(1 + \beta_1 \cos 2\pi t)$ , con  $\beta_0$  y  $\beta_1$  como parámetros libres. Este modelo reproduce las oscilaciones recurrentes anuales, bianuales y trianuales observadas en las series temporales, pero al incorporar el forzante  $\beta(t)$  produce un corrimiento apreciable del tamaño crítico de comuna [Bolker 1995], resultando del orden de  $5 \times 10^6$  habitantes. Modelos posteriores incorporan distinción de cohortes por edades de los individuos, con tasas de contacto diferenciadas  $\beta_{ij}$  entre las cohortes  $i$  y  $j$  [Schenzle 1984, Bolker 1993], que también impacta en el fenómeno de persistencia.

### 1.3.1. Sensibilidad de la persistencia con la distribución de tiempos

Para resolver la inconsistencia en el tamaño crítico de comuna de los modelos que introducen estacionalidad y cohortes, en [Keeling 1997b, Keeling 1998, Lloyd 2001] se propone considerar distribuciones de tiempos infectivos más realistas que las exponenciales, asociadas con las tasas constantes de transición de los modelos tradicionales. El uso de variables aleatorias con distribución exponencial obedece a una simplificación matemática más que a una correspondencia con la realidad. Una forma de lograr distribuciones menos dispersas es reemplazando las exponenciales por distribuciones más generales del tipo Gamma:

$$P(t) = \frac{1}{\Gamma(k)\theta^k} t^{k-1} e^{-\frac{t}{\theta}}, \tag{1.4}$$

siendo  $k$  el factor de forma y  $\theta$  el de escala. Una variable con distribución Gamma puede obtenerse usando el hecho de que la suma de  $n$  variables con distribución exponencial ( $\Gamma(k = 1, \theta)$ ) da por resultado una variable con distribución  $\text{Gamma}(k = n, \theta)$  [Bailey 1975]. El método tradicional para alcanzar una distribución Gamma para el tiempo infectivo en el contexto del modelo determinista SEIR, consiste en subdividir el estado infectivo en  $n$  subestados

[Bailey 1964, Bailey 1975]:

$$\begin{aligned}
 di_1/dt &= \sigma e - (\mu + n\gamma)i_1; \\
 di_2/dt &= n\gamma i_1 - (\mu + n\gamma)i_2; \\
 &\cdot \\
 &\cdot \\
 &\cdot \\
 di_n/dt &= n\gamma i_{n-1} - (\mu + n\gamma)i_n.
 \end{aligned}
 \tag{1.5}$$

La combinación de la separación de los individuos en cohortes junto con la introducción de distribuciones de tiempo infectivo más realistas da lugar a un mejor ajuste del patrón de persistencia del sarampión [Bolker 1995, Keeling 1997b, Keeling 1998]. Sin embargo, en este último modelo, las tasas de contacto  $\beta_{ij}$  entre las diferentes cohortes constituyen los elementos de una *matriz de mezcla* y son parámetros de ajuste del modelo. Los propios autores declaran en la información suplementaria de un trabajo posterior [Conlan 2009]: “*Selection of the most appropriate mixing matrix is very much an art rather than a science.*”

### 1.3.2. Efecto de la estructura de red social en la persistencia

Tal como queda claro del análisis precedente, el patrón de persistencia del sarampión resulta muy sensible a las distintas variantes que incrementan el nivel de complejidad del modelo. Hemos destacado que al abandonar paulatinamente el paradigma de mezcla homogénea en la población, al introducir cohortes con tasas de interacción diferenciadas, se afecta el umbral de persistencia. Por ello, es razonable esperar que el tamaño crítico de comuna debería verse afectado al sustituir la hipótesis de mezcla homogénea por una red social de interacciones. Sin embargo, este camino no ha sido muy explorado, con excepción de [Verdasca 2005] donde se analiza la incidencia de la propiedad de mundo pequeño en la persistencia del sarampión sobre una red cuadrada con reconexión de enlaces.

En esta tesis nos proponemos explorar la incidencia de las correlaciones locales entre los individuos en el patrón de persistencia del sarampión sobre una red compleja.

## 1.4. Estructura de la tesis

En el Capítulo 2 daremos una rápida introducción a la teoría de redes complejas, con el objetivo de presentar las variables topológicas usuales que utilizaremos a lo largo de esta tesis. El Capítulo 3 introduce el modelo de diseminación de opiniones para la vacunación en su forma simplificada con red estática. En el Capítulo 4 generalizaremos el modelo del capítulo precedente incorporando enlaces personales y no-personales, un mecanismo de adaptación en la red social y la acción de campañas de vacunación. El Capítulo 5 desarrolla el modelo original de generación de redes sociales presentado en esta Introducción. Haciendo uso de las redes generadas mediante el modelo del capítulo anterior, en el Capítulo 6 estudiamos el efecto de las correlaciones locales de grado en la persistencia del sarampión. Por último, el Capítulo 7 contiene nuestras conclusiones finales y perspectivas de trabajo futuro.

# Breve introducción a las redes complejas

## Índice

<b>2.1. Redes estáticas: propiedades topológicas . . . . .</b>	<b>19</b>
2.1.1. Distribución de grado . . . . .	20
2.1.2. Distancia media entre nodos . . . . .	21
2.1.3. Coeficiente de clusterización . . . . .	23
2.1.4. Grado medio de primeros vecinos . . . . .	24
2.1.5. Estructura de comunas . . . . .	26
<b>2.2. Clasificación topológica de redes complejas . . . . .</b>	<b>28</b>
2.2.1. Redes aleatorias . . . . .	28
2.2.2. Redes de Mundo Pequeño . . . . .	30
2.2.3. Redes libres de escala . . . . .	31

Definiremos aquí las propiedades estructurales fundamentales de las redes complejas. Daremos una descripción introductoria de su clasificación a partir de dichas características. En este capítulo definiremos algunas de las características topológicas esenciales que serán utilizadas a lo largo de la presente tesis.

## 2.1. Redes estáticas: propiedades topológicas

El estudio riguroso del paradigma de las redes complejas cobró vigor dentro de la comunidad científica durante los últimos años del siglo pasado. En aquellos años los esfuerzos se concentraron fundamentalmente en la descripción de sus propiedades topológicas, de modo que fueron inicialmente analizadas como objetos estáticos. Las redes complejas se distinguen de las regulares debido a su

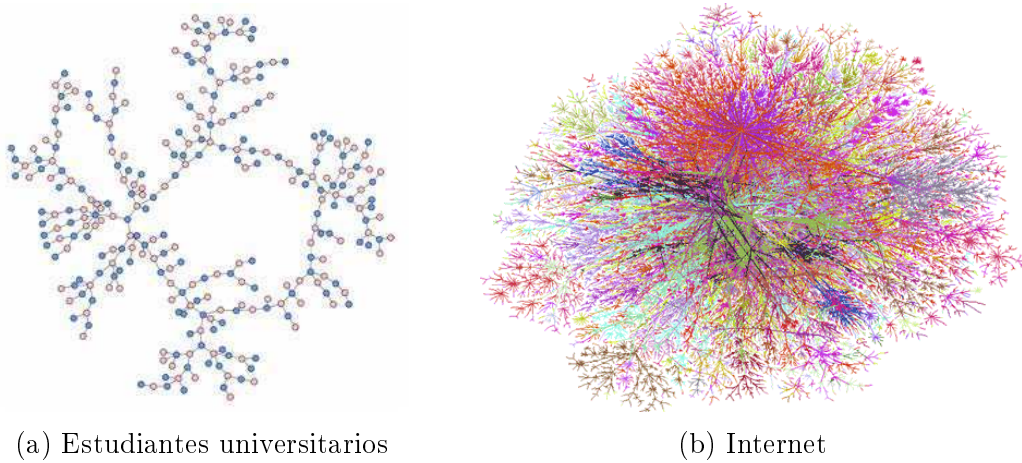


Figura 2.1: Ejemplos de redes complejas. (a) Red de relaciones sexuales entre estudiantes de una universidad norteamericana [Bearman 2004]. (b) Red de 100000 ruteadores de Internet y sus conexiones físicas [Barabási 2001].

patrón de conectividad heterogéneo. Constituyen en su esencia una representación matemática de la estructura de un sistema complejo, motivo por el cual han despertado mucho interés en la comunidad científica.

Una red o grafo  $G(N, L)$  es un conjunto de  $N$  nodos unidos a través de  $L$  enlaces (Figura 2.1). Los nodos pueden representar individuos, computadoras o regiones espaciales; mientras que los enlaces representan, por ejemplo, vínculos personales, similitudes o conexiones físicas entre los nodos. Un enlace  $l_{ij}$  entre los nodos  $i$  y  $j$  puede ser dirigido (redes dirigidas) o no-dirigido (redes no-dirigidas). Un enlace dirigido tiene un sentido prefijado, por ejemplo,  $l_{i \rightarrow j}$  corresponde a un enlace que comunica al nodo  $i$  con el  $j$ . Por otra parte, los enlaces también pueden tener un *peso* asociado que da cuenta de la intensidad del vínculo o de su frecuencia. En este último caso, la red se denomina **red pesada**. En esta tesis trabajaremos con **redes no-dirigidas** y **no-pesadas** salvo donde se declare explícitamente lo contrario.

### 2.1.1. Distribución de grado

El grado  $k_i$  correspondiente a un nodo  $i$  se define como la cantidad de primeros vecinos de  $i$ . La distribución de grados  $P(k)$  da cuenta de la probabilidad de seleccionar un nodo de grado  $k$  y es quizá la característica topológica más difundida en la clasificación de las redes complejas. Sea  $N_k$  la cantidad total de nodos con grado  $k$  en  $G(N, L)$ , entonces  $P(k)$  se define formalmente como

[Albert 2002, Newman 2003d]:

$$P(k) = \lim_{N \rightarrow \infty} \frac{N_k}{N}. \quad (2.1)$$

Las redes con distribución de grado libre de escala (**redes libres de escala**)  $P(k) \sim k^{-\xi}$  constituyen un subgrupo muy importante debido a que son ubicuas en la naturaleza [Barabasi 1999a, Albert 2002]. La World Wide Web (WWW), redes de contactos sexuales y las redes de colaboración entre científicos son algunos ejemplos de redes libres de escala. Las redes con distribución de grado del tipo Poisson ( $P(k) = \mu^k \exp(-\mu)/k!$ , siendo  $\mu$  la media de la distribución) constituyen otro caso paradigmático dado que surgen naturalmente del modelo de Erdős-Rényi (**ER**) para la generación de **redes aleatorias** [Erdős 1959, Erdős 1960, Newman 2003d].

### 2.1.2. Distancia media entre nodos

Se define la distancia  $d(i, j)$  entre dos nodos  $i$  y  $j$  como la mínima cantidad de enlaces que es necesario recorrer para comunicarlos. Luego, la distancia mínima media entre cualesquiera par de nodos  $i, j \in G(N, L)$  viene dada por

$$\bar{l} = \frac{1}{N(N-1)} \sum_{i \neq j} d(i, j), \quad (2.2)$$

donde se asume  $d(i, i) = 0$ . Es sencillo deducir que para una red regular  $\bar{l} \sim N^{1/D}$  siendo  $D$  la dimensión de la red (por ejemplo, para una red bidimensional  $\bar{l} \sim N^{1/2}$ ). Muy por el contrario, puede demostrarse que para el caso de una red aleatoria de Erdős-Rényi se tiene [Fronczak 2004]

$$\bar{l}_{ER} = \frac{\ln N - \gamma}{\ln \langle k \rangle} + \frac{1}{2} \approx \frac{\ln N}{\ln \langle k \rangle}, \quad (2.3)$$

siendo  $\gamma \simeq 0.5772$  la constante de Euler-Mascheroni y  $\langle k \rangle$  el grado medio de la red definido como

$$\langle k \rangle = \sum_k kP(k). \quad (2.4)$$

La expresión de la ecuación 2.3 muestra que  $\bar{l}_{ER}$  disminuye sensiblemente con respecto al obtenido para el caso de redes regulares. Algo semejante ocurre para



las redes aleatorias con distribución de grado  $P(k) \sim k^{-\xi}$ , para las que se obtiene

$$\bar{l} \sim \begin{cases} \ln N & \text{si } \xi > 3 \\ \frac{\ln N}{\ln \ln N} & \text{si } \xi = 3 \\ \ln \ln N & \text{si } 2 < \xi < 3 \end{cases} \quad (2.5)$$

por lo que el comportamiento en cuanto a  $\bar{l}$  es semejante al caso de las redes aleatorias **ER** cuando el exponente  $\xi > 3$ , se observa una relación de escala atípica para el caso  $\xi = 3$  [Cohen 2003, Fronczak 2004], mientras que se comporta como el doble logaritmo de  $N$  cuando  $2 < \xi < 3$  [Cohen 2003].

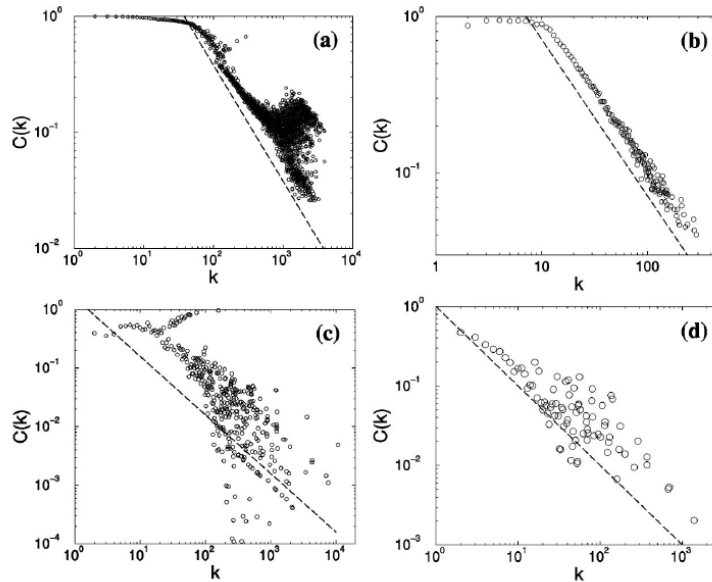


Figura 2.2: Leyes de escala para  $\bar{C}(k)$  observadas en cuatro redes reales analizadas [Ravasz 2003]. (a) Red de actores: dos actores se encuentran conectados si actuaron en la misma película de acuerdo con la base de [www.IMDB.com](http://www.IMDB.com). (b) Red semántica: conecta dos palabras del idioma Inglés si son sinónimos de acuerdo con el diccionario Merrian Webster. (c) La WWW, de acuerdo con los datos recopilados en [Albert 1999]. (d) La red de Internet como sistema autónomo: cada nodo representa un dominio, mientras que los enlaces comunicaciones físicas entre los mismos. Las líneas punteadas en cada gráfico corresponden a la ley de potencias de la ec. 2.10 con exponente  $-1$ .

### 2.1.3. Coeficiente de clusterización

Otra medida frecuentemente empleada en la caracterización topológica de las redes complejas es el coeficiente de clusterización  $C$ . Constituye una medida alternativa de la transitividad de la red, es decir, de la probabilidad de que exista un enlace entre dos nodos  $j$  y  $k$  cuando ambos poseen un vecino común  $i$ . En otras palabras, el coeficiente de clusterización suele definirse con una frase elocuente al indicar que informa de la frecuencia con la que “*los amigos de mis amigos son también mis amigos*” [Boguñá 2004].

Puede definirse el **coeficiente de clusterización local** para un nodo  $i$  como la fracción de enlaces entre sus primeros vecinos,  $j \in N_{nn}(i)$ , respecto del total de enlaces posibles entre los mismos, es decir,

$$C_i = \sum_{j,k \in N_{nn}(i)} \frac{a_{jk}}{k_i(k_i - 1)}, \quad (2.6)$$

donde  $k_i$  es el grado del nodo  $i$  y  $a_{jk}$  son los elementos de la denominada **matriz de adyacencias**  $A$  asociada a la red, tales que  $a_{jk} = 1$  si existe un enlace entre los nodos  $j$  y  $k$ , o  $a_{jk} = 0$  en caso contrario. Entonces, el valor medio del coeficiente de clusterización para toda la red puede expresarse como:

$$\bar{C} = \frac{1}{N} \sum_{i \in G} C_i. \quad (2.7)$$

El **coeficiente de clusterización medio**  $\bar{C}$  toma valores  $0 \leq \bar{C} \leq 1$ . En general,  $0.1 \leq \bar{C} \lesssim 1$  para las redes de origen social, en las que se ha observado una fuerte tendencia hacia los contactos transitivos [Kossinets 2006, Szell 2010a, Szell 2010b, Klimek 2013]. Esta última circunstancia conduce a una forma alternativa de definir  $\bar{C}$  como la probabilidad de encontrar enlaces cíclicos entre 3 nodos (triángulos). El mecanismo de formación de estos enlaces cíclicos de 3 nodos recibe el nombre de “**mecanismo de cierre triádico**” y será objeto de estudio en el Capítulo 5.

Para el caso de las redes aleatorias,  $\bar{C}$  toma valores muy pequeños dado que los enlaces transitivos implican la presencia de correlaciones locales de 3 nodos, que se ven explícitamente frustradas precisamente por la naturaleza aleatoria de la red (aunque esta afirmación es solo completamente válida cuando se considera un ensamble de realizaciones, ya que pueden presentarse correlaciones locales en realizaciones aisladas, ver [Colomer-de Simon 2012]). Puede demostrarse que  $\bar{C}$  para una red aleatoria toma el siguiente valor constante

[Newman 2003c, Serrano 2005]:

$$\bar{C} = \frac{\langle k(k-1) \rangle^2}{N\langle k \rangle^3} \simeq \frac{\langle k \rangle}{N}, \quad (2.8)$$

por lo que  $\bar{C}$  se comporta esencialmente como  $N^{-1}$ , de modo que para redes diluidas ( $\langle k \rangle \ll N$ ) alcanza valores extremadamente pequeños.

Otra forma de estudiar la clusterización de una red consiste en definir el espectro del coeficiente de clusterización medio en función del grado de los nodos  $\bar{C}(k)$ ,

$$\bar{C}(k) = \frac{1}{NP(k)} \sum_{i \in \text{Deg}(k)} C_i, \quad (2.9)$$

donde  $\text{Deg}(k)$  representa el subconjunto de nodos con grado  $k$ . Para el caso particular de las redes aleatorias  $\bar{C}(k)$  resulta independiente de  $k$  y toma el valor dado por la ec. 2.8. No obstante, en muchas redes reales analizadas se ha observado un comportamiento que responde a una función de escala del tipo [Vázquez 2002, Ravasz 2003, Klimek 2013]

$$\bar{C}(k) \sim k^{-\beta}, \quad (2.10)$$

donde el exponente toma típicamente valores  $\beta \lesssim 1$  (Figura 2.2). Es interesante notar que esta dependencia con  $k$  indica una organización jerárquica en los nodos [Ravasz 2003], que deviene en una estructura de **módulos o comunas** interconectados [Newman 2004, Medus 2005] (ver subsección 2.1.5).

#### 2.1.4. Grado medio de primeros vecinos

El grado medio de los primeros vecinos de los nodos de grado  $k$ ,  $\bar{k}_{nn}(k)$ , es otra característica topológica usual que indica la correlación de grados entre nodos. Por ello, constituye una medida indirecta de la probabilidad condicional  $P(k'|k)$  de tener un enlace hacia un nodo de grado  $k'$  cuando el nodo origen es de grado  $k$ . Luego,  $\bar{k}_{nn}(k)$  se define como [Pastor-Satorras 2001a]

$$\bar{k}_{nn}(k) = \sum_{k'} k' P(k'|k). \quad (2.11)$$

Para el caso particular de redes sin correlación de grados (por ejemplo, redes aleatorias de **ER**), usando que  $P(k'|k) = k'P(k')/\langle k \rangle$ , se obtiene  $\bar{k}_{nn}(k) = \langle k^2 \rangle / \langle k \rangle$ , resultando *independiente* de  $k$ . De modo que una dependencia de  $\bar{k}_{nn}(k)$  con  $k$

es un indicador de correlación de grados.

Existen dos casos bien delimitados [Newman 2002a]:

- (a)  $\bar{k}_{nn}(k)$  creciente: se dice que la red presenta *asortatividad en grados*, esto es, los nodos tienden a conectarse con otros de su misma clase en cuanto al grado se refiere (bajo con bajo, alto con alto);
- (b)  $\bar{k}_{nn}(k)$  decreciente: en este caso diremos que existe *desasortatividad en grados*, por lo que los nodos de grado bajo (alto) tenderán a establecer enlaces con los de grado alto (bajo).

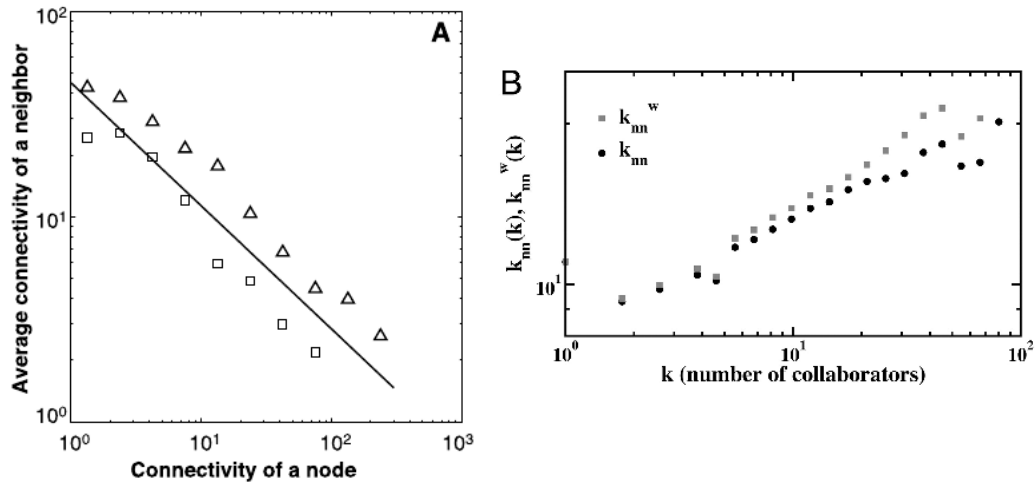


Figura 2.3: Grado medio de primeros vecinos  $\bar{k}_{nn}(k)$  para dos redes reales. (A) Red de proteínas, contemplando solo las interacciones físicas (triángulos) o las regulatorias (cuadrados) [Maslov 2002]. En este último caso, la línea sólida representa una ley de potencias  $\bar{k}_{nn}(k) \sim k^{-0.6}$ . (B) Red de colaboración entre científicos de la materia condensada, donde los enlaces representan colaboración en publicaciones, en sus versiones pesada (cuadrados) y no pesada (círculos) [Barrat 2004].

Una forma de cuantificar la asortatividad es a través del *coeficiente de asortatividad*  $r$ , definido como el coeficiente de correlación de Pearson de los grados correspondientes a los nodos en los extremos de un enlace [Newman 2002a, Newman 2003a]

$$r = \frac{L^{-1} \sum_i j_i k_i - [L^{-1} \sum_i \frac{1}{2}(j_i + k_i)]^2}{L^{-1} \sum_i \frac{1}{2}(j_i^2 + k_i^2) - [L^{-1} \sum_i \frac{1}{2}(j_i + k_i)]^2}, \quad (2.12)$$

Cuadro 2.1: Tabla de asortatividades para algunas redes analizadas en la literatura [Newman 2002a, Newman 2003a].

	Red	Tipo	$N$	$r$	$\sigma_r$
Sociales	Coautores de Física	no-dirigida	52909	0.363	0.002
	Coautores de Biología	no-dirigida	1520251	0.127	0.0004
	Coautores de Matemática	no-dirigida	253339	0.120	0.002
	Actores de películas	no-dirigida	449913	0.208	0.0002
	Relaciones de estudiantes	no-dirigida	573	-0.029	0.037
Tecnológicas	Red eléctrica	no-dirigida	4941	0.003	0.013
	Internet	no-dirigida	10696	-0.189	0.002
	WWW	dirigida	269504	-0.067	0.0002
Biológicas	Interacciones de proteínas	no-dirigida	2115	-0.156	0.010
	Red metabólica	no-dirigida	765	-0.240	0.007
	Red neuronal	dirigida	307	-0.226	0.016

siendo  $j_i$  y  $k_i$  los grados de los nodos ubicados en los extremos del enlace  $i$ , donde  $i = 1, \dots, L$ . Naturalmente, los valores que puede tomar el coeficiente de asortatividad son  $-1 \leq r \leq 1$ . En particular, para una red aleatoria (sin correlaciones de grado) es sencillo inferir que  $r = 0$ . Mientras,  $r > 0$  indica correlación de grados positiva y por ello asortatividad. En cambio,  $r < 0$  está asociado al comportamiento deasortativo en grados.

Las redes de origen social suelen presentar correlación positiva de grados y por ello son asortativas ( $r > 0$ ). Por el contrario, muchas de las redes tecnológicas y biológicas, tales como Internet, la WWW, la red eléctrica, algunas redes metabólicas y de neuronas, presentan correlación negativa de grados ( $r < 0$ ) (ver Cuadro 2.1).

Al igual que ocurre con  $\bar{C}(k)$ , en muchos casos observados  $\bar{k}_{nn}(k)$  también responde a una ley de escalas del tipo  $\bar{k}_{nn}(k) \sim k^{-\alpha}$  (ver Figura 2.3).

### 2.1.5. Estructura de comunas

El hecho de que muchas redes complejas analizadas en la práctica presenten correlaciones locales intensas, conduce a la idea de clusters de nodos más densamente conectados entre sí que hacia el resto de los nodos de la red. Esta última es precisamente la definición de *comuna* [Newman 2004, Radicchi 2004]. El particionado de redes complejas en comunas ha recibido mucha atención recientemente, debido a que constituye un modo de abordar una descripción de “grano grueso” de la red original. Esta técnica permite lidiar con las enormes

redes disponibles en la actualidad, al aportar un cambio de escala que simplifica el problema original. Cómo encontrar la mejor partición en comunas para una red dada? Esta es la pregunta que ha motivado una parte importante de la investigación en el área de las redes complejas en los últimos 10 años.

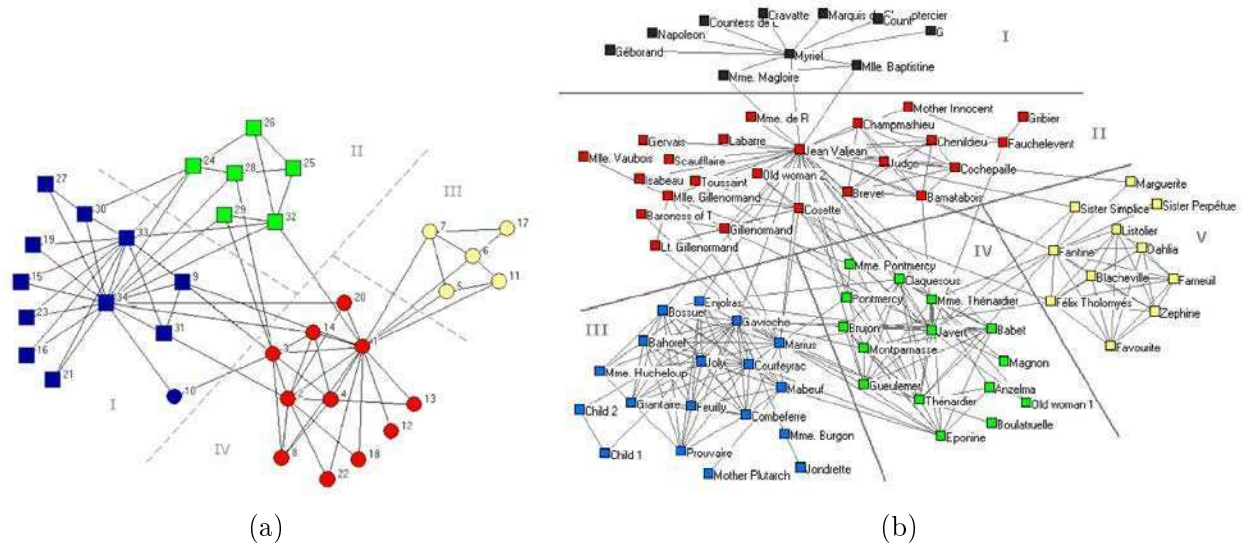


Figura 2.4: Partición óptima para dos redes típicas de la literatura. (a) Red del club de karate de Zachary [Zachary 1977]: partición óptima correspondiente a  $Q_N = 0.4197$  [Medus 2005]. (b) Partición óptima ( $Q_N = 0.546$  [Medus 2005]) para la red de coaparición de los personajes de la obra Los Miserables de Victor Hugo, según recopilación de Donald Knuth ([Knuth 1993]).

El factor de mérito que define la calidad del particionado en comunas de una red se denomina *modularidad de Newman*  $Q_N$ . La modularidad  $Q_N$  mide la densidad de enlaces entre los nodos pertenecientes a un cluster o comuna, relativa a la que se obtendría para el caso en el que los enlaces fueran conectados uniformemente al azar. Consideremos una red  $G(N, L)$  para la que se propone una división en  $K$  comunas. Si definimos  $\mathbf{e}_{ij}$  como la fracción de enlaces que unen las comunas  $i$  y  $j$ , entonces  $Q_N$  resulta:

$$Q_N = \sum_{i=1}^K (\mathbf{e}_{ii} - \mathbf{a}_i^2) \quad (2.13)$$

siendo  $\mathbf{a}_i = \sum_j \mathbf{e}_{ij}$ . De este modo,  $\mathbf{e}_{ii}$  representa la fracción de enlaces presentes en la comuna  $i$ , mientras que  $\mathbf{a}_i^2$  corresponde a la fracción que eventualmente se tendría para la misma comuna si los enlaces fueran dispuestos uniformemente al azar. Es sencillo notar que  $Q_N \simeq 0$  para las redes aleatorias, mientras que

$Q_N = 1$  es el máximo valor posible para  $Q_N$ , siendo  $0.3 < Q_N < 0.7$  los valores típicos para muchas de las redes reales analizadas en la literatura [Newman 2004]. Claramente, mientras mayor sea el valor de  $Q_N$ , mejor será la partición en comunas encontrada. En la Figura 2.4 se presentan algunos ejemplos concretos de particiones óptimas para redes previamente analizadas en la literatura.

Diversos algoritmos han sido sugeridos para encontrar el particionado óptimo. El algoritmo de Girvan-Newman o de “remoción de enlaces” [Newman 2004] consiste en ubicar y eliminar los enlaces de mayor “tránsito” o “betweenness”, entendiéndose que serán estos los que comunican a las eventuales comunas. De este modo se obtiene un subconjunto  $\{P_1, P_2, \dots, P_N\}$  de particiones accesibles, de entre todas las posibles para la red  $G(N, L)$ , del cual se escoge aquella que arroja mayor modularidad  $Q_N$  como la partición óptima. Esta clase de algoritmos solo sirven para encontrar la partición óptima de entre las que le son accesibles y no necesariamente la partición óptima global, tal como hemos demostrado en un trabajo previo a esta tesis [Medus 2005]. En dicho trabajo, hemos propuesto un algoritmo alternativo basado en la técnica de optimización global denominada “recocido simulado”, aplicada sobre el factor de mérito  $Q_N$  a fin de encontrar la partición óptima en el espacio de todas las particiones posibles de  $G(N, L)$ . Más aún, en trabajos recientes hemos propuesto una definición alternativa al factor de mérito  $Q_N$  más próxima a la definición intuitiva de comuna dada al comienzo de esta sección [Medus 2009, Dorso 2010].

## 2.2. Clasificación topológica de redes complejas

A partir de las definiciones previas es posible realizar una caracterización topológica de las redes complejas. En esta sección definiremos algunas de las familias de redes más representativas junto con sus características distintivas.

### 2.2.1. Redes aleatorias

Dentro de estas se destacan las generadas a partir del algoritmo propuesto por Erdős y Rényi [Erdős 1959]. En esencia, consiste en disponer enlaces al azar con probabilidad de ocupación constante  $p$ , sobre una red compuesta por  $N$  nodos inicialmente aislados (ver ejemplo en Fig. 2.5). De este modo, en el límite de  $N \rightarrow \infty$  manteniendo  $pN = \text{constante}$ , la distribución de grados responde a una Poisson:

$$P(k) = \frac{(pN)^k}{k!} \exp(-pN), \quad (2.14)$$

con media  $\langle k \rangle = pN$ .

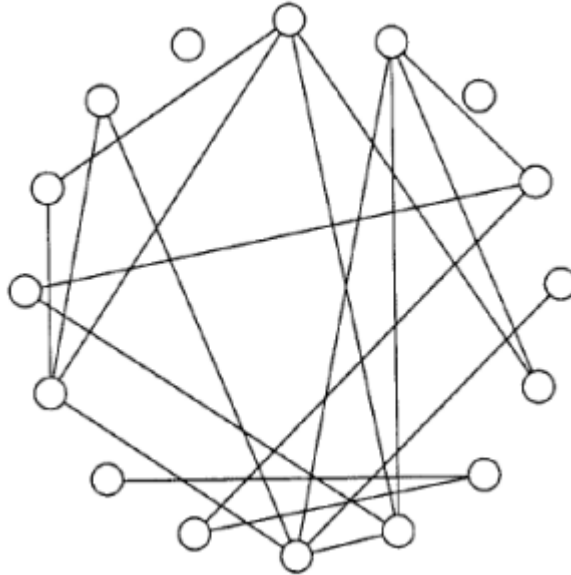


Figura 2.5: Red aleatoria de Erdős-Rényi con  $N = 16$  y  $p = 1/7$  (extraído de [Newman 2002b]).

Existen otras variantes de redes aleatorias con distribución de grado arbitraria  $P(k)$ . Estas son obtenidas mediante el denominado *modelo configuracional* [Molloy 1995, Molloy 1998, Newman 2001b], que consiste en asignar grados a los nodos con probabilidad dada por  $P(k)$ , para luego establecer contactos al azar entre los mismos. No obstante, una característica común a todas las redes aleatorias es su baja correlación de grados, que a su vez se traduce en un  $\bar{C}$  también pequeño, tal como se ha mostrado previamente en la ecuación 2.8 ( $\bar{C} \propto N^{-1}$ ). Otra característica común a las redes aleatorias es su baja distancia media entre nodos  $\bar{l} \propto \ln N$ .

Si bien es cierto que las redes aleatorias no constituyen, generalmente, un buen modelo de redes sociales o tecnológicas, son habitualmente utilizadas como una primera aproximación debido a que brindan la posibilidad de tratamiento analítico, por ejemplo, a través del formalismo de *función generatriz* [Newman 2001b].

El experimento seminal del psicólogo social Stanley Milgram publicado en el año 1967 [Milgram 1967] mostró que, en promedio, es necesario atravesar tan solo 6 enlaces para comunicar a un individuo cualquiera con otro que oficia de objetivo. Este fenómeno recibió el nombre de “seis grados de separación”, haciendo mención al concepto introducido por el autor húngaro Frigyes Karinthy en su cuento “Chain-Links” del año 1929 [Karinthy 1929]. Esta “cercanía” entre nodos



de una red es capturada por los modelos de redes aleatorias. Sin embargo, muchas de las redes sociales y tecnológicas analizadas en la práctica también presentan alto  $\bar{C}$  [Newman 2001a, Newman 2003d].

### 2.2.2. Redes de Mundo Pequeño

Dado que muchas de las redes analizadas en la literatura, en particular las redes sociales, presentan como características distintivas distancia pequeña ( $\sim \ln N$ ) y alto  $\bar{C}$  ( $\gg 1/N$ ) es necesario proponer un modelo de crecimiento de redes complejas que introduzcan estas dos propiedades. Dicho modelo fue propuesto por Duncan Watts y Steve Strogatz en 1998 [Watts 1998].

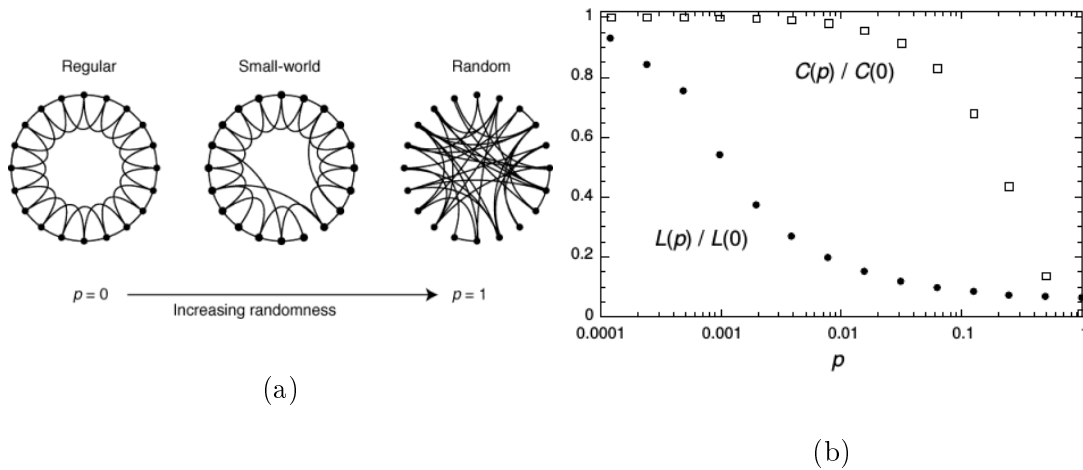


Figura 2.6: (a) Acción del proceso de reconexión con probabilidad  $p$  sobre la red del tipo anillo regular [Watts 1998]. Al aumentar  $p$ , se incrementa la cantidad de atajos, disminuyendo  $\bar{l}$  pero conservando el valor alto de  $\bar{C}$ . (b) Efecto del incremento de  $p$  sobre el clustering  $\bar{C}$  y sobre la distancia media  $\bar{L}$  también, ambos normalizados respecto de sus valores a  $p = 0$ .

La idea del modelo reside en que la distancia media pequeña es una característica natural de las redes aleatorias, mientras que la alta clusterización es propia de las redes regulares. De modo que las redes que conjugan ambas características, denominadas *redes de mundo pequeño*, deben estar a mitad de camino entre las aleatorias y las regulares. Es así como Watts y Strogatz propusieron un modelo que parte de una red del tipo anillo regular (con alto  $\bar{C}$  pero también alto  $\bar{l}$ ) sobre la que se realizan *reconexiones* al azar de enlaces con probabilidad  $p$ . Dichas reconexiones son las responsables de incorporar aleatoriedad a la red, disminuyendo el  $\bar{l}$  al constituirse en “atajos” (shortcuts). La secuencia de la Figura 2.6-(a) ejemplifica el principio fundamental del modelo.

Las redes obtenidas a partir del modelo de Watts-Strogatz presentan bajo  $\bar{l}$ , aunque conservando alta clusterización para valores pequeños de  $p$ . En el gráfico de la Figura 2.6-(b) se muestra la evolución simultánea de  $\bar{C}$  y  $\bar{l}$  en función de  $p$  para el modelo de Watts-Strogatz. Puede observarse que para una probabilidad de reconexión  $p = 0.01$  la red resultando presenta las características de mundo pequeño. Las redes de colaboración entre científicos, de coactores y de amistad, entre muchas otras redes sociales, pueden catalogarse como redes de mundo pequeño. Sin embargo, muchas de ellas presentan a su vez distribuciones de grado  $P(k)$  de cola pesada que, en principio, no son reproducidas a partir del modelo de Watts-Strogatz.

### 2.2.3. Redes libres de escala

Muchas redes observadas en la práctica presentan distribución de grado con cola del tipo ley de potencias  $P(k) \sim k^{-\gamma}$ , frecuentemente con  $2 \leq \gamma \leq 3$ , por lo que reciben el nombre de redes *libres de escala*. La red de servidores de Internet, la WWW, la red de actores de películas de IMDB, ciertas redes metabólicas y diferentes redes de colaboración entre científicos constituyen algunos ejemplos paradigmáticos de este comportamiento. Es esperable que este fenómeno emergente sea el resultado de una combinación de mecanismos microscópicos lo suficientemente sencillos y generales. El primer modelo de crecimiento de redes libres de escala, introducido por Barabasi y Albert en 1999 [Barabasi 1999b], propone una tasa de incremento del grado proporcional al grado actual de los nodos. En otras palabras, aquellos nodos con mayor grado  $k$  tendrán mayores chances de incorporar nuevos vecinos. Este mecanismo se denomina *enlace preferencial* y ha sido propuesto previamente en otros contextos para justificar distribuciones del tipo ley de potencias [Yule 1925, Simon 1955]. Puede demostrarse que a partir de una tasa de enlace preferencial lineal del tipo  $k/2L$ , se obtiene

$$P(k) \sim k^{-3}. \quad (2.15)$$

El mecanismo de enlace preferencial constituye el paradigma aceptado por la comunidad científica para generar redes libres de escala. Sin embargo, aún se discute el origen de tal mecanismo. Una de las hipótesis propuestas sostiene que el origen del enlace preferencial en las redes sociales se debe a la tendencia hacia la formación de triángulos entre nodos, es decir, a la mayor frecuencia de aparición de enlaces transitivos que tienden a cerrar triángulos de amistad [Davidsen 2002, Holme 2002]. Recientemente, ha sido propuesto un mecanismo alternativo que

justifica la aparición de enlace preferencial debido a una distribución heterogénea (power-law) de actividad social entre los individuos [Perra 2012, Muchnik 2013]. Volveremos sobre estas hipótesis más adelante, cuando propongamos un modelo de crecimiento de redes que conjuga la evolución temporal microscópica con las propiedades emergentes de la red acumulada en el tiempo.

Como ya hemos destacado en el capítulo introductorio, la estructura topológica de la red incide en sus características de propagación. Para el caso particular de una enfermedad con dinámica SIS propagándose sobre una red libre de escala con  $P(k) \sim k^{-\gamma}$ , puede verse que el umbral de epidemia responde a  $\lambda_c = \langle k \rangle / \langle k^2 \rangle$  que resulta  $\lambda_c = 0$  cuando  $\gamma \leq 3$  debido a que  $\langle k^2 \rangle \rightarrow \infty$  [Pastor-Satorras 2001b]. Este es tan solo un ejemplo que denota el fenómeno que pretendemos enfatizar en esta tesis: la importancia del estudio de la estructura de las redes complejas y su íntima relación con los fenómenos de propagación.

# Diseminación de opiniones y vacunación: modelo de red estática

---

## Índice

---

<b>3.1. Modelo de red social</b>	<b>34</b>
<b>3.2. Modelo de diseminación cultural y vacunación.</b>	<b>35</b>
3.2.1. Diseminación de la opinión respecto de la vacunación	36
3.2.2. Movimientos anti-vacunación	38
<b>3.3. Descripción algorítmica del modelo</b>	<b>38</b>
<b>3.4. Transición vacunadores a no-vacunadores</b>	<b>40</b>
<b>3.5. Análisis de brotes de sarampión</b>	<b>42</b>
<b>3.6. Análisis de clusters</b>	<b>43</b>
3.6.1. Incidencia sobre el cluster máximo	45
3.6.2. Clusters máximos: Propiedades topológicas	46
<b>3.7. Alta cobertura de la vacuna</b>	<b>47</b>
<b>3.8. Conclusiones preliminares</b>	<b>49</b>

---

El entorno social incide en la génesis del sistema de creencias de los individuos. Más aún, muchos comportamientos humanos son adquiridos por un proceso de imitación. En este capítulo analizamos el impacto del entorno social en la postura respecto de la vacunación de los individuos a través de un modelo basado en agentes sociales (ABM) dispuestos en una red compleja. Para ello, consideramos que los agentes interactúan por imitación pesada por afinidad con otros de su entorno. De este modo, introducimos el concepto de homofilia a través de un mecanismo inspirado en el modelo de Axelrod [Axelrod 1997].

Analizamos la incidencia de los grupos anti-vacunación en la diseminación de la postura de no-vacunación. Mostramos que la dinámica sencilla del mecanismo propuesto da lugar a la formación de clusters de sujetos no vacunados, por lo

tanto susceptibles de contraer y propagar una enfermedad infecciosa transmitida por contacto. Consideramos el sarampión como caso de estudio.

Como mostramos en la sección 3.7 del presente capítulo, la probabilidad de brote para el modelo aquí propuesto resulta notablemente superior a la esperada para modelos que asumen distribución uniforme de los agentes infectados, afectando de este modo el efecto de inmunidad de grupo.

### 3.1. Modelo de red social

Las redes complejas constituyen una representación del entramado social al considerar las heterogeneidades de los contactos entre los agentes sociales. Los nodos de la red representan a dichos agentes, mientras que los enlaces dan cuenta de los vínculos entre los mismos. Estos vínculos pueden indicar lazos familiares, de amistad, relaciones laborales, simple proximidad espacial, etc. De este modo, cada agente se encuentra inmerso en un contexto social heterogéneo que luego influirá en su cosmovisión.

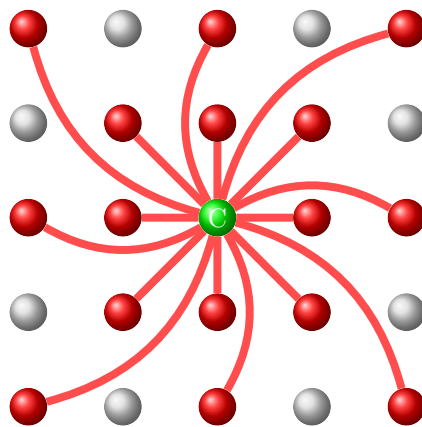


Figura 3.1: Red cuadrada estática de contactos entre agentes sociales. Considerando el nodo central C (verde), sus enlaces hacia los nodos restantes (rojos), pueden transmitir tanto opiniones como una enfermedad infecciosa.

En el contexto de nuestro modelo, asumimos que los agentes sociales representan grupos familiares donde los padres deben decidir acerca de la vacunación

de su único hijo/a. En esta implementación, los agentes sociales están dispuestos en una red cuadrada de  $N = 50 \times 50$  nodos con enlaces a primeros y segundos vecinos ( $3N+2N$ ) más el agregado de enlaces adicionales para completar una red con grado uniforme  $k = \langle k \rangle = 16$  (Figura 3.1). En este capítulo solo consideraremos enlaces homogéneos y estáticos, mientras que en el próximo introduciremos contactos personales (transmiten tanto opiniones como enfermedades) y no-personales (solo transmiten opiniones), así como también un mecanismo de reconexión adaptativa de enlaces.

## 3.2. Modelo de diseminación cultural y vacunación.

Consideramos una condición inicial sin grandes brotes epidémicos recientes de sarampión. También consideramos una política de vacunación voluntaria, tal es el caso del Reino Unido, o una donde sea posible eximirse de la misma esgrimiendo una objeción de conciencia, como sucede en algunos estados de EEUU. En este escenario, la decisión de vacunar a no a sus hijos recae en los padres. Asimismo, las últimas grandes epidemias de sarampión en estos países desarrollados datan de la década de 1960, por lo que los padres actuales carecen de registros vivenciales de sus consecuencias en su memoria inmediata. Ante estas circunstancias de baja percepción de riesgo inminente de contagio, nuestra hipótesis es que la decisión estará fundada mayormente en un proceso de imitación social a partir de las posturas que los agentes decisores recojan de su entorno.

Los modelos más sencillos consideran una opción binaria respecto de la vacunación, que cada agente adopta por simple opinión mayoritaria de su entorno [Salathé 2008]. Modelos más complejos asumen agentes racionales que, además de implementar el proceso de imitación social, evalúan el riesgo de contagio y los efectos nocivos de la vacuna en el contexto de teoría de juegos [Fu 2011, Xia 2013, Cardillo 2013]. En ambos casos se considera que la incidencia de los contactos es homogénea, de modo que todos los vecinos inmediatos son igualmente influyentes.

La similitud en rasgos culturales, étnicos o generacionales actúa a modo de catalizador de las relaciones sociales. El grado de afinidad o similitud entre individuos es un concepto que ha sido introducido en las ciencias sociales, donde recibe el nombre de *homofilia* [Lazarsfeld 1954, McPherson 2001]. El modelo aquí propuesto incorpora este grado adicional de heterogeneidad debido a la homo-

filia entre agentes. Para ello, implementamos una variante original del modelo de agentes propuesto por el politólogo y matemático Robert Axelrod en 1997 [Axelrod 1997], quien lo ideó con el objetivo de explicar el origen del consenso social a través de la diseminación cultural y la formación de dominios de convergencia cultural (ver 1.1.2).

### 3.2.1. Diseminación de la opinión respecto de la vacunación

Los modelos más extendidos en la actualidad consideran a la imitación social como el mecanismo responsable de la diseminación de la postura respecto de la vacunación. Muchos de ellos proponen un contexto de teoría de juegos que comprende agentes racionales con información local (de su entorno) y/o global, quienes deciden su posición óptima respecto de la vacunación de entre dos posibilidades: vacunar o no vacunar [Fu 2011, Mbah 2012, Xia 2013, Cardillo 2013]. Naturalmente, aquellos agentes que optan por vacunar reciben el beneficio de la inmunidad pero asumen, por ejemplo, el costo de los efectos secundarios de la vacuna. Elegir no vacunar evita los costos asociados con la vacuna, pero implica el riesgo de contagio de sarampión. En el contexto de estos modelos, la estrategia óptima de cada agente dependerá entonces del costo de la vacunación y del riesgo latente de contagio. No obstante, cuando se considera la estructura de red social, la teoría de juegos por sí sola no es suficiente para justificar la agrupación de individuos no-vacunados, responsable de algunos de los brotes recientes en países desarrollados. Es por ello que en [Mbah 2012] se recurre a una combinación de individuos racionales que buscan optimizar su beneficio, con otros que eligen un mecanismo, muchas veces irracional, de imitación social que conduce a la formación de clusters de no-vacunados.

Nuestro enfoque se diferencia notablemente de los precedentes al asumir que la conducta de vacunación se propaga como una de las características culturales en una variante original del modelo de Axelrod. De este modo, nos alejamos del paradigma de agentes racionales en buscar de estrategias óptimas, conjugados con agentes irracionales que practican la imitación social.

El acervo cultural de cada agente social, en este caso considerando el conjunto de sujetos decisores que integran la familia en cuestión, vendrá definido por  $F + 1$  características, donde la característica adicional representa la postura binaria de vacunación. Luego, la multiplicidad de rasgos culturales asociada con la postura respecto de la vacunación será  $q = 2$ , mientras que el resto de las características

tendrán multiplicidad homogénea  $q \geq 2$ . Al margen de esta distinción, la postura respecto de la vacunación se involucra en la dinámica de adaptación por imitación como una característica cultural más. El grado de homofilia entre los agentes  $i$  y  $j$  se define como la cantidad de características en las que coinciden  $h_t(i, j)$  a tiempo  $t$ , al igual que para el modelo de Axelrod. Las  $F$  características culturales, excluyendo la postura respecto de la vacunación, se incorporan en el modelo con el fin de desarrollar homofilia heterogénea entre los agentes. De este modo, la diseminación de las diferentes opiniones, en particular la de vacunación, se dará con mayor probabilidad entre aquellos agentes conectados que presenten alto grado de homofilia.

El objetivo es que los agentes sociales adapten su postura respecto de la vacunación imitando a aquellos de su entorno con los que tienen alto grado de homofilia. Esto es: buscamos que la postura respecto de la vacunación se imite debido a una afinidad previa en otras características culturales. Para ello, incorporamos un *umbral de homofilia* ( $\kappa$ ) a partir del cual dos agentes comenzarán a interactuar. En nuestro caso, elegimos  $\kappa = 2$ , aunque también hemos considerado  $\kappa = 3$  obteniendo cualitativamente idénticos resultados. Con  $\kappa = 2$  evitamos la interacción social entre dos agentes que solo comparten la postura binaria de vacunación y, con ello, la casi segura evolución hacia un equilibrio absorbente de consenso.

Luego, la probabilidad de interacción  $P(i \rightarrow j, t)$  entre los agentes  $i$  y  $j$  vendrá dada por:

$$P(i \rightarrow j, t) = \begin{cases} 0 & \text{si } h_t(i, j) < \kappa \\ \frac{h_t(i, j)}{F+1} & \text{si } h_t(i, j) \geq \kappa \end{cases} \quad (3.1)$$

Siguiendo el modelo introducido por Axelrod, definimos un vector de opiniones  $\mathbf{V}^i(t) = (V_1^i(t), V_2^i(t), \dots, V_{F+1}^i(t))$  con  $\mathbf{V}^i(t) \in \{0, 1\} \times \mathbb{I}^F$ , siendo la componente  $V_{F+1}^i(t)$  la correspondiente a la postura respecto de la vacunación, de modo que:

$$V_{F+1}^i(t) = \begin{cases} 0 & \text{si el agente } i \text{ es } \mathbf{vacunador} \\ 1 & \text{si el agente } i \text{ es } \mathbf{no-vacunador}. \end{cases} \quad (3.2)$$

Luego, el proceso de adaptación por imitación social del vector de opiniones  $\mathbf{V}^i(t)$  será el descrito en la ecuación 3.3 del capítulo 1,

$$V_n^i(t+1) = \begin{cases} V_n^j(t) & \text{con probabilidad } P(i \rightarrow j, t+1) \\ V_n^i(t) & \text{con probabilidad } 1 - P(i \rightarrow j, t+1) \end{cases} . \quad (3.3)$$



### 3.2.2. Movimientos anti-vacunación

Uno de nuestros objetivos es analizar la incidencia de los grupos anti-vacunación en la cobertura de la vacuna TV. Estos están representados por un conjunto de agentes  $S_{stub}$  que adoptan una postura de no-vacunación de forma obstinada. Luego, la postura respecto de la vacunación de todos los agentes pertenecientes al movimiento anti-vacunación será

$$V_{F+1}^i(t) = 1 \quad \forall i \in S_{stub} \quad (3.4)$$

para todo tiempo  $t$ . Estos agentes son dispuestos al azar en la red social y, dado que constituyen un grupo claramente minoritario, de aquí en más asumiremos que involucra solo al 1 % de la población total. Consideraremos una única configuración espacial inicial de no-vacunadores obstinados con el fin de poder comparar distintas realizaciones del modelo, independizándonos así de su distribución inicial. Dado que con un 1 % de agentes obstinados aún estamos alejados del umbral de percolación de la red considerada, la probabilidad de formación de clusters grandes de agentes pertenecientes a  $S_{stub}$  es despreciable y no tendrá incidencia en la evolución del mecanismo de diseminación.

## 3.3. Descripción algorítmica del modelo

Para culminar esta introducción del modelo, daremos su descripción algorítmica completa especificando las condiciones iniciales y los aspectos computacionales relevantes. Dado que buscamos reproducir la situación inicial realista de amplio consenso en favor de la vacunación, comenzaremos asumiendo un escenario donde todos los agentes son vacunadores ( $V_{F+1}^i(t) = 0$ ), salvo el 1 % de no-vacunadores obstinados pertenecientes a  $S_{stub}$ . Estos últimos son elegidos uniformemente al azar y satisfacen  $V_{F+1}^j(t) = 1 \quad \forall j \in S_{stub}$ , para todo tiempo  $t$ .

En todo los casos hemos fijado  $F = 10$ . Para explicar el motivo de esta elección es suficiente con remitirse a las secciones de los periódicos: La Nación de Argentina cuenta con 10 secciones fijas diariamente en su versión online, mientras que New York Times (EEUU) también cuenta con 12 secciones diarias fijas y 8 variables, algo semejante ocurre con El País de España, siempre refiriéndonos a sus versiones online. Este hecho refleja que los temas de interés general son relativamente escasos: política local e internacional, economía y negocios, deportes, espectáculos y arte, son algunos de ellos. De modo que  $F = 10$  refleja adecuadamente la cantidad de temas respecto de los cuales los agentes sociales

manifiestan opinión. Finalmente, el umbral de homofilia se estableció en  $\kappa = 2$ , tal como se señaló con anterioridad.

Consideramos una red social, anteriormente descrita, compuesta por  $N = 2500$  agentes (red cuadrada  $3N+2N$  de  $50 \times 50$  con 8 enlaces adicionales por nodo para completar  $\langle k \rangle = k_i = 16 \forall i \in \{1, \dots, N\}$ , como se muestra en la Figura 3.1) y  $L = 20000$  enlaces. Salvo la postura respecto de la vacunación, el resto de las características de los agentes son inicialmente elegidas uniformemente al azar en  $\{1, 2, 3, \dots, q\}$ . Una vez definidas las condiciones iniciales del modelo, se comienza con la actualización del sistema en cada paso (aquí simbolizado como un tiempo adimensional discreto  $t$ ) del siguiente algoritmo:

1. Un agente fuente  $i$  es elegido uniformemente al azar, mientras que uno de sus primeros vecinos  $j$  es elegido como target, también uniformemente al azar.
2. Se obtiene  $h_t(i, j)$  y la probabilidad de interacción  $P(i \rightarrow j, t)$ . Luego, se elige al azar una característica  $n$  en la que  $i$  y  $j$  no coincidan.
3. Si  $n \neq F + 1$ , es decir que la característica no es la postura respecto de la vacunación, el agente  $i$  sigue el proceso de adaptación descrito en la ec. 3.3.
4. Si  $n = F + 1$ , caben dos alternativas:
  - $i \in S_{stub}$ , en este caso  $i$  conserva su postura respecto de la vacunación ( $V_{F+1}^i(t+1) = V_{F+1}^i(t)$ )
  - $i \notin S_{stub}$ , en cuyo caso el agente sigue el proceso de adaptación descrito en la ec. 3.3.
5. Condición de detención: consideramos dos posibles condiciones de detención:
  - i. El algoritmo se detiene cuando alcanza una configuración absorbente, es decir, cuando no son posibles transiciones adicionales.
  - ii. En los casos donde prevalece la opinión de vacunación, el accionar de los agentes  $S_{stub}$  puede dar lugar a pequeñas fluctuaciones intermitentes en la postura respecto de la vacunación en su entorno. Es por ello que fijamos como condición de detención adicional la imposibilidad de realizar transiciones para todas las características culturales,

exceptuando pequeñas fluctuaciones en la postura respecto de la vacunación.

6.  $t \rightarrow t + 1$ . Retorna al paso 1.

Con respecto a la condición de estabilidad del ítem 5, puede darse el caso en el que un agente  $i$  no-vacunador obstinado resulte rodeado por otros vacunadores con los que acuerda en todo su acervo cultural, salvo en la postura frente a la vacunación. Ocurre entonces que los agentes pertenecientes al entorno de  $i$  ( $nn(i)$ ) podrán imitar intermitentemente las postura de no-vacunación de  $i$ , para luego retornar a la postura predominante de su entorno.

### 3.4. Transición vacunadores a no-vacunadores

El observable que nos ocupa es la fracción neta de vacunadores y no-vacunadores en la red social. En la Figura 3.2-A se presenta la fracción de no-vacunadores, promediada sobre 1000 realizaciones, como función de  $q$ .

Para valores pequeños de  $q$ , los agentes no-vacunadores obstinados logran diseminar su ideario, alcanzando un amplio consenso no-vacunador en la población, siendo este el estado absorbente monocultural del sistema. Sin embargo, al tomar valores próximos a  $q \approx 170$ , el consenso no-vacunador comienza a ceder, las interacciones sociales resultan más difíciles de establecerse debido al incremento del número de rasgos culturales, dando lugar a dominios de consenso no-vacunador alrededor de los agentes obstinados, coexistiendo con dominios integrados por vacunadores. Este hecho se manifiesta en la Figura 3.3, donde se muestra un gráfico de densidad correspondiente al promedio sobre 500 realizaciones del modelo. La intensidad de color rojo da cuenta de la fracción de realizaciones en las que cada agente alcanzó el equilibrio como no-vacunador.

La transición de fase fuera del equilibrio presentada por el sistema es característica del modelo de Axelrod [Klemm 2003]. La fase de consenso no-vacunador corresponde al estado absorbente monocultural donde todos los agentes comparten los mismos rasgos culturales. En cuanto a la fase de consenso vacunador (exceptuando a los agentes pertenecientes a  $S_{stub}$ ), si bien existe consenso en cuanto a la opinión de vacunación, no ocurre lo propio respecto del resto de las características culturales, por lo que la fase corresponde a polarización cultural total. En este último caso, el consenso vacunador proviene de la configuración inicial, para la cual el 99% de los agentes comienzan el proceso con opinión pro-vacunación.

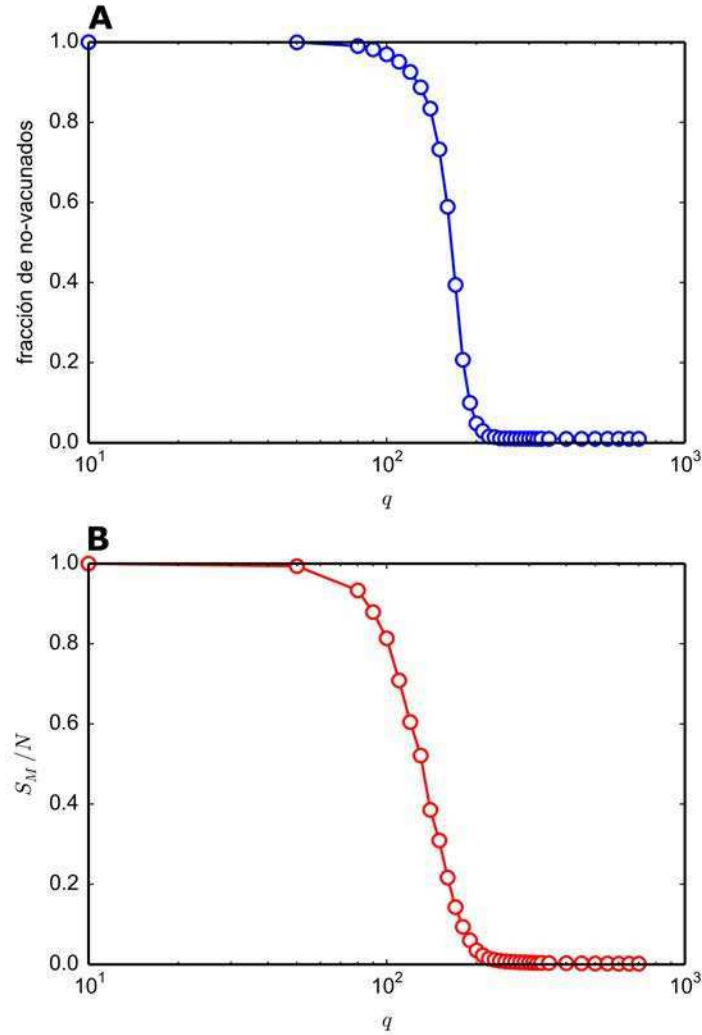


Figura 3.2: (A) Transición de consenso no-vacunador a vacunador en función de la multiplicidad de rasgos culturales  $q$ . (B) Tamaño normalizado del cluster máximo monocultural  $S_M/N$  para las primeras  $F$  opiniones en función de  $q$ .

En la Figura 3.2-B mostramos el tamaño normalizado correspondiente al dominio mono-cultural de mayor población  $S_M/N$  (con  $S_M$  el tamaño del cluster máximo) considerando solo las primeras  $F$  opiniones, salvo la de vacunación, en función de  $q$ . Puede observarse la transición de consenso a polarización cultural total mencionada en el párrafo anterior. Con lo cual, se evidencia que la fracción final de no-vacunadores se encuentra condicionada por la posibilidad de formación de dominios de consenso cultural características del modelo de Axelrod.

Tal como hemos señalado en la introducción (ver Capítulo 1), la formación de clusters o dominios de no-vacunadores es responsable de los brotes recientes

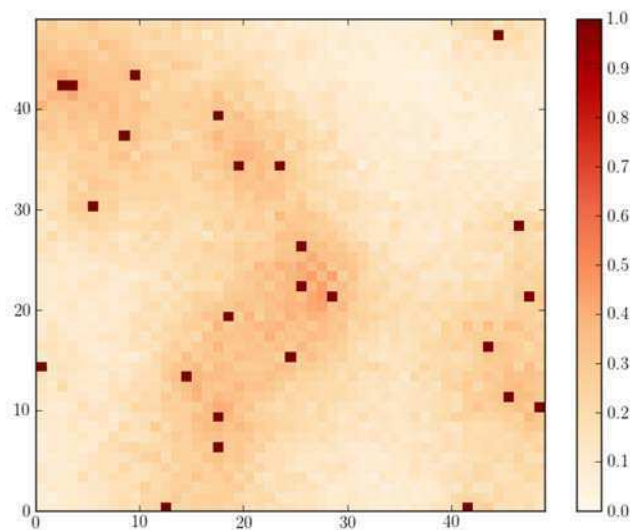


Figura 3.3: Mapa de calor realizado a partir de 500 realizaciones del modelo de diseminación cultural. La intensidad de rojo indica la fracción de realizaciones en las que el agente en cuestión resultó no-vacunador. Los sitios más oscuros corresponden a los agentes obstinados  $S_{stub}$ .

de sarampión en países con amplia cobertura de la vacuna TV. Mostramos aquí que esto se debe a que la decisión de no vacunar se disemina predominantemente en grupos que comparten un acervo cultural. No obstante, los clusters de no-vacunadores coexisten con otros de individuos vacunados. Dicha situación de coexistencia de dominios con distintas posturas se corresponde con el comportamiento observado en la región crítica de transición en las Figuras 3.2-A y 3.2-B. Es por ello que en lo que resta del presente capítulo fijaremos  $q = 170$  con el objeto de situarnos en la región crítica con coexistencia de ambas posturas de vacunación.

### 3.5. Análisis de brotes de sarampión

En la sección precedente se observó la coexistencia de dominios de vacunadores y no-vacunadores originados como resultado del proceso de diseminación cultural. Ahora veremos cómo esto impacta en el tamaño medio de un eventual brote de sarampión. En particular, nos interesa contrastar nuestros resultados con la hipótesis nula que consiste en asumir distribución al azar de los individuos no-vacunados en la red social. Llamaremos a este enfoque alternativo *modelo*

*de percolación*. Este último es el que naturalmente se considera en los modelos de campo medio, aunque también suele asumirse en los modelos de redes complejas más sencillos. Consiste en distribuir al azar sobre la red la misma cantidad de no-vacunados obtenida del modelo de diseminación cultural. Es importante notar que tanto para el modelo de diseminación cultural como para el de percolación, consideramos las mismas posiciones de los agentes  $S_{stub}$  con el objeto de no introducir factores exógenos que deriven en sesgos indeseados.

Como se describió en la Introducción, el sarampión puede representarse por un modelo compartimental del tipo SEIR, es decir, los agentes solo pueden pertenecer a una de las siguientes categorías: susceptibles (S), expuestos (E), infectivos (I) o refractarios (R). Los detalles del modelo epidemiológico utilizado pueden encontrarse en el Apéndice A. Es importante destacar que solo los individuos no-vacunados pueden contraer la enfermedad (asumimos que la vacuna tiene una eficiencia del 100%) Por otra parte, recordemos que cada agente representa a una familia con un hijo, de modo que este último es el único que puede contraer la enfermedad (asumimos que los padres ya contrajeron el sarampión en su infancia o, en su defecto, han sido vacunados).

El brote comienza con un único infectado elegido al azar de entre los agentes susceptibles, es decir, los no-vacunados. Se efectuaron 10000 realizaciones del modelo de propagación del sarampión para cada una de las 500 configuraciones de equilibrio obtenidas del mecanismo de diseminación cultural. En la Figura 3.4 se presenta el tamaño medio del brote (i.e., la media de la cantidad total de infectados) en función de la población total de no-vacunados, tanto para nuestro modelo como para la aproximación de percolación. Puede verse claramente que nuestro modelo da lugar a un tamaño medio de brote superior al de percolación. Este efecto se ve magnificado para poblaciones pequeñas de no-vacunados.

El sarampión, como toda enfermedad infecciosa por contacto entre humanos, se transmite fundamentalmente entre individuos conectados. Es por ello que la raíz del comportamiento mostrado en la Figura 3.4 debe buscarse en la estructura de clusters de no-vacunados. Por ello, en la siguiente sección estudiaremos la distribución de tamaños de clusters.

### 3.6. Análisis de clusters

Naturalmente, el tamaño máximo de un brote de sarampión estará condicionado por la cantidad de agentes no-vacunados que se encuentren físicamente conectados mediante un enlace, o por una sucesión de los mismos, con el infec-

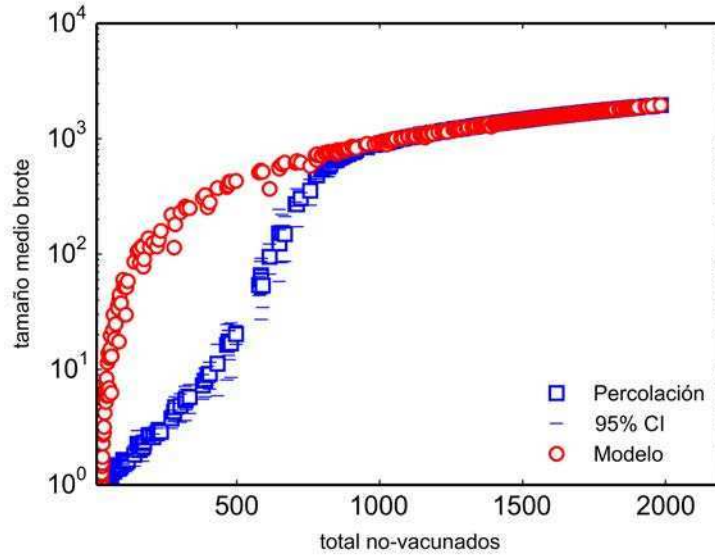


Figura 3.4: Tamaño medio del brote de sarampión sobre 10000 realizaciones del modelo de propagación, en función de la población de no-vacunados obtenidos mediante el modelo de diseminación cultural (círculos rojos) y para la aproximación de percolación (cuadrados azules).

tado inicial. Esto es, el tamaño del brote será a lo sumo la población del cluster de no-vacunados al que pertenece el infectado inicial. De esto se infiere que la distribución de tamaños de clusters juega un rol fundamental para develar el origen de los resultados de la sección anterior.

En la Figura 3.5 se muestra la distribución de tamaños de clusters de no-vacunados para el acumulado de las 500 realizaciones del proceso de diseminación cultural. Los resultados son comparados con los correspondientes al modelo de percolación, considerando en cada caso exactamente las mismas poblaciones finales de no-vacunados. Para el caso del modelo de percolación, se efectuaron 1000 realizaciones por cada una de las 500 poblaciones finales de no-vacunados de nuestro modelo. Del análisis visual de la Figura 3.5 puede concluirse que los clusters de mayor tamaño son mucho más frecuentes para el modelo de diseminación cultural que para el modelo de percolación equivalente. Diremos entonces que nuestro modelo es propenso al clustering de no-vacunados como producto de la dinámica del proceso de imitación social por homofilia. Una forma de cuantificar las diferencias entre ambos modelos es a través del cálculo del estadístico  $z_S$ , definido como:

$$z_S(b) = \frac{\rho(b) - \hat{\rho}_{perco}(b)}{\hat{\sigma}_{perco}(b)} \quad (3.5)$$

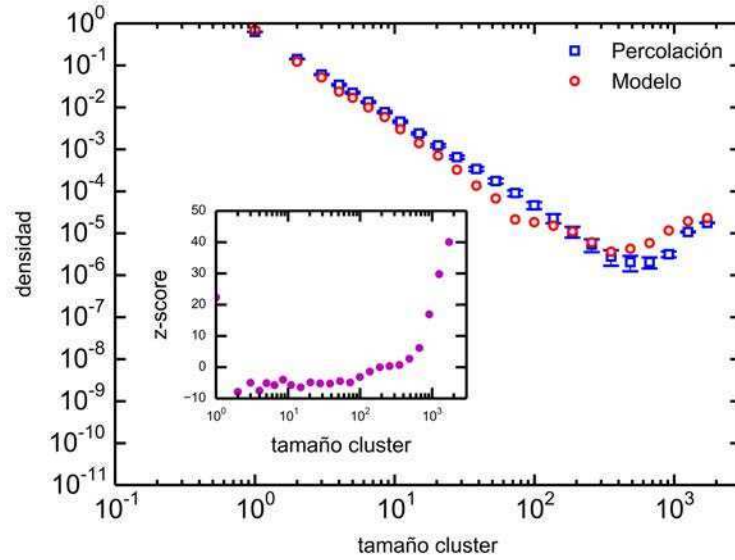


Figura 3.5: Distribución de tamaños de clusters para el modelo de diseminación cultural (círculos rojos) y para la aproximación de percolación (cuadrados azules). Inset: estadístico  $z$  como función del tamaño de cluster. Se observa la preeminencia de clusters de gran tamaño en el modelo de diseminación cultural respecto del modelo nulo de percolación.

siendo  $\rho(b)$  la densidad correspondiente al bin  $b$  dada por el modelo, mientras que  $\hat{\rho}_{perco}(b)$  y  $\hat{\sigma}_{perco}(b)$  corresponden, respectivamente, a la media y la desviación standard de la densidad para las realizaciones de percolación en el mismo bin  $b$ . En el inset de la Figura 3.5 se observa que  $z_S$  aumenta con el tamaño de cluster, por lo cual se concluye cuantitativamente la tendencia a la clusterización para nuestro modelo al compararlo con la aproximación de percolación.

### 3.6.1. Incidencia sobre el cluster máximo

En la sección precedente observamos que nuestro modelo presenta mayor tendencia a la formación de clusters que aquel que asume distribución aleatoria de los no-vacunados (modelo de percolación). Cabe ahora preguntarnos si es que se observarán diferencias en la incidencia acumulada del sarampión cuando éste se propaga sobre el cluster máximo para nuestro modelo y sobre uno de idéntica población, pero obtenido a través del modelo de percolación. Para ello seleccionamos realizaciones de nuestro modelo y del modelo de percolación donde el tamaño del cluster máximo cae en el mismo bin de la distribución. Luego, comparamos la cantidad media de infectados sobre 10000 realizaciones del modelo



de propagación de sarampión en cada uno de ellos. La Figura 3.6 muestra que la cantidad media de infectados sobre los clusters máximos correspondientes al modelo de diseminación cultural es superior a la obtenida mediante la aproximación de percolación sobre clusters de tamaño equivalente, efecto que nuevamente se ve magnificado para poblaciones pequeñas por debajo del umbral de percolación.

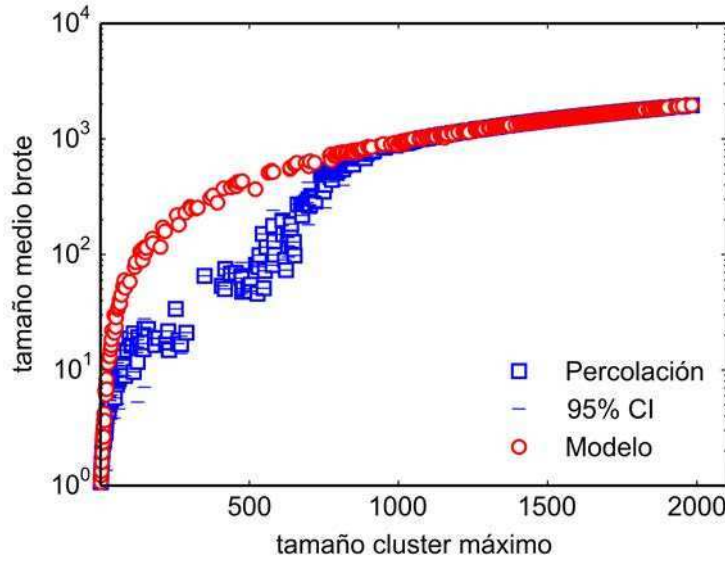


Figura 3.6: Tamaño medio del brote de sarampión sobre 10000 realizaciones del modelo de propagación, en función de la población de no-vacunados correspondientes a los clusters máximos obtenidos mediante el modelo de diseminación cultural (círculos rojos) y para la aproximación de percolación (cuadrados azules).

En la sección precedente mostramos que a igual cantidad de no-vacunados, el modelo de diseminación cultural arroja mayor incidencia de sarampión que la correspondiente aproximación de percolación. Vinculamos este resultado a la mayor presencia de clusters en el primero. No obstante, en la presente sección observamos que el modelo de diseminación cultural también presenta mayor incidencia de sarampión que el de percolación cuando se comparan clusters máximos de igual tamaño.

### 3.6.2. Clusters máximos: Propiedades topológicas

La discrepancia señalada en la sección anterior tiene su origen en las características topológicas de los clusters máximos en ambos modelos. Para verificarlo, se estudió el grado medio  $\langle k \rangle$ , el coeficiente de clusterización  $\bar{C}$ , la distancia me-

dia entre nodos  $\bar{l}$  y la modularidad  $Q$  sobre los clusters máximos provenientes del modelo de diseminación cultural y de la aproximación de percolación.

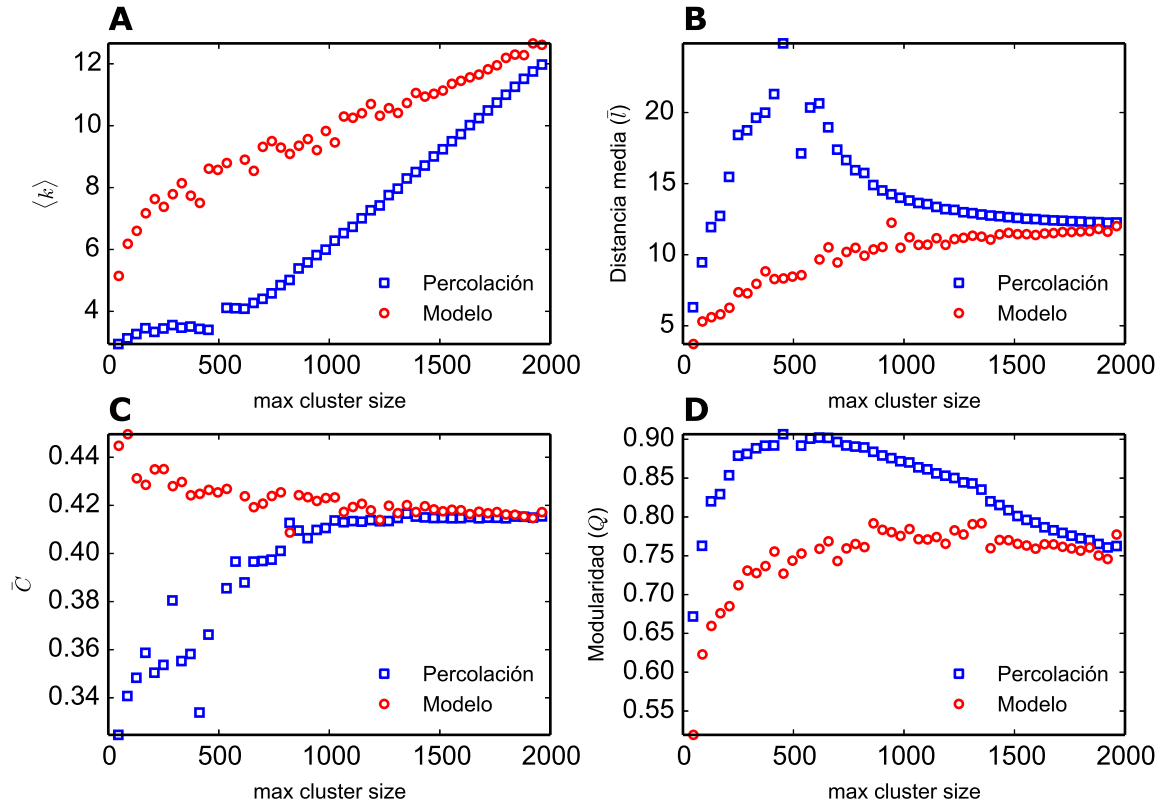


Figura 3.7: Propiedades topológicas de los clusters máximos en función de su tamaño. (A) Grado medio  $\langle k \rangle$ ; (B) distancia media  $\bar{l}$ ; (C) coeficiente de clusterización medio  $\bar{C}$  y (D) modularidad  $Q$ .

En la Figura 3.7 puede observarse que el grado medio  $\langle k \rangle$  es mayor para los clusters máximos del modelo de diseminación cultural. Por otra parte, la distancia media  $\bar{l}$  entre nodos es más pequeña y el  $\bar{C}$  levemente mayor, mientras que  $Q$  es levemente menor al obtenido para el modelo de percolación. Entonces, el mayor  $\langle k \rangle$  y la distancia media  $\bar{l}$  más corta facilitan el encuentro entre infectados y susceptibles para el caso de los clusters obtenidos a partir del mecanismo de diseminación cultural, aumentando así la probabilidad de contagio.

### 3.7. Alta cobertura de la vacuna

Como hemos señalado previamente, los brotes recientes de sarampión en países desarrollados se han producido en un contexto de alta cobertura promedio

de la vacuna (ver Capítulo 1). En la Figura 3.4 se evidencia que cuanto menor es la cantidad de no-vacunados (mayor cobertura de la vacuna), mayores son las discrepancias para el tamaño medio de brote entre nuestro modelo y aquel que asume distribución al azar de los no-vacunados. En esta sección veremos cómo este hecho se manifiesta para un caso particular que satisface el umbral de cobertura de la vacuna  $TV$  del 95 % recomendado por la OMS [WHO 2012].

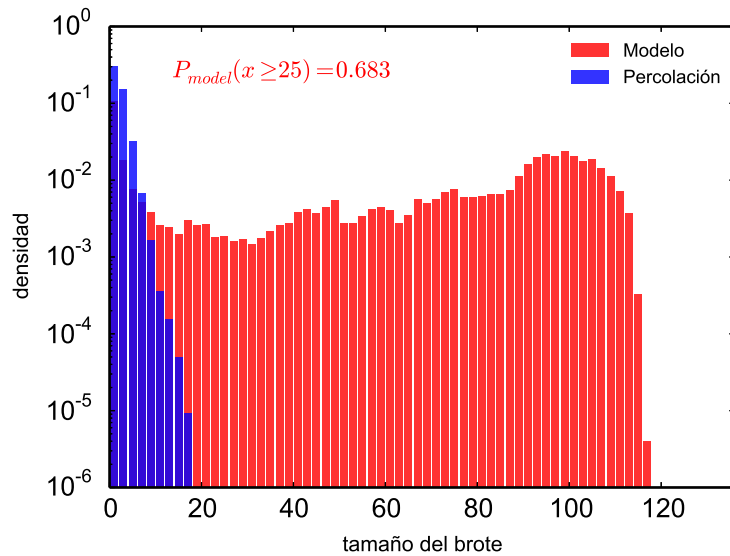


Figura 3.8: Histograma de tamaños de brotes de sarampión bajo condiciones de 95 % de cobertura de la vacuna para nuestro modelo de diseminación cultural (barras rojas) y para la aproximación de percolación (barras azules).

Consideramos 100 realizaciones del modelo de diseminación cultural, con los parámetros previamente establecidos, para las cuales la cobertura final de la vacuna resulte entre 94.6 % y 95.4 %, es decir, entre 115 y 135 no-vacunados en total. Para cada población final de no-vacunados también se estudió la aproximación de percolación. En ambos modelos, y sobre cada una de las configuraciones finales de no-vacunados, se efectuaron 10000 realizaciones del algoritmo de propagación de sarampión. En la Figura 3.8 se presentan los histogramas para el tamaño de los brotes, tanto para el modelo de diseminación cultural como para la aproximación de percolación. Claramente, el modelo de diseminación cultural da lugar a brotes de mayor tamaño que el correspondiente a percolación. Si definimos una epidemia como un brote que abarca al menos el 1 % de la población total (en nuestro caso  $N = 2500$ ), la distribución acumulada a derecha  $P(x \geq 25)$  de tener

un brote involucrando a más de 25 agentes, se define como:

$$P(x \geq 25) = \sum_{n \geq 25} p(n) \quad (3.6)$$

donde  $p(n)$  representa la probabilidad de obtener un brote comprendiendo exactamente  $n$  infectados. En el caso de la Figura 3.8,  $P(x \geq 25) = 0.683$  para el modelo de diseminación cultural, mientras que dicha probabilidad resulta nula para la aproximación de percolación. Adicionalmente, en la Figura 3.9 presenta-

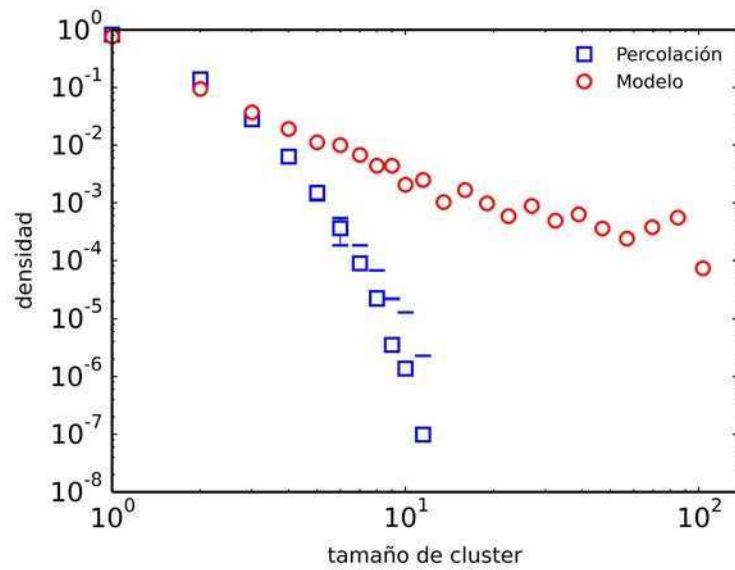


Figura 3.9: Distribución de tamaños de clusters bajo alta cobertura de la vacuna, para el modelo de diseminación cultural (círculos rojos) y para la aproximación de percolación (cuadrados azules). En este último caso se presentan los intervalos de confianza del 95 %.

mos la distribución de tamaños de clusters, restringiéndonos al caso particular de alta cobertura aquí estudiado. Se evidencia notablemente el efecto de la clusterización como un fenómeno que impacta fuertemente en la probabilidad de brote. Más aún, dicho impacto se magnifica en la comparación con el modelo de percolación cuando se considera alta cobertura de la vacuna.

### 3.8. Conclusiones preliminares

En este capítulo presentamos un modelo de propagación de opinión de vacunación sobre una red compleja estática, fundado en el modelo de diseminación

cultural de Axelrod [Axelrod 1997]. Dicho modelo incorpora la homofilia entre agentes sociales como un factor de peso en el mecanismo de imitación social, que conduce a la formación y coexistencia de dominios monoculturales con posturas divergentes.

Observamos que el mecanismo de diseminación cultural sobre la red condiciona la población de no-vacunados y conduce a su clusterización. Dicho efecto actúa como antagonista del fenómeno de inmunidad de grupo y se magnifica en situaciones de amplia cobertura de la vacuna. Esto fue probado explícitamente para el caso del 95 % de cobertura sugerida por la OMS para la vacuna TV contra el sarampión, rubeola y paperas.

# Diseminación de opiniones y vacunación: modelo generalizado

---

## Índice

---

<b>4.1. Contactos personales y de opinión . . . . .</b>	<b>52</b>
<b>4.2. Enlaces adaptativos: mecanismo de reconexión . . . . .</b>	<b>53</b>
<b>4.3. Descripción algorítmica del modelo generalizado . . . . .</b>	<b>55</b>
<b>4.4. Red estática: efectos de la heterogeneidad de enlaces . . . . .</b>	<b>56</b>
<b>4.5. Red adaptativa . . . . .</b>	<b>58</b>
4.5.1. Formación de clusters . . . . .	59
4.5.2. Clusters máximos . . . . .	60
<b>4.6. Acción de las campañas de vacunación . . . . .</b>	<b>62</b>
4.6.1. Modelo generalizado con campo externo . . . . .	63
4.6.2. Barrido en $q$ y $\phi$ . . . . .	64
4.6.3. Acción de $\phi$ sobre los modelos con red estática y adaptativa	67
4.6.4. Incidencia de $\phi$ en la topología de los clusters máximos. . . . .	68
<b>4.7. Modelo generalizado con alta cobertura . . . . .</b>	<b>70</b>
<b>4.8. Conclusiones preliminares . . . . .</b>	<b>71</b>

---

En este capítulo introduciremos mejoras al modelo del capítulo anterior. Una de ellas será la distinción entre enlaces que representan contacto personal entre agentes, de aquellos que solo simbolizan un intercambio de opiniones sin necesidad de contacto. De este modo, consideramos la intermediación de las nuevas tecnologías y su incidencia en la formación de opiniones, aunque sin la posibilidad de transmitir una enfermedad infecciosa.

Finalmente, consideramos la acción de las campañas de información acerca de la vacunación como un factor atenuador del impacto de los grupos anti-vacunación.

## 4.1. Contactos personales y de opinión

Las múltiples alternativas de comunicación sin necesidad de contacto personal constituyen una característica distintiva de nuestra era. Los canales de comunicación no convencionales tales como algunas redes sociales online, mensajería instantánea o microblogging (Twitter), entre otros, constituyen una fuente de información insoslayable que impacta fuertemente en nuestras opiniones [Lerman 2010, Conover 2011, Bond 2012, Centola 2011]. En el Capítulo 3 asumimos una correspondencia exacta entre los enlaces que propagaban opiniones y enfermedades infecciosas transmitidas por contacto directo (en nuestro caso, sarampión). Sin embargo, por lo dicho hasta aquí, resulta evidente que la diseminación de opiniones puede producirse sin que medie un contacto personal que, en cambio, es indispensable para el contagio del sarampión.

Con el objetivo de incorporar una distinción entre las dos modalidades de propagación previamente señaladas, proponemos un modelo original incluyendo dos tipos de enlaces bien diferenciados en nuestra red social:

- a. Enlaces personales: son aquellos que implican proximidad física entre agentes, tales como vínculos familiares, amistades cercanas, relaciones laborales, entre otros. A través de ellos pueden transmitirse tanto opiniones como enfermedades infecciosas.
- b. Enlaces de opinión no-personales: representan relaciones a la distancia, tales como aquellas mediadas por tecnologías de la información. Algunos ejemplos: amistades no-personales en redes sociales online, seguidores de Twitter, suscripciones a blogs, etc.

De este modo, la red sobre la que trabajaremos en este capítulo será estructuralmente idéntica a la introducida en el Capítulo 3, salvo que presentará ambos tipos de enlaces. Los 8 primeros vecinos de un nodo dado  $i$  constituirán su entorno cercano, por lo cual los enlaces hacia ellos serán *enlaces personales*. Por otra parte, 8 de los segundos vecinos de  $i$  constituirán su entorno de opinión, por lo que  $i$  se unirá a ellos a través de *enlaces de opinión*. Por simplicidad, asumiremos que los enlaces no son direccionales, de modo que tanto la influencia social como el contagio pueden darse en ambos sentidos indistintamente. La Figura 4.1 muestra una representación pictórica de la red que usaremos a lo largo del presente capítulo.

Además de heterogeneidad en la naturaleza de los enlaces, el modelo de diseminación de opiniones introducido en este capítulo también involucrará un

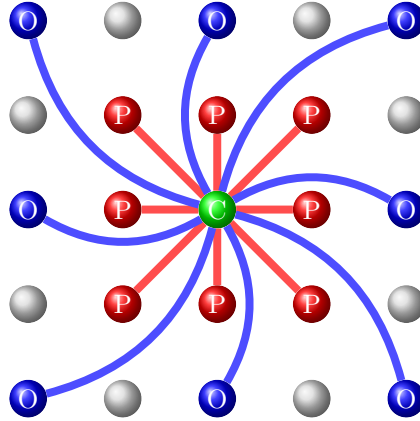


Figura 4.1: Red utilizada en el modelo. Enlaces personales (en rojo) y enlaces no-personales de opinión (en azul).

mecanismo activo de adaptación por reconexionado (rewiring) de enlaces. Es por ello que deberemos distinguir entre dos subclases de redes sociales:

- **Red estática:** sin posibilidad de reconexionado, aunque con heterogeneidad de enlaces (personales y no-personales).
- **Red adaptativa:** aquella que incluye el mecanismo de reconexionado de enlaces.

## 4.2. Enlaces adaptativos: mecanismo de reconexionado

El mecanismo de diseminación cultural del Capítulo 3 considera la adaptación del acervo cultural de cada agente  $i$ , representado por el vector  $\mathbf{V}^i(t)$ , a través de su interacción con un conjunto de vecinos permanentes. Sin embargo, el entorno de los agentes sociales reales no permanece inalterado con el transcurso del tiempo, sino que también sufre modificaciones que acompañan el proceso de adaptación de opiniones. Con el fin de incorporar este aspecto en nuestro modelo, introduciremos un mecanismo de reconexión de enlaces en la versión generaliza-



da de nuestro modelo de diseminación cultural aplicado a la propagación de la conducta de vacunación.

La reconexión de enlaces constituye un paso de adaptación de la red a nivel estructural. Así como los agentes actualizan su opinión al interactuar con su entorno, a través del reconexionado pueden sustituir a los miembros de su entorno por otros con los que presenten mayor grado de similitud u homofilia. Este último constituye un proceso de adaptación activa y por ello a esta clase de redes se las denomina *redes adaptativas*.

Estableceremos una distinción entre los mecanismos de reconexionado de enlaces de opinión y personales. Los primeros son volátiles por definición, generalmente no representan compromiso personal y son absolutamente discrecionales, de modo que pueden actualizarse en su totalidad y sin restricción. En cuanto a los enlaces personales, dado que representan interacciones fuertes entre individuos, deben ser más estables que los primeros, es decir, menos propensos al reconexionado. Es por ello que en este caso solo seleccionaremos una pequeña fracción  $p_{pc}$  de enlaces personales adaptativos.

A fin de conservar las cantidades totales de enlaces personales y de opinión, los reconexionados no pueden darse de cualquier modo. En virtud de ello, el reconexionado diferencial de un enlace  $l_{ij}$ , con nodos fuente  $i$  y destino  $j$ , seguirá las siguientes prescripciones:

- i. Si  $l_{ij}$  es enlace de opinión: se elige cualquier agente  $k$  de la red que no sea vecino de  $i$ . Si el grado de homofilia  $h_t(i, j) < h_t(i, k)$ , entonces el reconexionado se acepta con probabilidad 1 y el enlace  $l_{ij}$  es borrado y sustituido por otro enlace de opinión  $l_{ik}$ . En caso contrario, se rechaza el reconexionado y se implementa el mecanismo de imitación social descrito en el Capítulo 3.
- ii. Si  $l_{ij}$  es enlace personal adaptativo: se elige un agente  $k$  de entre los contactos de opinión del agente  $i$ . Si  $h_t(i, j) < h_t(i, k)$ , entonces el enlace personal adaptativo  $l_{ij}$  se convierte en enlace de opinión, mientras que el enlace  $l_{ik}$  se recategoriza como personal adaptativo.

Cabe destacar que la prescripción de reconexionado para enlaces personales adaptativos tiene la característica favorable de promover un enlace de opinión a enlace personal, contemplando una afinidad previa entre agentes conectados, en lugar de seleccionar cualquier otro agente de la red uniformemente al azar.

### 4.3. Descripción algorítmica del modelo generalizado

La descripción algorítmica del modelo es esencialmente la desarrollada en el Capítulo 3 con el agregado de las prescripciones de reconexión diferencial ya señaladas. Nuevamente, consideramos una red inicialmente cuadrada con  $N = 2500$  nodos, cada uno de ellos con enlaces personales a primeros y segundos vecinos ( $k_p = 8$ ), más el agregado de 8 enlaces no-personales de opinión por nodo ( $k_{op} = 8$ ) tal como se mostró en la Figura 4.1. Adicionalmente, al igual que en el Capítulo 3, fijamos  $F = 10$  (más la opinión binaria asociada con la vacunación) para el modelo de diseminación cultural. Asimismo, consideramos un umbral de interacción por homofilia  $\kappa = 2$ , de modo que dos agentes solo podrán interactuar si presentan coincidencia en al menos dos de sus opiniones (ver ec. 3.5).

En lo que respecta a la subred de enlaces personales, se elige de entre el total de enlaces personales un fracción  $p_{pc} = 0.1$  de enlaces personales adaptativos, que permanecerán en dicha condición durante toda la evolución del algoritmo. Estos son los únicos enlaces personales que pueden ser reconectados como parte del mecanismo de diseminación de opiniones sobre la red adaptativa. Por otra parte, todos los enlaces no-personales, exclusivamente de opinión, son adaptativos.

La condición inicial del sistema es la misma que para el modelo del capítulo anterior: las opiniones son elegidas uniformemente al azar en  $\{1, 2, 3, \dots, q\}$  para cada agente, con excepción de la opinión de vacunación. En este último caso, los agentes comienzan con una opinión pro-vacunación ( $V_{F+1}^i(t) = 0$ ), salvo el 1% de no-vacunadores obstinados pertenecientes a  $S_{stub}$  cuya postura contra la vacunación está fija. Finalmente, la descripción algorítmica del modelo para cada paso de la evolución es:

1. Un agente fuente  $i$  es elegido uniformemente al azar, mientras que uno de sus primeros vecinos  $j$  es elegido como “objetivo”, también uniformemente al azar.
2. Variantes:
  - i. **Red estática:** seguir en el ítem 3.
  - ii. **Red adaptativa:** seguir las prescripciones indicadas en la sección 4.2. Si se concreta el reconexión del enlace  $l_{ij}$ , seguir en el ítem 6. En caso contrario, iniciar el mecanismo de imitación social en el ítem 3.

3. Se obtiene  $h_t(i, j)$  y la probabilidad de interacción  $P(i \rightarrow j, t)$ . Luego, se elige al azar una característica  $n$  en la que  $i$  y  $j$  no coincidan.
4. Si  $n \neq F + 1$ , es decir que la característica no es la postura respecto de la vacunación, el agente  $i$  sigue el proceso de adaptación descrito en la ec. 3.3.
5. Si  $n = F + 1$ , caben dos alternativas:
  - $i \in S_{stub}$ , en este caso  $i$  conserva su postura respecto de la vacunación ( $V_{F+1}^i(t+1) = V_{F+1}^i(t)$ )
  - $i \notin S_{stub}$ , en cuyo caso el agente sigue el proceso de adaptación descrito en la ec. 3.3.
6. Condición de detención: consideramos dos posibles condiciones de detención:
  - i. El algoritmo se detiene cuando alcanza una configuración absorbente, es decir, cuando no son posibles transiciones adicionales.
  - ii. En los casos donde prevalece la opinión de vacunación, el accionar de los agentes  $S_{stub}$  puede dar lugar a pequeñas fluctuaciones intermitentes en la postura respecto de la vacunación en su entorno. Es por ello que fijamos como condición de detención adicional la imposibilidad de realizar transiciones para todas las características culturales, exceptuando pequeñas fluctuaciones en la postura respecto de la vacunación.
7.  $t \rightarrow t + 1$ . Retorna al paso 1.

En cuanto al modelo epidemiológico para la propagación de sarampión, el mismo se describe en el Apéndice A. A continuación, daremos una descripción de ambas variantes del modelo generalizado (i.e., estática y adaptativa), destacando sus similitudes y diferencias, tanto entre ellos como con respecto al modelo del capítulo anterior.

#### 4.4. Red estática: efectos de la heterogeneidad de enlaces

Para el modelo generalizado introducido en el presente capítulo las redes estáticas son aquellas que no admiten reconexión de enlaces, aunque con-

servan las heterogeneidades de enlaces a través de la distinción entre enlaces personales y enlaces no-personales de opinión. En el capítulo precedente hemos considerado solo enlaces homogéneos capaces de transmitir tanto opiniones como enfermedades. Naturalmente, la distinción entre enlaces personales y de opinión no produce ninguna modificación en lo que respecta a la diseminación de la conducta de vacunación en el Capítulo 3, siempre que la red permanezca estática (sin reconexionado), dado que la red sobre la que se propagan las opiniones es esencialmente la misma. Sin embargo, cabe preguntarse cuál será el efecto de la heterogeneidad de enlaces en la propagación de la enfermedad sobre la red estática, dado que ahora tiene lugar exclusivamente sobre una subred de enlaces personales. Con dicho objetivo, realizamos 10000 simulaciones del modelo de propagación de sarampión sobre las redes estáticas obtenidas mediante el modelo generalizado. En la Figura 4.2 presentamos la comparación entre los resultados del modelo del Capítulo 3 y el aquí introducido, ambos correspondiendo a red estática, para el tamaño medio del brote de sarampión en función de la cantidad total de no-vacunados.

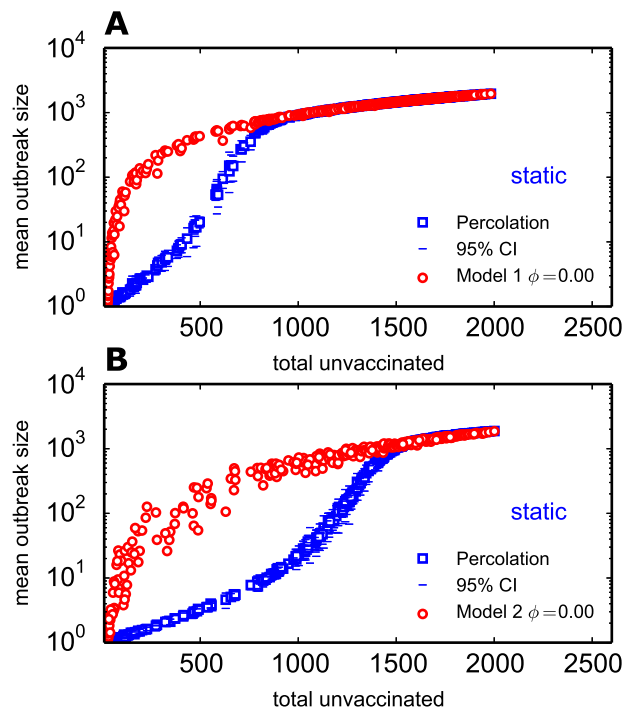


Figura 4.2: Tamaño medio del brote en función de la población de no-vacunados para: (A) red estática con enlaces homogéneos y (B) red estática con enlaces heterogéneos personales y de opinión. En ambos gráficos se presentan los resultados para la aproximación de percolación (cuadrados azules).

Las discrepancias observadas en la la Figura 4.2 se deben a la diferencia en el grado medio de enlaces personales para los nodos: mientras que para la red con enlaces homogéneos todos los enlaces son equivalentes ( $\langle k \rangle = 16$ ), para la red con enlaces heterogéneos  $\langle k_p \rangle = 8$ , siendo  $\langle k_p \rangle$  el grado medio considerando exclusivamente enlaces personales. Resumiendo, la cantidad de enlaces capaces de propagar la enfermedad para la subred constituida por enlaces personales es exactamente la mitad que para el caso de la red con enlaces homogéneos. Dicha diferencia entre los modelos es estructural y ocasiona una modificación en el tamaño del umbral de percolación, resultando mayor para el caso de la red con enlaces heterogéneos.

A modo de conclusión preliminar, diremos que el modelo más realista con enlaces heterogéneos resulta optimista al compararlo con el del capítulo precedente, al menos en lo que respecta a la prevalencia del sarampión. Sin embargo, aún en este escenario optimista, la separación respecto de la aproximación de percolación (discutida en el Capítulo 3) es significativa, tal como se observa en la Figura 4.2. Continuaremos nuestro análisis del modelo generalizado, ahora incorporando la variante de red adaptativa.

## 4.5. Red adaptativa

En esta sección analizaremos los efectos de considerar el reconexionado de enlaces como un mecanismo de adaptación adicional durante la diseminación de opiniones. Naturalmente, la estructura de la red social evolucionará en el tiempo lo cual, como hemos notado a lo largo de estos primeros capítulos, ocasiona a su vez cambios en sus propiedades de propagación. El resultado es una realimentación entre el mecanismo de adaptación de enlaces y la dinámica del modelo de Axelrod. Cabe entonces preguntarnos de qué modo el reconexionado de enlaces afectará la transición entre el consenso vacunador y el no-vacunador en nuestro modelo.

En la Figura 4.3 mostramos la fracción de no-vacunadores en la red en función de  $q$  de forma totalmente análoga a la Figura 3.2 del Capítulo 3. Nuevamente, consideramos  $N = 2500$ ,  $F = 10$ ,  $\langle k \rangle = 16$  y  $\langle k_p \rangle = 8$ .

La transición mostrada en la Figura 4.3 resulta mucho más lenta que la correspondiente a la red estática. Con ello podemos concluir que la región crítica se ensancha, lo que resulta deseable para sintonizar nuestro sistema en la transición. Por otra parte, el valor crítico ahora se encuentra próximo a  $q_C = 300$ , mayor al obtenido para la red estática ( $q_C = 170$ ). Otro aspecto a destacar de la Fi-

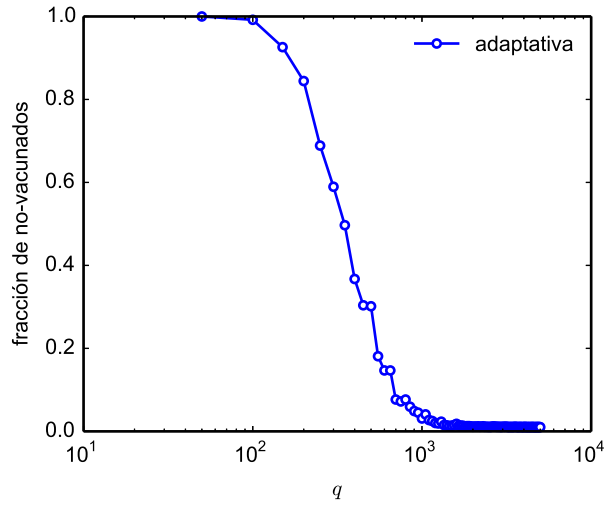


Figura 4.3: Fracción de no-vacunados en función de  $q$  para el modelo generalizado con red adaptativa.

gura 4.3 es que tanto para red estática como adaptativa, el modelo generalizado arroja aproximadamente la misma cantidad final media de no-vacunados para  $q = q_C$ . De aquí en más, fijaremos  $q = 300$  para analizar el comportamiento del modelo generalizado bajo red adaptativa, salvo que se indique explícitamente lo contrario.

#### 4.5.1. Formación de clusters

El aspecto más destacado del modelo del capítulo 3 es la formación de dominios o clusters de agentes no-vacunados con una frecuencia mucho mayor que la obtenida cuando la misma población de no-vacunados es distribuida al azar (aproximación de percolación). Cabe preguntarnos si el modelo con red adaptativa da lugar a un comportamiento semejante.

La Figura 4.4 muestra la distribución de tamaños de clusters para el modelo generalizado con red adaptativa correspondiente a  $q = 300$  y siempre con  $F = 10$ . Es importante destacar que, en el contexto del modelo generalizado, **solo estudiamos clusters para la subred de enlaces personales, que son los únicos relevantes para la transmisión del sarampión**. Tal como en el caso del capítulo precedente, comparamos los resultados con el modelo de percolación a fin de discernir si el comportamiento observado constituye o no un evento fortuito. Puede verse en la Figura 4.4-(A) que efectivamente el modelo de diseminación de opiniones aún conduce a la formación de clusters con mayor fre-

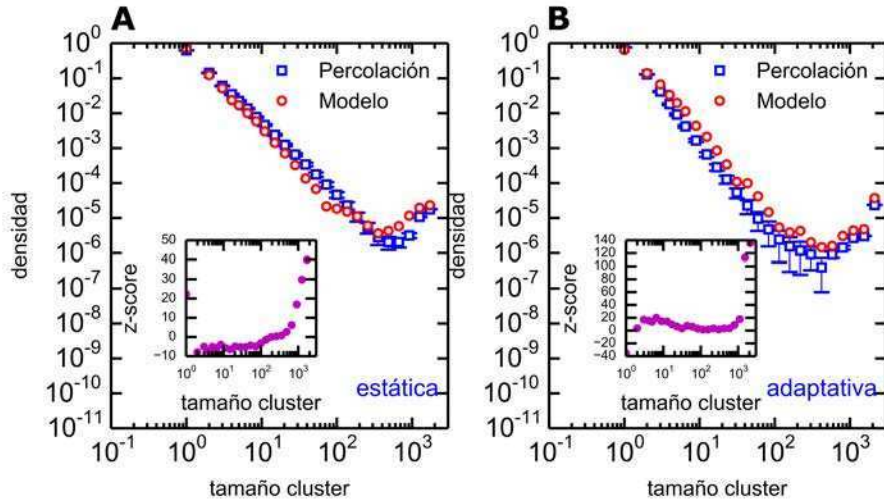


Figura 4.4: Distribución de tamaños de clusters para: (A) modelo generalizado con red estática y (B) modelo generalizado con red adaptativa. En ambos casos se presenta la comparación con la aproximación de percolación y la magnitud del z-score (insets).

cuencia que el modelo de percolación. Más aún, tanto el tamaño absoluto como la frecuencia de los clusters máximos resulta mayor para el caso del modelo con red adaptativa (ver Figura 4.4-(B)). De modo que el modelo de diseminación de opiniones con red adaptativa no solo conduce a la formación de clusters de no-vacunados sino que a su vez resulta más propenso a tal comportamiento. Cabe destacar que cuando comparamos el modelo de red adaptativa con la aproximación de percolación, consideramos la red final luego del proceso de adaptación por reconexión de enlaces. Por este motivo, la red sobre la cual se distribuyen al azar los individuos no-vacunados en el modelo de percolación ya ha sido modificada por el modelo generalizado con red adaptativa (las discrepancias entre ambos serían aún mayores al comparar con el modelo de percolación sobre la red inicial).

### 4.5.2. Clusters máximos

En el Capítulo 3 hemos analizado el tamaño de los brotes sobre el cluster de mayor tamaño  $S_M$  (cluster máximo). Aquí haremos lo propio con los clusters máximos provenientes del modelo con red estática y adaptativa para el modelo generalizado. En la Figura 4.5 se observa el tamaño medio del brote de sarampión en función del tamaño del cluster máximo. La Figura 4.5-(A) presenta la

comparación entre el modelo generalizado para red estática y la aproximación de percolación, mientras que la Figura 4.5-(B) es la equivalente para el modelo generalizado con red adaptativa. Conviene recordar aquí que partimos de un único infectado perteneciente al cluster máximo y efectuamos 10000 realizaciones del algoritmo de propagación de sarampión.

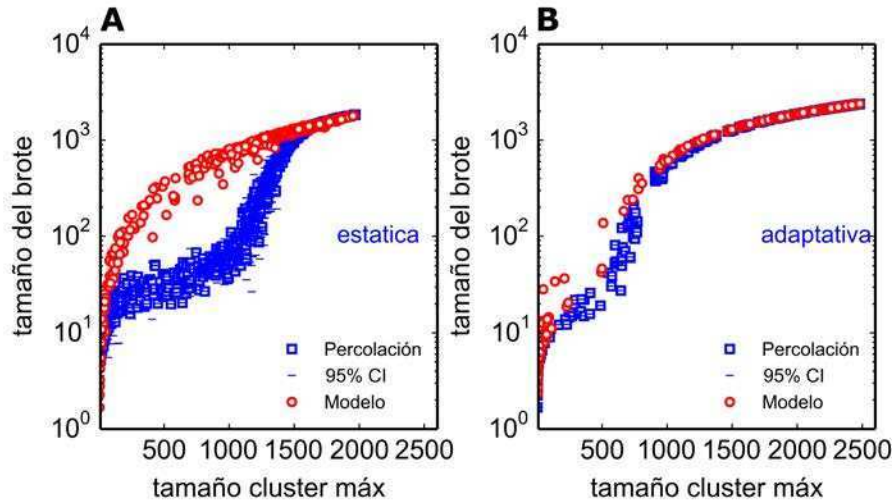


Figura 4.5: Tamaño del brote en función del tamaño del cluster máximo para: (A) modelo generalizado con red estática y (B) modelo generalizado con red adaptativa.

De la comparación entre las Figuras 4.5-(A) y (B) se observa que la red estática presenta diferencias significativas respecto de la aproximación de percolación, efecto que no se presenta de manera tan marcada para el caso de la red adaptativa. Para explicar este último resultado es necesario recurrir al análisis de la estructura de los clusters máximos.

En la Figura 4.6 presentamos las propiedades topológicas de los clusters máximos para el caso del modelo generalizado con red estática (Figura 4.6-(A)) y con red adaptativa (Figura 4.6-(B)), en ambos casos comparamos los indicadores topológicos con los obtenidos para clusters máximos de igual tamaño a partir del modelo de percolación. Puede verse claramente que las discrepancias entre las propiedades de los clusters obtenidos a partir de nuestro modelo y los de la aproximación de percolación, disminuyen para el caso con red adaptativa. Es precisamente por este hecho que las diferencias respecto del modelo de percolación para los tamaños de brotes de la Figura 4.5-(B) también disminuyen apreciablemente respecto de los observados en la Figura 4.5-(A).



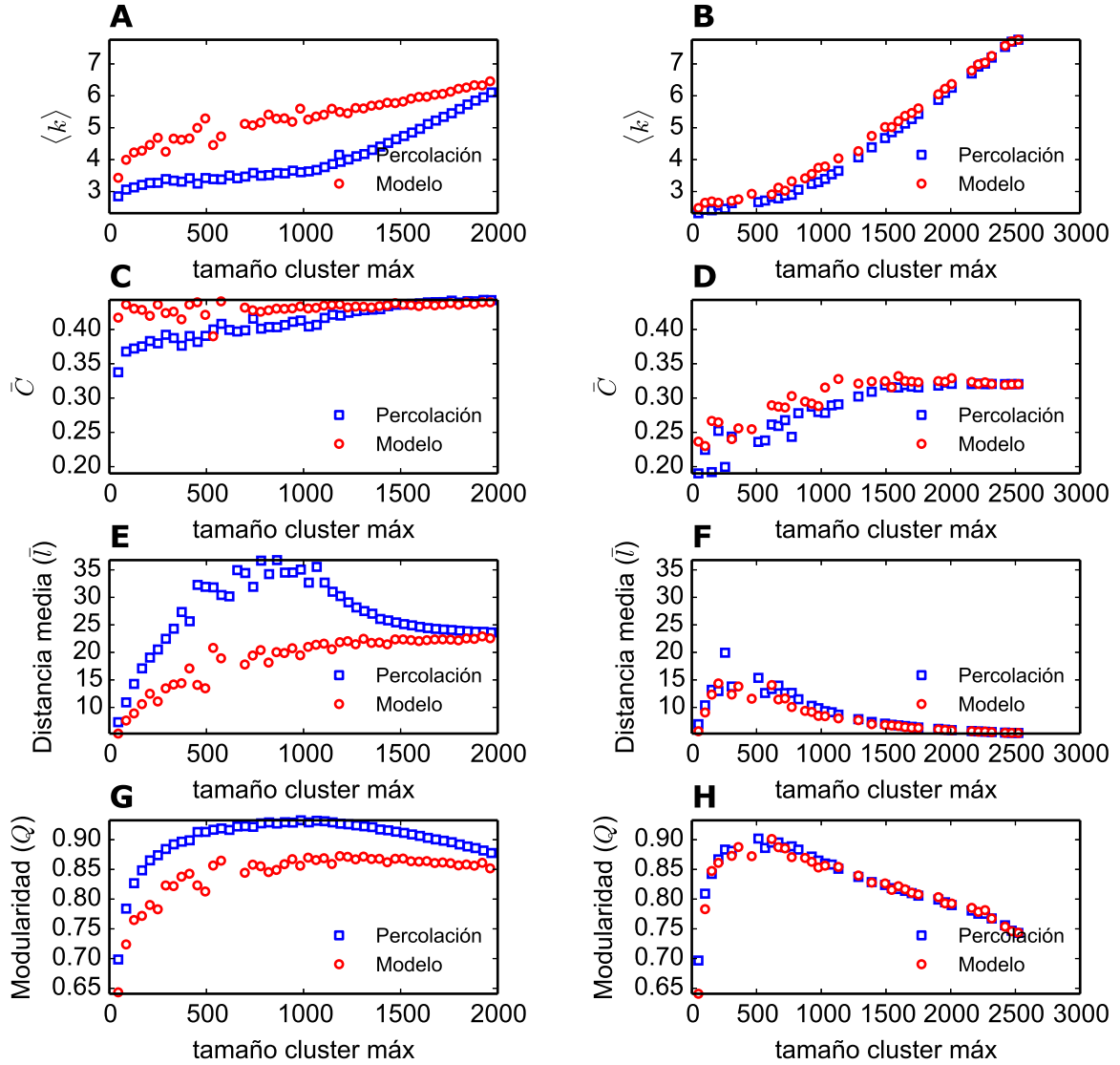


Figura 4.6: Propiedades topológicas de los clusters máximos. Modelo generalizado con (A,C,E,G) red estática y (B,D,F,H) con red adaptativa. En ambos casos se presenta la comparación con las propiedades de los clusters máximos provenientes de la aproximación de percolación.

## 4.6. Acción de las campañas de vacunación

La disponibilidad de vacunas es una condición necesaria para alcanzar una cobertura amplia. Gracias a la implementación de programas de vacunación a nivel global por parte de la OMS, UNICEF y otros organismos de salud pública [PLAN 2012], esta meta puede alcanzarse. El Plan Estratégico de la OMS 2012-

2020 [PLAN 2012] contempla entre sus objetivos un aspecto fundamental:

- “Comunicar y comprometerse a construir confianza y demanda pública para la inmunización.”

Este es precisamente el aspecto en el que nos centramos en la presente sección. Claramente, la disponibilidad de vacunas no es suficiente si las personas las rechazan. El rechazo proviene mayoritariamente de factores culturales [Salmon 2006, Saint-Victor 2013], de conciencia [Pearce 2008] y a una percepción errónea del riesgo asociado con la vacuna **TV**. Por supuesto, lo que nosotros sostenemos aquí es que esta percepción del riesgo no está fundada en una opinión informada por parte de los agentes decisores, sino en un mecanismo de diseminación de opiniones por imitación social.

#### 4.6.1. Modelo generalizado con campo externo

Si bien los agentes adoptan una postura de vacunación bajo la influencia de su entorno próximo, también son permeables a campañas de promoción de los beneficios de la vacunación, llevadas a cabo por las autoridades sanitarias. Incorporamos este factor adicional a nuestro modelo de agentes en la forma de un campo externo global  $\phi \geq 0$ , al cual están expuestos todos los agentes. Dicho campo  $\phi$  tiene por función introducir un sesgo en la probabilidad de transición de vacunador hacia no-vacunador en el mecanismo de imitación social.

La implementación de la acción de  $\phi$  en nuestro modelo generalizado es la siguiente: Sea  $i$  un “agente fuente” vacunador ( $V_{F+1}^i(t) = 0$ ) que interactúa con un “agente destino” no-vacunador  $j$  ( $V_{F+1}^j(t) = 1$ ) a través del mecanismo de imitación. Si la característica elegida para el mecanismo de adaptación es la postura de vacunación, dicha postura se actualiza según

$$V_{F+1}^i(t+1) = \begin{cases} V_{F+1}^j(t) & \text{con probabilidad } (1 - \phi) \\ V_{F+1}^i(t) & \text{con probabilidad } \phi \end{cases} \quad (4.1)$$

Esto es: el agente vacunador (fuente) aún conserva una chance de persistir en su postura, con probabilidad  $\phi$ , al interactuar mediante el mecanismo de imitación social con un agente no-vacunador (target o destino). En cuanto a la implementación sobre el modelo generalizado con red adaptativa, cabe enfatizar que el campo externo  $\phi$  no interviene en forma directa con el mecanismo de reconexión de enlaces.

Como resulta evidente de la ecuación 4.1, cuando  $\phi = 0$  se recupera el mecanismo de imitación social tradicional. Por otra parte, es importante destacar que la campaña aquí modelada es solo de información favorable y concientización hacia la vacunación en el sentido propuesto por la OMS. La acción de  $\phi$  no pretende imponer la vacunación compulsiva ni tiene incidencia en la posición de aquellos que han tomado la decisión de no-vacunar, sino que solo actúa sobre aquellos agentes vacunadores para que no modifiquen su postura en el futuro.

#### 4.6.2. Barrido en $q$ y $\phi$

El hecho de incorporar un campo externo  $\phi$  naturalmente amplía nuestro espacio de parámetros. Es por ello necesario implementar un barrido tanto sobre  $q$  como para nuestro nuevo parámetro  $\phi$ . En la Figura 4.7 presentamos la población media de agentes no-vacunados para nuestro modelo generalizado, tanto para red estática (Figura 4.7-(A)) como adaptativa (Figura 4.7-(B)), en función de  $q$  y  $\phi$ , conservando fijo  $F = 10$ . En ambos casos se destaca el efecto mitigador de  $\phi$  por sobre la acción del 1% de agentes obstinados no-vacunadores. Una mínima intervención masiva sostenida en el tiempo presenta efectos notables en la evolución de la postura de vacunación.

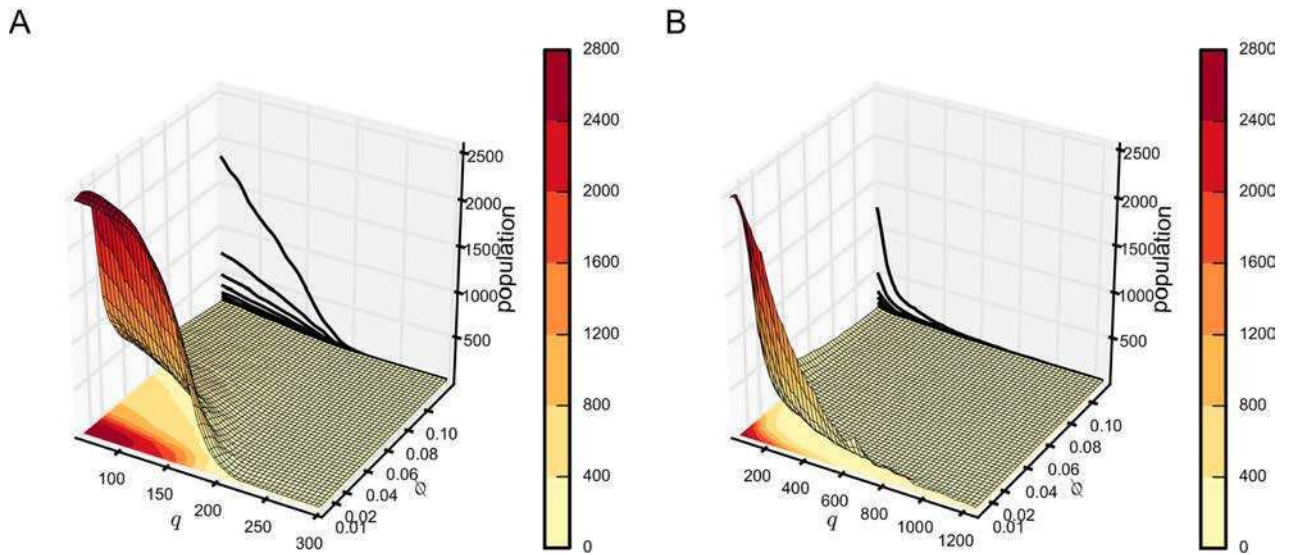


Figura 4.7: Población de no-vacunados para barrido en  $q$  y  $\phi$ . Modelo generalizado correspondiente a (A) red estática y (B) red adaptativa.

La presencia de un campo externo  $\phi > 0$  frustra la posibilidad de consenso no-vacunador absoluto observado en las transiciones previas con  $q < q_C$  (ver Figura

4.9). Para valores pequeños de  $q$  pueden aún formarse dominios de consenso si se analizan solo las primeras  $F$  opiniones, sin embargo, esto no ocurre para el caso de la opinión  $F + 1$  referida a la vacunación. Lo que ocurre para  $q < q_C$  es una competencia entre dos posturas antagónicas sobre un dominio de convergencia en las primeras  $F$  opiniones.

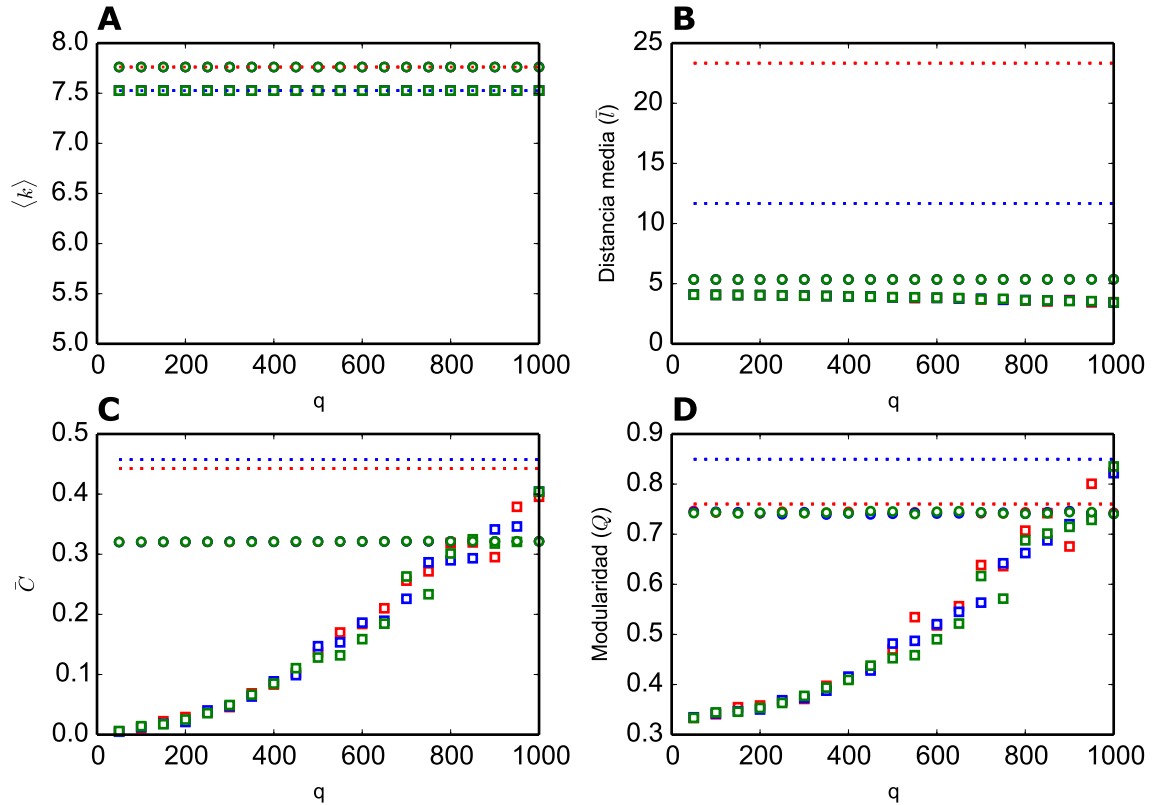


Figura 4.8: Topología de las subredes de enlaces personales y de opinión con campo externo  $\phi$ . Los círculos (cuadrados) rojos, azules y verdes corresponden a las subredes de enlaces personales (de opinión) y  $\phi = 0.00, 0.01$  y  $0.02$ , respectivamente. Las líneas de puntos indican los valores correspondientes para las subredes iniciales  $G_{pers}$  (línea de puntos rojos) y  $G_{op}$  (línea de puntos azules).

Cabe también preguntarnos cómo se modifica la topología de la red al barrer  $q$  para  $\phi = 0.00, 0.01$  y  $0.02$ , debido al proceso de reconexión de enlaces en el caso del modelo generalizado con red adaptativa. Recordemos que la red se encuentra conformada por una subred de contactos personales ( $G_{pers}$ ) y otra integrada por contactos exclusivamente de opinión ( $G_{op}$ ). La Figura 4.8 muestra el grado medio ( $\langle k \rangle$ ), la longitud de camino media ( $\ell$ ), el coeficiente de clusterización medio ( $\bar{C}$ ) y la modularidad ( $Q$ ), todos en función de  $q$  y promediados sobre

100 realizaciones del modelo, tanto para la subred de contactos personales como para la correspondiente a contactos exclusivamente de opinión. El valor de  $\ell$  se obtiene sobre el cluster de mayor tamaño en aquellos casos donde la subred se fragmenta (evento muy improbable para el caso de la subred de contactos personales  $G_{pers}$ ). Como puede observarse en la Figura 4.8,  $q$  no afecta la topología de  $G_{pers}$  pero sí lo hace con la de  $G_{op}$ . Por la formulación del modelo,  $\langle k \rangle$  permanece constante para ambas subredes (es un poco menor para  $G_{op}$  debido a las diferencias con  $G_{pers}$  cerca de los bordes). El valor de  $\ell$  permanece casi inalterable ante el incremento de  $q$ , muy por debajo de su valor inicial, para ambas subredes. Esto implica que la cantidad de reconexionados aceptados es en todos los casos suficiente para disminuir el  $\ell$  con respecto al inicial. El  $\bar{C}$  crece con  $q$  para el caso de  $G_{op}$ . Esto muestra que para  $q$  pequeños la probabilidad de aceptación del reconexionado se incrementa, dando lugar a una topología compatible con una subred aleatoria. Sin embargo, al incrementar  $q$ ,  $\bar{C}$  comienza a converger al valor correspondiente para la red inicial (indicados en líneas punteadas en la Figura 4.8) debido a la reducción drástica en la probabilidad de aceptación, tanto de pasos de adaptación pasiva como activa, esta última por reconexionado de enlaces. El gráfico para la modularidad  $Q$  en función de  $q$  también sustenta esta hipótesis. Cabe destacar que los indicadores topológicos de  $G_{pers}$  le adjudican propiedades de mundo pequeño. Por otra parte, en todos los casos se concluye la independencia de la topología final de la red respecto del campo externo  $\phi$ , cuyos resultados se presentan para  $\phi = 0.00$  (símbolos rojos),  $\phi = 0.01$  (símbolos azules) y  $\phi = 0.02$  (símbolos verdes).

Estudiar la dinámica compleja de este modelo para todo el espacio de parámetros es un proceso muy extenso que se encuentra en sus comienzos. De todos modos, nuestro objetivo es analizar la formación de clusters de no-vacunadores debido a un proceso de diseminación cultural, aún en situación de alta cobertura de la vacunación, y en la región crítica donde dichos dominios alcanzan un equilibrio metaestable.

La acción de las campañas de información masiva en materia de salud pública no ha sido un aspecto muy explorado en los modelos de campo medio, con excepción de [d'Onofrio 2012]. Según nuestro leal saber y entender, esta constituye la primera oportunidad en la que se trata el tema sobre un modelo con estructura heterogénea de red social [Medus 2014b].

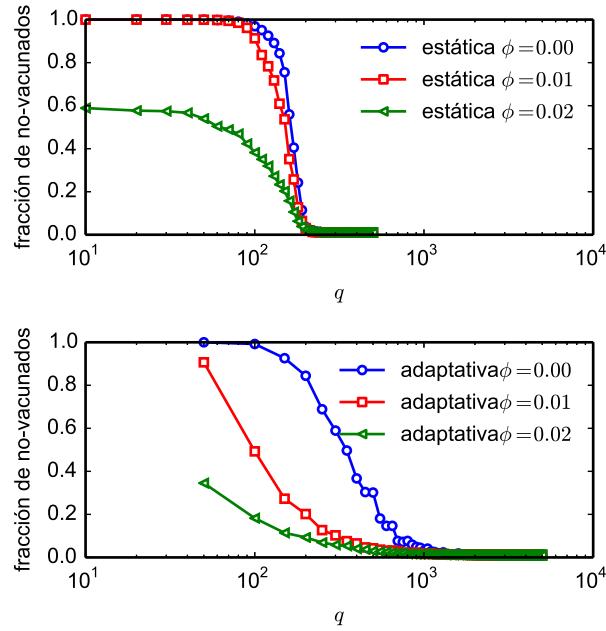


Figura 4.9: Dependencia de la fracción de no-vacunados con  $q$  para ambas variantes del modelo generalizado de diseminación de opiniones.

### 4.6.3. Acción de $\phi$ sobre los modelos con red estática y adaptativa

Cabe preguntarnos qué diferencias cuantitativas concretas existen entre la acción mitigadora del campo externo  $\phi$ , sobre los modelos de red estática y adaptativa. Para ello, efectuamos 500 realizaciones tanto del modelo de diseminación de opiniones con red estática como adaptativa, con  $\phi \in \{0.0, 0.1, 0.2\}$ . A su vez, sobre cada una de las 500 configuraciones finales de no-vacunados efectuamos 10000 realizaciones del modelo de propagación de sarampión, registrando en cada caso el tamaño final del brote. Para cuantificar la acción mitigadora del campo externo, calculamos la fracción de realizaciones donde la cantidad de infectados totales supera el umbral del 1% de la población total (dado  $N = 2500$ , dicho umbral será 25 agentes). Dicha distribución acumulada a derecha fue definida en la ec. 3.6 del Capítulo 3 y notada como  $P(x \geq 25)$ , es decir, la probabilidad de tener un brote que involucre a 25 o más agentes.

En la Figura 4.10 mostramos  $P(x \geq 25)$  para  $\phi \in \{0.0, 0.1, 0.2\}$ , tanto para el modelo con red estática como adaptativa. El intervalo de confianza binomial para cada punto resulta despreciable dado el gran tamaño de la muestra. Una conclusión que se extrae inmediatamente del gráfico de la Figura 4.10 es que el

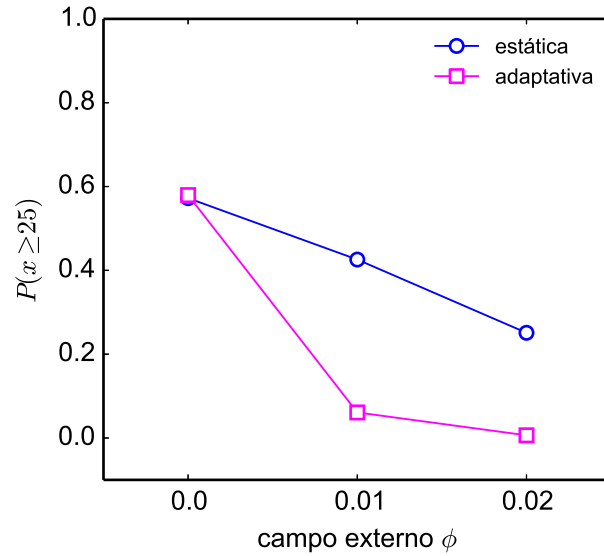


Figura 4.10: Comparación de la incidencia del campo externo en los modelos con red estática y adaptativa a partir de  $P(x \geq 25)$ .

efecto mitigador de  $\phi$  resulta mayor en el caso correspondiente a red adaptativa. Para hallar los orígenes de este fenómeno será necesario abordar el estudio del efecto de  $\phi$  sobre la estructura típica de los clusters de no-vacunados arrojados por ambas variantes del modelo generalizado.

#### 4.6.4. Incidencia de $\phi$ en la topología de los clusters máximos.

Hemos visto que la cantidad final de no-vacunados a  $\phi = 0$  es equivalente tanto para el modelo generalizado con red estática como adaptativa en  $q = q_C$ . Sin embargo, acabamos de ver que al considerar  $\phi > 0$ , la probabilidad de brote de sarampión disminuye más abruptamente para el caso del modelo con red adaptativa. Esto no se debe a que la red adaptativa cambie su topología al aumentar  $\phi$ , dado que estos cambios no son significativos, tal como mostramos en la Figura 4.8. El motivo de las diferencias observadas se obtendrá al comparar la topología de los clusters máximos (correspondientes a la subred de enlaces personales  $G_{pers}$ ) para el modelo generalizado con red estática y adaptativa. La Figura 4.11 compara  $\langle k \rangle$ ,  $\ell$ ,  $\bar{C}$  y  $Q$  para  $\phi = 0.01$  (el caso correspondiente a  $\phi = 0.02$  es totalmente análogo y la Figura 4.6 muestra el de  $\phi = 0$ ). Tal como se observa en el gráfico para  $\phi = 0.01$ , el modelo con red adaptativa arroja típicamente

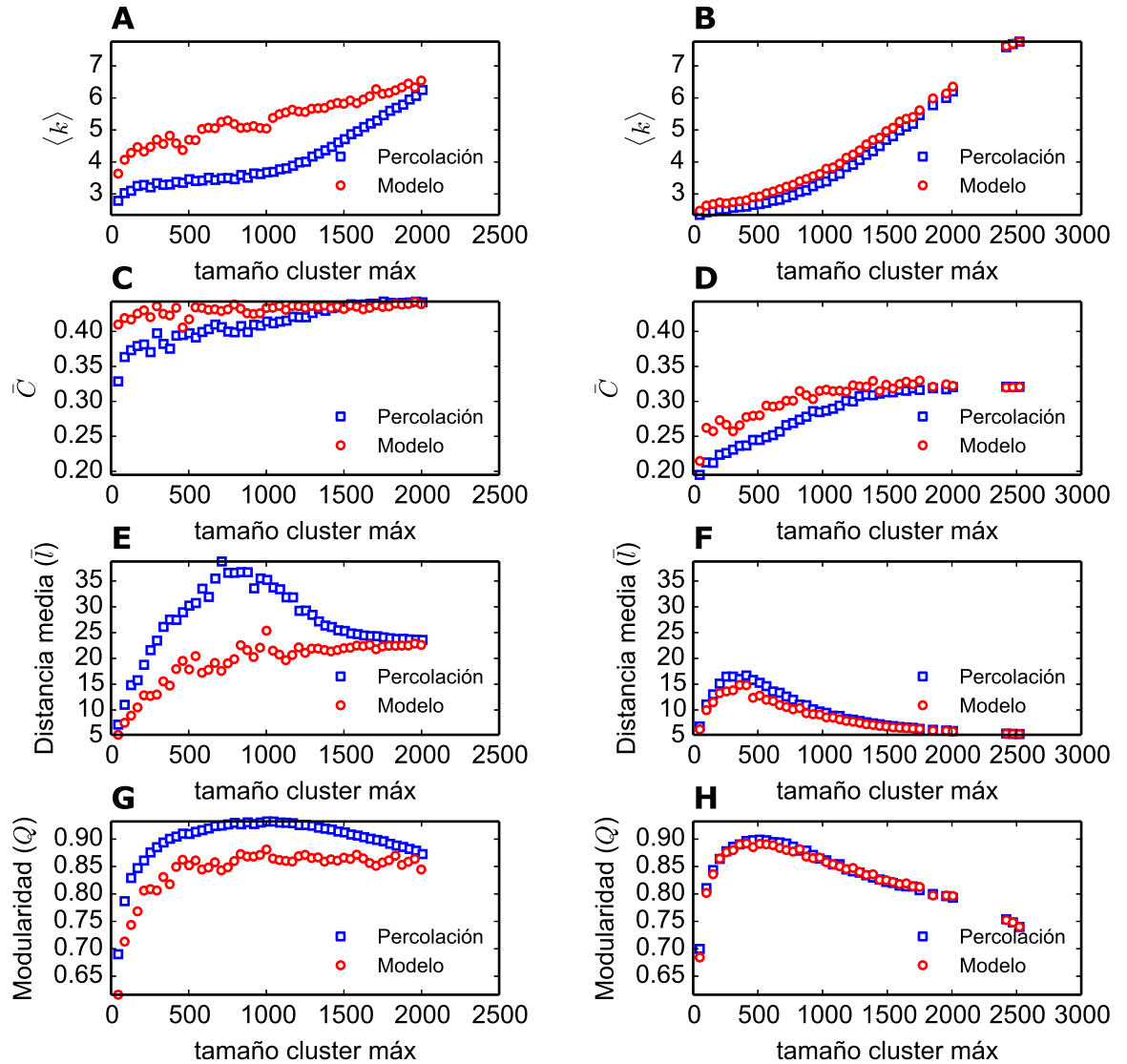


Figura 4.11: Propiedades topológicas de los clusters máximos con campo externo  $\phi = 0.01$ . Modelo generalizado con (A,C,E,G) red estática y (B,D,F,H) con red adaptativa. En ambos casos se presenta la comparación con las propiedades de los clusters máximos provenientes de la aproximación de percolación.

clusters con grado medio menor que el de red estática (comparando clusters del mismo tamaño). Por otra parte, la distancia media entre nodos ( $\bar{\ell}$ ) resulta menor para el caso correspondiente a red adaptativa, el coeficiente de clusterización ( $\bar{C}$ ) es levemente menor y la modularidad ( $Q$ ) alcanza valores altos en ambos casos. De modo que el mecanismo de reconexionado incorporado en el modelo con red adaptativa da lugar a clusters compatibles con la característica de árbol, tal co-



mo puede verse en el ejemplo de la Figura 4.12. Las diferencias estructurales aquí

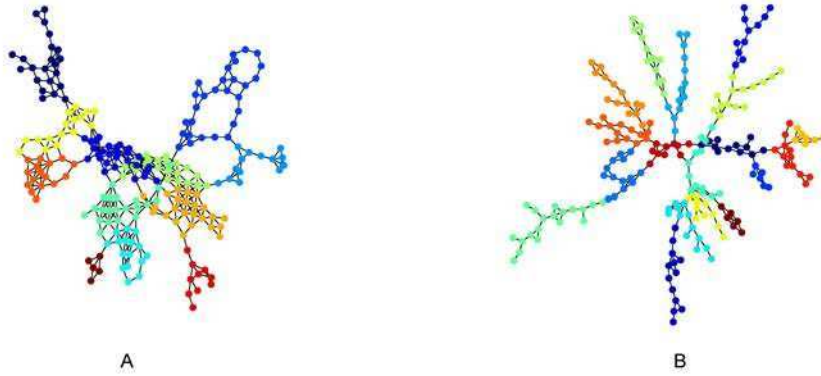


Figura 4.12: Ejemplos de clusters máximos integrados por aproximadamente 200 nodos para: (A) modelo con red estática y (B) con red adaptativa, ambos para  $\phi = 0.01$ . Los distintos colores distinguen las comunas obtenidas a partir de la mejor partición de acuerdo con el método de Louvain [Blondel 2008] incluido en el paquete Networkx de Python para la optimización de la modularidad [Hagberg 2008].

analizadas hacen que los clusters de no-vacunadores provenientes del modelo con red adaptativa, resulten más frágiles que los del modelo con red estática ante la acción del campo externo  $\phi$ . Resumiendo, aunque el modelo generalizado con red adaptativa promueva la clusterización aún más allá que para el caso con red estática, los clusters de no-vacunados alcanzados mediante el primero resultan más frágiles que aquellos alcanzados a través el segundo.

## 4.7. Modelo generalizado con alta cobertura

En el capítulo anterior analizamos la incidencia del modelo de diseminación cultural en aquellos casos con una cobertura  $P_{cob}$  próxima a la meta de 95% ( $94.6\% \leq P_{cob} \leq 95.4\%$ ) recomendada por la OMS. La Figura 4.13 muestra que el modelo generalizado presenta una probabilidad de brote notablemente superior a la que se obtiene a partir de la aproximación de percolación, tanto para red estática como dinámica. Asimismo, puede observarse el efecto mitigador del campo externo  $\phi$  sobre la probabilidad de brote de tamaño mayor al 1% de la población total ( $P(x \geq 25)$ ).

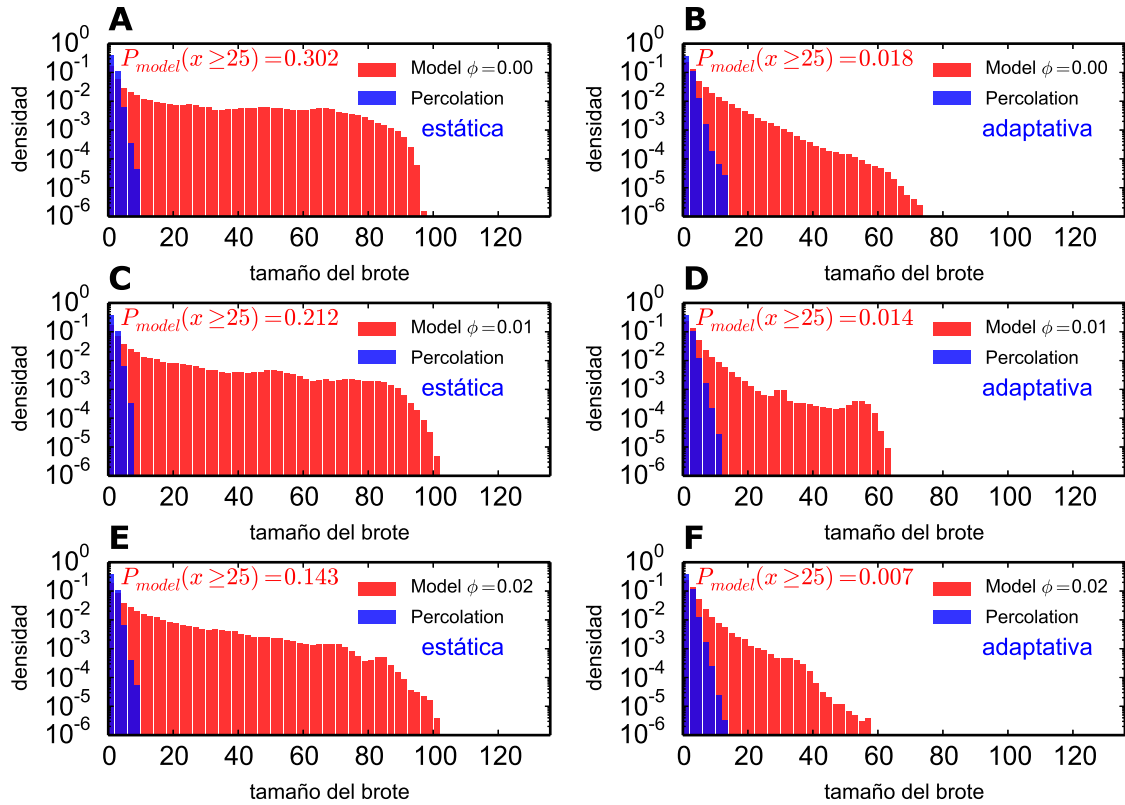


Figura 4.13: Histograma de tamaños de brotes de sarampión bajo condiciones de 95 % de cobertura de la vacuna para el modelo generalizado de diseminación cultural (barras rojas) y para la aproximación de percolación (barras azules). Se presentan los resultados para  $\phi = 0.00, 0.01, 0.02$ , tanto para el caso del modelo con red estática como adaptativa.

## 4.8. Conclusiones preliminares

En el presente capítulo hemos mostrado que la distinción realista entre enlaces personales y no-personales impacta en la incidencia acumulada del sarampión al modificar estructuralmente el subgrafo sobre el cual se propaga. No obstante, aún se manifiesta el fenómeno de clusterización de no-vacunadores, vinculado al mecanismo de diseminación cultural, observado en el capítulo anterior.

Adicionalmente, hemos incluido la variante con adaptación estructural de la red por medio del reconexión de enlaces como una variante del modelo generalizado propuesto. Mostramos que dicha variante con red adaptativa, por un lado favorece la formación de clusters de gran tamaño, pero sus propiedades topológicas los tornan estructuralmente más frágiles que los correspondientes a

la variante estática del modelo. Por este hecho, encontramos que la acción de una campaña de información en favor de la vacunación, introducida a través del campo externo  $\phi$ , produce una mayor disminución en la probabilidad de brote sobre el modelo generalizado con red adaptativa.

# Generación de redes sociales con base empírica

---

## Índice

---

<b>5.1. Actividad social como generador de heterogeneidades estructurales . . . . .</b>	<b>74</b>
<b>5.2. Modelo estocástico generalizado . . . . .</b>	<b>76</b>
5.2.1. Efectos de memoria no-markovianos: cierre triádico . . . . .	77
5.2.2. Formulación analítica del modelo . . . . .	78
5.2.3. Población constante ( $\gamma = 0$ ) . . . . .	81
5.2.4. Población creciente . . . . .	84
<b>5.3. Correlaciones de grado de alto orden . . . . .</b>	<b>86</b>
5.3.1. Grado medio de primeros vecinos . . . . .	86
<b>5.4. Coeficiente de clusterización . . . . .</b>	<b>90</b>
<b>5.5. Casos de estudio . . . . .</b>	<b>91</b>
5.5.1. Propiedades topológicas . . . . .	93
5.5.2. Mecanismo <b>TC</b> en redes reales . . . . .	95

---

Tal como hemos señalado en los capítulos precedentes, la estructura de las redes sociales tiene un impacto directo en sus propiedades de propagación. Es por ello que las propiedades topológicas estáticas de las redes complejas fueron profusamente analizadas en la literatura científica durante los primeros años de este siglo.

Lejos de ser objetos estáticos, las redes complejas, en particular las de origen social, presentan una evolución constante que abarca la creación y, eventualmente, destrucción continua de enlaces y nodos. La imagen de nodos unidos por enlaces corresponde entonces tan solo a una representación acumulada en el tiempo de un proceso dinámico. Dicha evolución de la red adquiere relevancia cuando

su tiempo característico es comparable con el correspondiente al fenómeno (ya sea este una enfermedad infecciosa o una opinión) que se propaga sobre la misma. Por otra parte, recientemente la comunidad científica pudo acceder a la evolución microscópica de grandes redes online con buena resolución temporal, y comenzó a realizar experimentos concretos tendientes a dilucidar la dinámica de contactos entre individuos [Isella 2011]. La conjunción de todos estos elementos impulsó el estudio de las llamadas *redes variables en el tiempo* como una subárea del campo de las redes complejas.

Distintos modelos de redes variables en el tiempo acompañan la aparición de nuevos sets de datos empíricos. Sin embargo, dichos modelos asumen habitualmente hipótesis tendientes a simplificar su estudio analítico. Entre dichas hipótesis se encuentra la de modelar a los hombres como caminantes al azar [Starnini 2013a] o, en el mejor de los casos, como sujetos sin memoria de sus contactos previos [Perra 2012]. Como consecuencia de ello, las redes acumuladas en el tiempo arrojadas por dichos modelos no presentan la alta clusterización, ni el comportamiento asortativo, en suma, las correlaciones de grado características de las redes sociales estudiadas en su versión estática. Esto plantea una tensión entre dos representaciones de la misma realidad: la que reproduce la dinámica de los contactos y la que hace foco en su representación acumulada en el tiempo. El objetivo del presente capítulo es proponer un marco teórico que permita resolver este conflicto, aportando elementos para que las redes generadas algorítmicamente satisfagan las condiciones ampliamente observadas en las redes sociales acumuladas en el tiempo. El algoritmo aquí introducido nos permitirá entonces generar redes sociales acumuladas en el tiempo con base empírica.

## 5.1. Actividad social como generador de heterogeneidades estructurales

Muchas de las redes sociales analizadas en la literatura poseen ciertas características generales distintivas: pequeña distancia media entre nodos ( $\propto \ln N$ ) y alta clusterización al compararla con una red aleatoria con el mismo número de nodos  $N$  y enlaces  $L$ , fuerte estructura de comunas y distribuciones de grados  $P(k)$  que pueden abarcar desde distribución exponencial a ley de potencias (redes libres de escala). En cuanto a esta última propiedad, hemos mencionado en la Introducción (Capítulo 1) que el mecanismo más extendido para el crecimiento de redes libres de escala es el *enlace preferencial* [Barabasi 1999b]. Dicho

mecanismo está basado solo en la conectividad y requiere del crecimiento de la población por agregado de nodos y enlaces para justificar la formación de redes libres de escala. No obstante, aún se discuten sus orígenes microscópicos.

En el contexto de las redes sociales variables en el tiempo, se ha propuesto que la distribución de grados libre de escala puede tener origen en una característica intrínseca de los agentes sociales denominada *actividad social* ( $a$ ) [Perra 2012]. Algunos individuos resultan socialmente más activos que otros, por lo cual establecen mayor cantidad de vínculos sociales que sus pares. Luego, si la distribución de actividad social sigue una ley de potencias  $F(a) \propto a^{-\gamma}$ , puede mostrarse que dicha distribución es heredada por la distribución de grados de la red de contactos acumulados durante una ventana temporal  $T$ , es decir,  $P_T(k) \propto k^{-\gamma}$ . Recientes estudios han probado que la actividad social de individuos que participan en proyectos colaborativos tales como Wikipedia (actividad medida en términos de número de aportes) se correlaciona fuertemente con el grado que los mismos individuos adquieren en la red de seguidores [Muchnik 2013].

El modelo propuesto en [Perra 2012] considera un conjunto de  $N$  nodos inicialmente desconectados. A cada nodo  $i$  se le asigna una actividad social  $a_i$  de acuerdo con una distribución de probabilidad  $F(a)$ . El valor de  $a_i$  corresponde entonces a la tasa de activación del nodo  $i$ , por lo que  $a_i \delta t$  es la probabilidad que tiene el nodo  $i$  de activarse en el intervalo de tiempo  $\delta t$ . Cuando un nodo se activa, establece enlaces con otros  $m$  nodos elegidos al azar, estén estos últimos activos o no. Luego del paso temporal  $\delta t$ , todos los enlaces de la red son borrados y el proceso comienza nuevamente. De este modo, los nodos actúan como agentes markovianos sin memoria de sus contactos previos. Motivo por el cual la red acumulada temporalmente resulta una red aleatoria con distribución de grados

$$P_T(k) \propto F \left[ \frac{k}{mT} \right], \quad (5.1)$$

siendo  $T$  el tiempo acumulado. En la Figura 5.1 se muestran ejemplos consignados en [Perra 2012] para distribuciones de actividad  $F(x)$  empíricas obtenidas a partir de 3 bases de datos: la de publicaciones en Physical Review Letters (PRL), la base de películas IMDB y, por último, la de mensajes en la red Twitter. Puede verse que en los tres casos el comportamiento de  $F(x)$  presenta un comportamiento compatible con leyes de potencias.

El éxito del modelo de Nicola Perra et al radica entonces en la explicación microscópica del origen de la distribución libre de escala, dentro del paradigma de redes variables en el tiempo. Sin embargo, las redes acumuladas obtenidas a

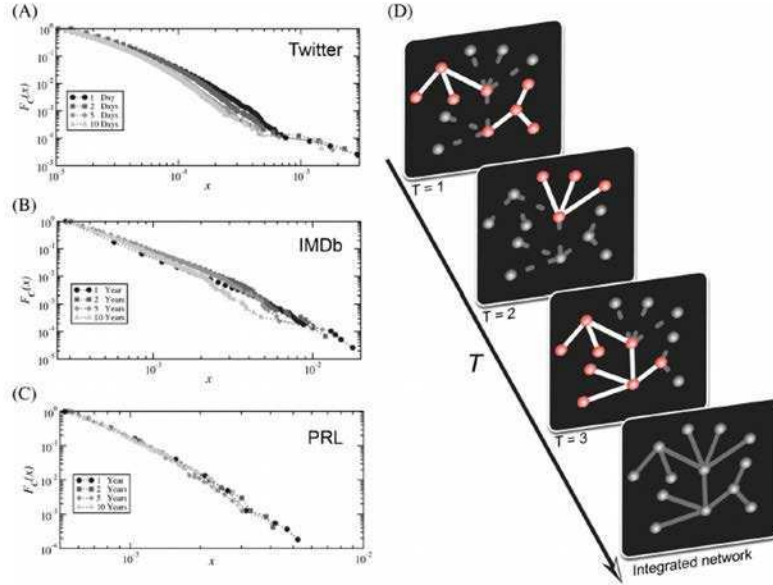


Figura 5.1: Distribuciones acumuladas de actividad de usuarios,  $F(x)$ , medidas empíricamente usando cuatro ventanas temporales diferentes en cada caso. El panel (A) muestra la distribución de actividad acumulada para la red de Twitter, (B) para la actividad de los actores de películas según la base IMDb, y el panel (C) para los autores de publicaciones en la revista Physical Review Letters (PRL). El panel (D) muestra una representación esquemática del modelo de actividad social considerando solo 13 nodos. Se muestra la red acumulada en el tiempo para 3 pasos temporales sucesivos.

partir de este modelo carecen de otras características fundamentales mencionadas al comienzo de esta sección. Es por ello que en la siguiente sección nos abocaremos a desarrollar un modelo alternativo no-Markoviano que conserve las ventajas del modelo anterior, aunque incorporando las características adicionales propias de las redes sociales reales.

## 5.2. Modelo estocástico generalizado

Sea  $G_t(N, L)$  el grafo dependiente del tiempo  $t$ , compuesto por  $N(t)$  nodos y  $L(t)$  enlaces. Como en el caso de las redes inducidas por actividad social (**AS**), también asumiremos que cada nodo  $i$  posee una actividad intrínseca  $a_i$  dada por una pdf  $F(a)$ . En el modelo **AS** los nodos se activan en cada paso temporal  $\delta t$  con una probabilidad dada por  $a_i \delta t$ . Aquí proponemos un modelo de crecimiento estocástico generalizado (**GSG**) para la red, donde  $a_i$  representa la tasa a la

cual emergen nuevos enlaces desde el nodo  $i$ . De este modo, en lugar de agregar múltiples enlaces en cada paso temporal, estos irán generándose uno por uno siguiendo un proceso de Poisson con tasa variable. Entonces, los enlaces serán incorporados con una tasa dependiente del tiempo dada por

$$\beta(t) = \sum_{i=1}^{N(t)} a_i . \quad (5.2)$$

Así como los enlaces son agregados, el modelo **GSG** también admite un proceso de eliminación de enlaces, de modo que a cada enlace  $l_j$  se le asigna una tasa de borrado  $\mu_j$  proveniente de una pdf  $F_D(\mu)$ . En consecuencia, la tasa de borrado total en función del tiempo será

$$\eta(t) = \sum_{j=1}^{L(t)} \mu_j . \quad (5.3)$$

Si bien nuestro modelo admite la creación y borrado de enlaces en función del tiempo, aquí nos interesa analizar las propiedades topológicas de las redes acumuladas en el tiempo  $\mathbf{A}_T$ . Es por ello que los enlaces agregados permanecerán en nuestra descripción como si fueran estáticos. En cualquier caso, destacamos que el modelo **GSG** también incorpora un mecanismo de borrado de enlaces arbitrario.

Análogamente a lo que sucede con los enlaces, la población de nodos también puede sufrir cambios en el tiempo. Es por ello que **GSG** admite a su vez dos regímenes para los nodos: i) población constante o ii) población creciente con el tiempo. Cada nodo agregado se conecta inicialmente a un nodo existente a través de un único enlace.

### 5.2.1. Efectos de memoria no-markovianos: cierre triádico

Como hemos señalado, el problema del modelo **AS** propuesto por Perra et al tiene su origen en la suposición de agentes sin memoria de sus contactos previos. Es precisamente por ello que las redes acumuladas en el tiempo resultan esencialmente redes al azar con distribución de grados del tipo ley de potencias. Claramente, se trata de una hipótesis simplificadora sin correspondencia alguna con la observación empírica. Existe evidencia suficiente que da cuenta de correlaciones de grado de alto orden en redes sociales reales [Newman 2002a, Newman 2003b, Kossinets 2006]. Incorporar dichas correlacio-



nes de grado en nuestro modelo implica asumir memoria de los contactos previos en los agentes. Para ello, implementamos un mecanismo de conexión mixto que procede del siguiente modo para el agregado de un enlace  $l_{ij}$ :

1. Se elige el nodo fuente  $i$  con probabilidad proporcional a su actividad  $a_i$ .
2. El nodo de destino  $j$  del enlace es elegido de acuerdo con el siguiente procedimiento:
  - a. Mecanismo de *cierre triádico* (**TC**): con probabilidad  $q$  se elige un segundo vecino de  $i$ , que no se encuentre previamente vinculado a  $j$ , con el fin de cerrar un triángulo por transitividad.
  - b. Conexión aleatorio: con probabilidad  $1 - q$  se elige un nodo  $j$  uniformemente al azar.

El mecanismo de cierre triádico (**TC**) ha sido previamente utilizado con el objeto de agregar transtividad en modelos de crecimiento de redes [Kossinets 2006, Davidsen 2002, Holme 2002, Vázquez 2003, Toivonen 2006, Liben-Nowell 2007]. Por otra parte, tiene por fin reproducir el mecanismo de enlace que habitualmente se observa en las redes de amistad: los amigos de nuestros amigos con frecuencia también terminan siendo nuestros amigos. Es en este sentido en que nos referimos a *agentes con memoria*, dado que para que resulte posible establecer un vínculo entre segundos vecinos es necesario que los agentes recuerden todos sus contactos previos. Cabe destacar que el modelo **GSG** asume agentes ideales con memoria infinita en pos de simplificar el análisis ulterior.

### 5.2.2. Formulación analítica del modelo

El modelo **GSG** puede formularse matemáticamente como una proceso estocástico de tiempo continuo  $\{L_t\}_{t \in \mathbb{R}_{\geq 0}}$  con tasa de incorporación de enlaces  $\beta(t)$ . Este proceso de tiempo continuo puede representarse alternativamente de forma más simple como un proceso estocástico discreto embebido  $\{L_\ell\}_{\ell \in \mathbb{N}}$ , en términos del número de enlaces agregados  $\ell$ . En este último caso, se simplifica notablemente la formulación matemática y su posterior análisis. El proceso discreto embebido asume una tasa de incremento de la población de enlaces igual a 1, dado que por cada paso de la evolución se adiciona un enlace. De este modo, las tasas restantes del modelo estarán referidas a dicha tasa unitaria.

A fin de abordar la formulación más general del modelo, consideramos una tasa de crecimiento de la población de nodos  $\gamma$ . Por lo tanto, el número total de

enlaces vendrá dado por

$$L(\ell) = L_0 + (1 + \gamma)\ell, \quad (5.4)$$

donde  $\gamma\ell$  proviene de los enlaces asociados con los nodos agregados y  $L_0$  es la cantidad inicial de enlaces. Análogamente, la población de nodos se incrementa como

$$N(\ell) = N_0 + \gamma\ell, \quad (5.5)$$

siendo  $N_0$  la cantidad inicial de nodos en la red.

Definamos  $P_{k|a}(\ell, \ell_0)$  como la probabilidad de que un nodo con actividad  $a$  introducido en el paso  $\ell_0$  tenga grado  $k$  para un paso posterior  $\ell > \ell_0$ . Adicionalmente, definimos  $\bar{N}_{k|a}(\ell)$  como la cantidad media de nodos con grado  $k$  de entre aquellos que poseen tasa de actividad  $a$ , cuya expresión en términos de  $P_{k|a}(\ell, \ell_0)$  resulta:

$$\bar{N}_{k|a}(\ell) = \sum_{\ell_0=1}^{\ell} P_{k|a}(\ell, \ell_0). \quad (5.6)$$

En el Apéndice B mostraremos que  $\bar{N}_{k|a}(\ell)$  al propagador del modelo de variables ocultas propuesto en [Boguna 2003].

La evolución temporal de  $\bar{N}_{k|a}(\ell)$  puede formularse mediante una descripción continua a través del siguiente sistema de ecuaciones diferenciales acopladas [Krapivsky 2000, Krapivsky 2001]

$$\begin{aligned} \frac{d\bar{N}_{k|a}}{d\ell} = & q \left( \Theta(k-1, a, \ell) \bar{N}_{k-1|a} - \Theta(k, a, \ell) \bar{N}_{k|a} \right) + \frac{1-q}{N(\ell)} \left( \frac{a}{\langle a \rangle} + 1 \right) (\bar{N}_{k-1|a} - \bar{N}_{k|a}) \\ & + \frac{\gamma}{N(\ell)} (\bar{N}_{k-1|a} - \bar{N}_{k|a}) + \gamma \delta_{k1}, \end{aligned} \quad (5.7)$$

donde el primer término del miembro derecho (rhs) de la ecuación 5.7 corresponde a la contribución del mecanismo de cierre triádico **TC**, siendo  $\Theta(k, a, \ell)$  el núcleo de **TC**. El segundo término del rhs de la ecuación 5.7 corresponde a la contribución de los enlaces agregados al azar. Este último puede descomponerse en la contribución del nodo fuente elegido con probabilidad proporcional a su actividad

$$\frac{1-q}{N(\ell)} \left( \frac{a}{\langle a \rangle} \right) (\bar{N}_{k-1|a} - \bar{N}_{k|a}), \quad (5.8)$$

sumada al aporte correspondiente al nodo objetivo elegido uniformemente al azar

$$\frac{1-q}{N(\ell)} (\bar{N}_{k-1|a} - \bar{N}_{k|a}). \quad (5.9)$$

Continuando con la descripción de la ecuación 5.7, el último término del rhs corresponde a la contribución del extremo restante del único enlace que viene junto cada nodo de actividad  $a$  introducido al sistema. Cabe destacar que dicho extremo del enlace se conecta uniformemente al azar a cualquier nodo perteneciente a la red. Finalmente, el último término del rhs ( $\gamma\delta_{k1}$ ) representa el aporte a la población  $\bar{N}_{1|a}(\ell)$  de los nodos agregados con actividad  $a$  y grado inicial  $k = 1$ .

En cuanto al núcleo del mecanismo **TC** de la ec. 5.7, ( $\Theta(k, a, \ell)$ ), viene definido como

$$\Theta(k, a, \ell) = \frac{a}{\langle a \rangle N(\ell)} + \frac{k}{2L(\ell)}. \quad (5.10)$$

Uno de los extremos de cada enlace agregado mediante el mecanismo **TC** es conectado a un nodo fuente elegido al azar proporcionalmente a su actividad  $a$ , dando origen al primer término del rhs de la ec. 5.10. Mientras que el segundo término del rhs proviene de la conexión del extremo restante del enlace agregado a un segundo vecino del nodo fuente. Este último término representa el enlace preferencial que surge naturalmente a partir del mecanismo **TC**, resultado que se corresponde con los previamente reportados en [Holme 2002, Vázquez 2003].

Reagrupando términos, la ec. 5.7 puede reescribirse como

$$\frac{d\bar{N}_{k|a}}{d\ell} = \bar{N}_{k-1|a}\Phi_{\gamma,q}(k-1, a, \ell) - \bar{N}_{k|a}\Phi_{\gamma,q}(k, a, \ell) + \gamma\delta_{k1} \quad (5.11)$$

donde  $\Phi_{\gamma,q}(k, a, \ell)$  es el núcleo generalizado de conexión dado por:

$$\Phi_{\gamma,q}(k, a, \ell) = \frac{1}{N(\ell)} \left( \frac{a}{\langle a \rangle} + 1 - q + \gamma \right) + \frac{k}{2L(\ell)}q. \quad (5.12)$$

Luego, la expresión correspondiente al número medio de nodos con grado  $k$ , notada  $\bar{N}_k(\ell)$ , comprendiendo el aporte de todas las tasas de actividad posibles con pdf  $F(a)$ , queda definida como:

$$\bar{N}_k(\ell) = \int_{\Omega} F(a)\bar{N}_{k|a}(\ell)da \quad (5.13)$$

siendo  $\Omega$  el dominio de definición de  $F(a)$ . Por último, podemos definir la distribución de grados para una dada cantidad de enlaces agregados  $\ell$  como

$$P_{\ell}(k) = \bar{N}_k(\ell)/N(\ell). \quad (5.14)$$

El modelo **GSG** aquí descripto presenta una amplia versatilidad, permitiendo

modificar tanto la distribución de actividades como los regímenes de crecimiento de las poblaciones de nodos y enlaces. Es por ello que en las subsecciones siguientes estudiaremos las soluciones de la ec. 5.11 para población de nodos constante ( $\gamma = 0$ ) y creciente ( $\gamma > 0$ ). Alternativamente, estudiaremos dos pdf de actividad paradigmáticas: (i) actividad constante  $F(a) = \delta(a - a_0)$  y (ii) actividad del tipo ley de potencias  $F(a) \propto a^{-\xi}$ , que reproduce adecuadamente el patrón de actividad de los agentes para algunas redes sociales analizadas en la literatura [Perra 2012, Muchnik 2013].

### 5.2.3. Población constante ( $\gamma = 0$ )

En este caso particular, el crecimiento de la red solo tiene lugar a través del agregado de enlaces sobre una población constante de nodos. Sustituyendo  $\gamma = 0$  en la ec. 5.11 y luego de reemplazar en la misma la correspondiente expresión para  $\Phi_{0,q}(k, a, \ell)$ , se obtiene:

$$\begin{aligned} \frac{d\bar{N}_{k|a}}{d\ell} = & \bar{N}_{k-1|a} \left[ \frac{1}{N_0} \left( \frac{a}{\langle a \rangle} + 1 - q \right) + \frac{(k-1)q}{2L(\ell)} \right] - \\ & \bar{N}_{k|a} \left[ \frac{1}{N_0} \left( \frac{a}{\langle a \rangle} + 1 - q \right) + \frac{kq}{2L(\ell)} \right] \end{aligned} \quad (5.15)$$

donde se asume  $N(\ell) = N_0 \forall \ell \geq 0$  y la población de enlaces crece como  $L(\ell) = L_0 + \ell$ . En estas circunstancias, el grado  $k$  presenta una restricción natural dada por  $k \leq (N_0 - 1)$ . Asimismo, la solución asintótica a la ec. 5.15 adopta la forma trivial  $\bar{N}_{k|a} = N_0 \delta_{k, (N_0 - 1)} \forall a \geq 0$ . No obstante, en este caso nos interesa analizar la solución no-trivial transitoria. Esto se debe a que el régimen de población constante de nodos generalmente se encuentra asociado con el análisis de redes sociales sobre una ventana temporal lo suficientemente pequeña para que dicha hipótesis resulte adecuada.

Es posible resolver la ec. 5.15 por medio de un esquema iterativo como sigue:

$$\bar{N}_{k|a}(\ell) = \left( \bar{N}_{k|a}(0) + \int_0^\ell \frac{\bar{N}_{k-1|a}(\ell')}{\bar{\Pi}_{k|a}(\ell')} \Phi_{0,q}(k-1, a, \ell') d\ell' \right) \bar{\Pi}_{k|a}(\ell) \quad (5.16)$$

donde  $\bar{\Pi}_{k|a}(\ell) = A (L_0 + \ell)^{-kq/2} \exp(- (a/\langle a \rangle + 1 - q) \ell / N_0)$  es solución de la ecuación diferencial

$$\frac{d\bar{\Pi}_{k|a}(\ell)}{d\ell} = -\bar{\Pi}_{k|a}(\ell) \Phi_{0,q}(k, a, \ell). \quad (5.17)$$

Por otra parte, es posible encontrar soluciones analíticas cerradas a la ec. 5.15 en algunos casos particulares. Por ejemplo, si  $q = 0$  puede deducirse fácilmente que la solución a la ec. 5.15 resulta

$$\bar{N}_{k|a}(\ell) = N_0 \times \text{Pois}(k; \lambda = \ell(a + \langle a \rangle) / (\langle a \rangle N_0)), \quad (5.18)$$

siendo  $\text{Pois}(k; \lambda) = (\lambda^k / k!) \exp(-\lambda)$  la distribución de Poisson con media  $\lambda$ . Reemplazando la ec. 5.18 en la ec. 5.13 se obtiene

$$P_\ell(k) = \frac{\bar{N}_k(\ell)}{N_0} \sim \frac{1}{\ell} F\left(\frac{N_0}{\ell} k - 1\right), \quad (5.19)$$

recuperando así la solución del modelo de actividad social (**AS**) de la ec. 5.1 [Perra 2012].

Nuestro objetivo es encontrar soluciones aproximadas a la ec. 5.15 bajo condiciones extremas, que nos proporcionen una intuición del comportamiento exacto de  $\bar{N}_{k|a}(\ell)$  y, luego, de la distribución de grados asociada  $P_\ell(k)$ . Con tal fin, asumiremos inicialmente la condición extrema

$$\mathbf{a.} \quad \frac{1}{N_0} \left( \frac{a}{\langle a \rangle} + 1 - q \right) \gg q \frac{k}{2(L_0 + \ell)}, \quad (5.20)$$

en virtud de la cual la solución aproximada de la ec. 5.15 resulta

$$\bar{N}_{k|a}(\ell) \sim \frac{1}{k!} \left( \frac{a/\langle a \rangle + 1 - q}{N_0} \ell \right)^k e^{-\frac{a/\langle a \rangle + 1 - q}{N_0} \ell} \quad (5.21)$$

es decir, una Poisson con media  $\lambda = \ell(a/\langle a \rangle + 1 - q)/N_0$ . La solución de la ec. 5.21 resulta intuitiva dado que la condición **a.** implica despreciar la acción del término de enlace preferencial respecto del correspondiente a enlace al azar pesado por la actividad de los nodos.

Alternativamente, la condición opuesta a **a.**:

$$\mathbf{b.} \quad \frac{1}{N_0} \left( \frac{a}{\langle a \rangle} + 1 - q \right) \ll q \frac{k}{2(L_0 + \ell)}, \quad (5.22)$$

conduce a otra solución aproximada a la ec. 5.15, en este caso dada por

$$\bar{N}_{k|a}(\ell) \sim (L_0 + \ell)^{-q/2} \left[ 1 - \left( \frac{L_0 + \ell}{L_0} \right)^{-q/2} \right]^{k-1} \approx (L_0 + \ell)^{-q/2} e^{-(k-1) \left( \frac{L_0 + \ell}{L_0} \right)^{-q/2}}, \quad (5.23)$$

que muestra decaimiento exponencial independiente de la actividad  $a$ . En el caso de la condición 5.22, el mecanismo de enlace preferencial es el que se impone. Por ello, la solución con decaimiento exponencial es consistente con la esperada para un mecanismo de enlace preferencial con población de nodos constante [Barabasi 1999b].

La expresión de  $\bar{N}_k(\ell)$  se obtiene a partir de la ec. 5.13 para cada una de las pdf de actividad  $F(a)$  seleccionadas, en nuestro caso: constante ( $F(a) = \delta(a - a_0)$ ) o ley de potencias ( $F(a) \propto a^{-\xi}$ ).

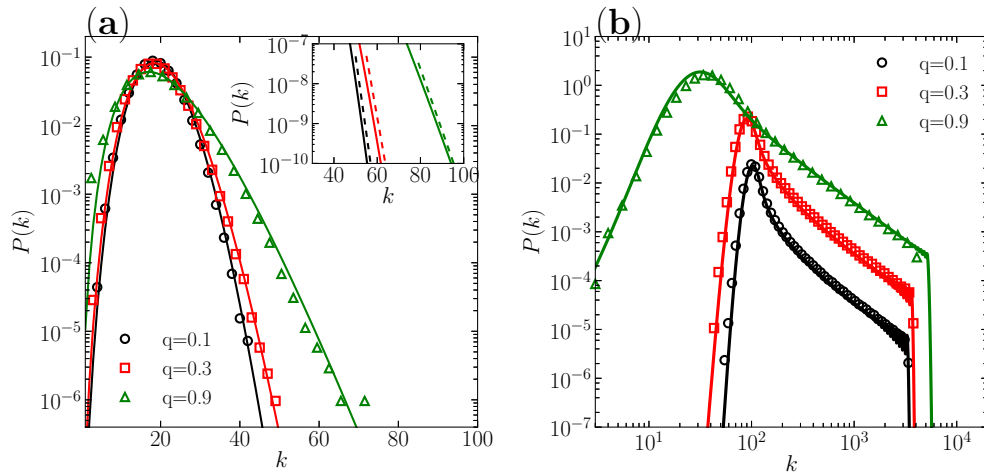


Figura 5.2: Distribución de grado  $P_\ell(k) = \bar{N}_k(\ell)/N_0$  para el régimen de población constante. **(a)** Distribución de actividad  $F(a) = \delta(a - a_0)$ , red acumulada en el tiempo con  $N = 10^5$  y  $\langle k \rangle = 20$ , **(b)** Distribución de actividad  $F(a) \propto a^{-1.5}$ , red acumulada con  $N = 10^5$  y  $\langle k \rangle = 200$ . En ambos casos los símbolos corresponden a los promedios sobre 100 simulaciones numéricas para:  $q = 0.1$  (círculos negros),  $q = 0.3$  (cuadrados rojos) y  $q = 0.9$  (triángulos verdes). Con el objeto de simplificar la visibilidad, los gráficos para distintos valores de  $q$  han sido desplazados verticalmente en todos los casos. Las líneas sólidas corresponden a la solución numérica de la ec. 5.16, luego integrada en la ec. 5.13, con el objeto de obtener  $P_\ell(k) = \bar{N}_k(\ell)/N(\ell)$ . **(Inset en (a))** Detalle para valores grandes de  $k$  a fin de comparar con el decaimiento exponencial de la solución aproximada dada por la ec. 5.23 (línea de trazos).

Ensayamos simulaciones computacionales del algoritmo **GSG** generando redes con una población constante de  $N = 10^5$  nodos, para las dos pdf de actividad  $F(a)$  consideradas. Simultáneamente, realizamos la integración numérica de las ecuaciones 5.15 y 5.14 con el fin de comparar las distribuciones de grados  $P_\ell(k) = \bar{N}_k(\ell)/N(\ell)$  con las correspondientes para las redes de las simulaciones de **GSG**. La Figura 5.2 muestra el muy buen acuerdo entre las soluciones

analíticas de  $P_\ell(k)$  y las obtenidas de las simulaciones del algoritmo **GSG**, para  $F(a) = \delta(a - a_0)$  (Figura 5.2-a) y  $F(a) \propto a^{-\xi}$  (Figura 5.2-b) con distintos valores de  $q$ .

#### 5.2.4. Población creciente

El caso alternativo al analizado en la subsección anterior corresponde a asumir población creciente, es decir, a considerar  $\gamma > 0$  en la ec. 5.7. Luego de agrupar términos, la ec. 5.7 puede reescribirse como:

$$\begin{aligned} \frac{d\bar{N}_{k|a}}{d\ell} = & \bar{N}_{k-1|a} \left[ \frac{1}{N(\ell)} \left( \frac{a}{\langle a \rangle} + 1 - q + \gamma \right) + \frac{(k-1)q}{2L(\ell)} \right] \\ & - \bar{N}_{k|a} \left[ \frac{1}{N(\ell)} \left( \frac{a}{\langle a \rangle} + 1 - q + \gamma \right) + \frac{kq}{2L(\ell)} \right] + \gamma \delta_{k1} \end{aligned} \quad (5.24)$$

donde ahora  $N(\ell) = N_0 + \gamma\ell$  y  $L(\ell) = L_0 + (1 + \gamma)\ell$ . Al contrario de lo que ocurría para el caso con población constante ( $\gamma = 0$ ) ahora es posible obtener soluciones asintóticas no-triviales a la ec. 5.24. Siguiendo [Krapivsky 2000] puede concluirse que esta clase de soluciones asumen la forma general  $\bar{N}_{k|a}(\ell) = n_{k|a}\ell$ , siendo  $n_{k|a}$  una función desconocida dependiente del grado  $k$  y de la tasa de actividad  $a$ . En el límite asintótico ( $\ell \rightarrow \infty$ ) la ec. 5.24 puede resolverse a través de un procedimiento iterativo arrojando la siguiente solución general:

$$n_{k|a} = n_{1|a} \prod_{j=1}^{k-1} \frac{2(\gamma+1)(a/\langle a \rangle + 1 - q + \gamma) + q\gamma j}{2(\gamma+1)(a/\langle a \rangle + 1 - q + 2\gamma) + q\gamma(j+1)} \quad (5.25)$$

siendo  $n_{1|a}$  la solución correspondiente al borde  $k = 1$ , dada por

$$n_{1|a} = \frac{2\gamma^2(\gamma+1)}{2(\gamma+1)(a/\langle a \rangle + 1 - q + 2\gamma) + q\gamma}. \quad (5.26)$$

Más allá de la expresión rigurosamente exacta para  $n_{k|a}$ , pretendemos adquirir intuición a partir del análisis de su comportamiento en escenarios aproximados. Al igual que para población constante, comenzaremos considerando la situación extrema en la que puede despreciarse el término de enlace preferencial en la ec. 5.24:

$$\mathbf{a.} \quad \left( \frac{a}{\langle a \rangle} + 1 - q + \gamma \right) \langle k \rangle_\ell \gg qk \quad (5.27)$$

siendo  $\langle k \rangle_\ell = 2L(\ell)/N(\ell)$  el grado medio luego de  $\ell$  enlaces agregados (o pasos del algoritmo **GSG**).

Introduciendo la aproximación 5.27 en la ec. 5.24 y sustituyendo  $\bar{N}_{k|a}(\ell) = n_{k|a}\ell$ , obtenemos la solución asintótica correspondiente a este escenario aproximado:

$$n_{k|a} \sim \left( 1 + \frac{\gamma}{a/\langle a \rangle + 1 - q + \gamma} \right)^{-(k-1)} \quad (5.28)$$

que muestra un claro decaimiento exponencial.

Veamos ahora la aproximación alternativa a 5.27,

$$\mathbf{b.} \quad \left( \frac{a}{\langle a \rangle} + 1 - q + \gamma \right) \langle k \rangle_\ell \ll qk. \quad (5.29)$$

Bajo la condición 5.29, la ec. 5.25 toma la siguiente forma asintótica

$$n_{k|a} \sim k^{-1 - \frac{2(\gamma+1)}{q}}, \quad (5.30)$$

mostrando un comportamiento del tipo ley de potencias con exponente  $\alpha = 1 + \frac{2(\gamma+1)}{q}$ , que resulta  $\alpha > 3$  cuando  $0 < q \leq 1$  y  $\gamma > 0$ . Una consecuencia importante es que  $n_{k|a}$  resulta independiente de  $a$ , de modo que  $\bar{N}_k(\ell)$  de la ec. 5.13 hereda idéntica propiedad, adoptando el mismo exponente en su forma asintótica:

$$\bar{N}_k \sim k^{-1 - \frac{2(\gamma+1)}{q}}. \quad (5.31)$$

La ec. 5.31 indica que el comportamiento asintótico del tipo ley de potencias mostrado por  $\bar{N}_k(\ell)$  resulta independiente de la pdf de actividad  $F(a)$  y se sostiene solo a partir del mecanismo **TC** de cierre triádico. Nuevamente, la combinación de un término de enlace preferencial con crecimiento poblacional da lugar a distribución de grados del tipo ley de potencias, tal como se establece en [Barabasi 1999b].

En la Figura 5.3 se observa el muy buen acuerdo entre los resultados de las simulaciones numéricas y la solución analítica de la ec. 5.24 para  $P(k)$ , bajo  $F(a) = \delta(a - a_0)$ . En línea de trazos se presentan superpuestas las soluciones aproximadas que provienen de las ec. 5.28 y 5.30. La Figura 5.4 presenta idéntica comparación, aunque en este caso considerando  $F(a) \propto a^{-\xi}$  con  $\xi = 1.5$ .

En las Figuras 5.2, 5.3 y 5.4 puede verse que el comportamiento de  $P(k)$  para  $F(a) \propto a^{-\xi}$ , es dominado por esta última ley de potencias en la región intermedia, lejos de los casos extremos analizados.



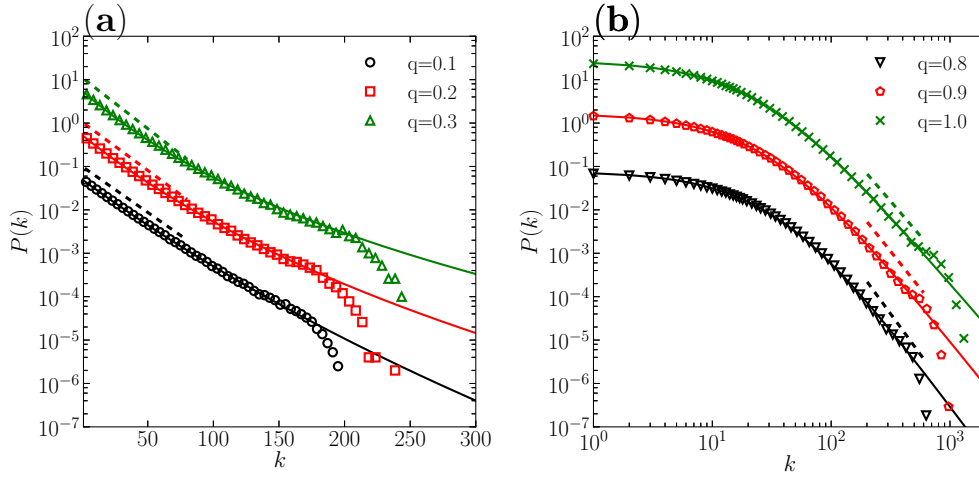


Figura 5.3: Distribuciones de grado  $P_\ell(k) = \bar{N}_k(\ell)/N(\ell)$  para régimen con población creciente y  $F(a) = \delta(a - a_0)$ , para redes acumuladas en el tiempo con población final  $N = 10^5$  y  $\langle k \rangle = 20$ . Los símbolos corresponden a promedios sobre 100 simulaciones numéricas para: **(a)**  $q \in \{0.1, 0.2, 0.3\}$ , y **(b)**  $q \in \{0.8, 0.9, 1.0\}$ . Las líneas sólidas corresponden a las soluciones numéricas de la ec. 5.24 apropiadamente normalizada para obtener  $P(k)$ . Las tendencias de las soluciones asintóticas de las ecs. 5.28 y 5.30 fueron graficadas mediante líneas de trazos. Nuevamente, los gráficos correspondientes a diferentes valores de  $q$  han sido convenientemente desplazados para facilitar su visualización.

### 5.3. Correlaciones de grado de alto orden

En esta sección analizaremos las características de las correlaciones de grado representadas a partir de dos indicadores: el grado medio de los primeros vecinos ( $k_{nn}(k)$ ) y el coeficiente de clusterización medio ( $\bar{C}(k)$ ), ambos en función del grado  $k$ . Tal como hemos mencionado, las correlaciones de grado de alto orden constituyen una signatura de las redes de origen social. Es por ello que otorgamos particular atención a este aspecto. No obstante, la formulación analítica en estos casos se torna rápidamente intratable, por lo que solo abordaremos un tratamiento numérico a partir de simulaciones de nuestro modelo.

#### 5.3.1. Grado medio de primeros vecinos

Una medición indirecta de la correlación de grados de segundo orden puede obtenerse a partir del cálculo del grado medio de los nodos de grado  $k$  ( $k_{nn}(k)$ ),

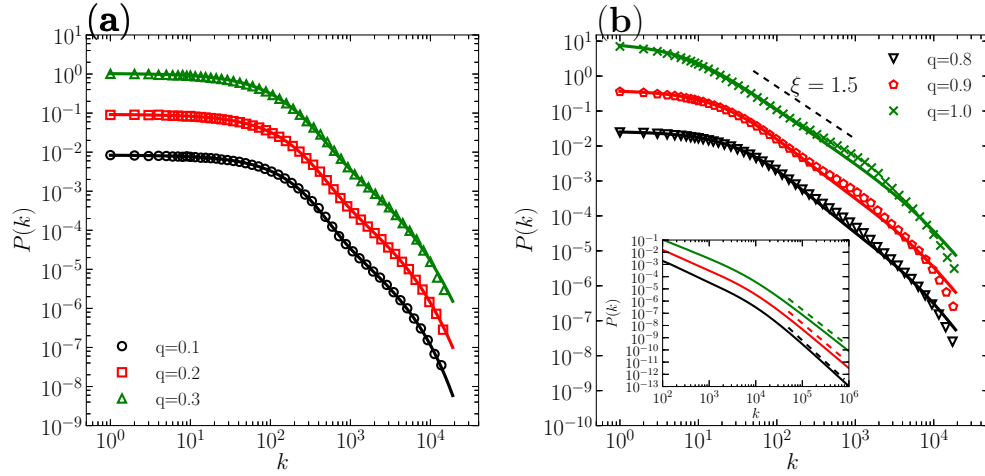


Figura 5.4: Distribución de grado  $P_\ell(k) = \bar{N}(\ell)_k/N(\ell)$  para el régimen de población creciente con  $F(a) \propto a^{-1.5}$ , redes con población final  $N = 10^5$  y  $\langle k \rangle = 200$ . Los símbolos corresponden a promedios sobre 100 simulaciones numéricas para: **(a)**  $q \in \{0.1, 0.2, 0.3\}$ , y **(b)**  $q \in \{0.8, 0.9, 1.0\}$ . Las líneas sólidas corresponden a las soluciones numéricas de la ec. 5.24, promediadas mediante la ec. 5.13 y apropiadamente normalizada para obtener  $P(k)$ . Con el objeto de simplificar la visibilidad, los gráficos para distintos valores de  $q$  han sido desplazados verticalmente en todos los casos. Las líneas sólidas corresponden a las soluciones numéricas de la ec. 5.24 apropiadamente normalizada para obtener  $P(k)$ . Las tendencias de las soluciones asintóticas de las ecs. 5.28 y 5.30 fueron graficadas mediante líneas de trazos. **(Inset en (b))** Detalle para valores grandes de  $k$  a fin de comparar con el decaimiento del tipo ley de potencias de la solución asintótica aproximada dada por la ec. 5.31 (línea de trazos).

formalmente definida como [Pastor-Satorras 2001a]:

$$\bar{k}_{nn}(k) = \sum_{k'} k' P(k'|k), \quad (5.32)$$

donde  $P(k'|k)$  es la probabilidad condicional que tiene un nodo de grado  $k$  de estar conectado a otro con grado  $k'$ . De esta forma, la dependencia de  $\bar{k}_{nn}$  con  $k$  da cuenta de la existencia de correlación grado-grado. En particular, en aquellos casos para los cuales  $\bar{k}_{nn}$  crece (decrece) con  $k$  diremos que la red presenta asortatividad (deasortatividad) de grado [Newman 2002a, Newman 2003a]. Asortatividad de grado es precisamente el comportamiento que observamos para el régimen de población creciente tanto con  $F(a)$  constante como ley de potencias, tal como se observa en la Figura 5.5 (b) y (d). Por el contrario, el modelo **GSG** con población constante y  $F(a) = \delta(a - a_0)$  presenta un gráfico cuasi-constante

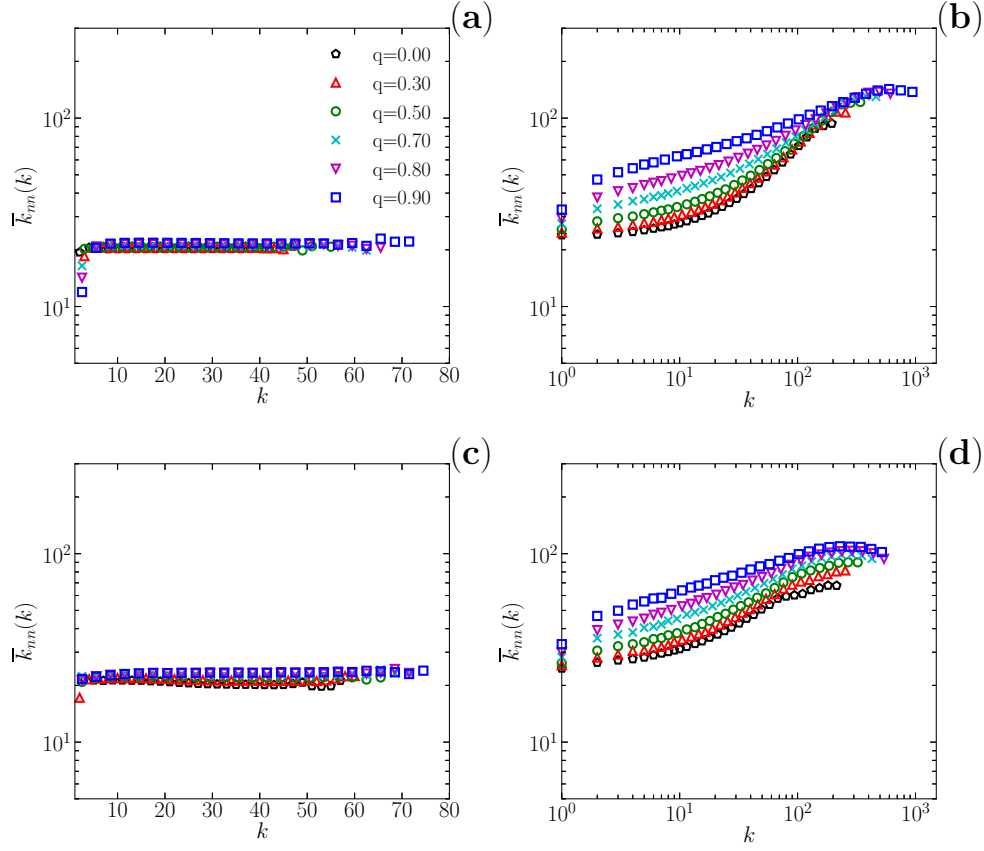


Figura 5.5: Grado medio de primeros vecinos de un nodo con grado  $k$ ,  $\bar{k}_{nn}(k)$ , para las redes obtenidas mediante el modelo **GSG**, con población final  $N = 10^5$  y  $\langle k \rangle = 20$  en (a) y (b), mientras que  $N = 10^4$  y  $\langle k \rangle = 20$  para (c) y (d). Los símbolos corresponden a los promedios sobre 100 simulaciones numéricas bajo las siguientes condiciones: (a) Población constante y pdf de actividad  $F(a) = \delta(a - a_0)$ ; (b) población creciente y pdf de actividad  $F(a) = \delta(a - a_0)$ ; (c) población constante y pdf de actividad  $F(a) \propto a^{-1.5}$ ; (d) población creciente y pdf de actividad  $F(a) \propto a^{-1.5}$ .

de  $\bar{k}_{nn}(k)$  (Figura 5.5-(a)). Este comportamiento es compatible con el correspondiente a redes aleatorias de Erdős-Rényi (ER), para las cuales  $\bar{k}_{nn}(k') = \langle k \rangle \forall k'$ , aunque en breve mostraremos que tal correspondencia es solo parcial. Para el caso alternativo con  $F(a) \propto a^{-1.5}$  el gráfico de  $\bar{k}_{nn}(k)$  es también plano aunque levemente deasortativo (Figura 5.5-(c)).

Para el caso particular con  $q = 0$  y  $\gamma = 0$ , puede demostrarse (ver Apéndice B para una derivación formal) que se satisface la siguiente ley de escala que

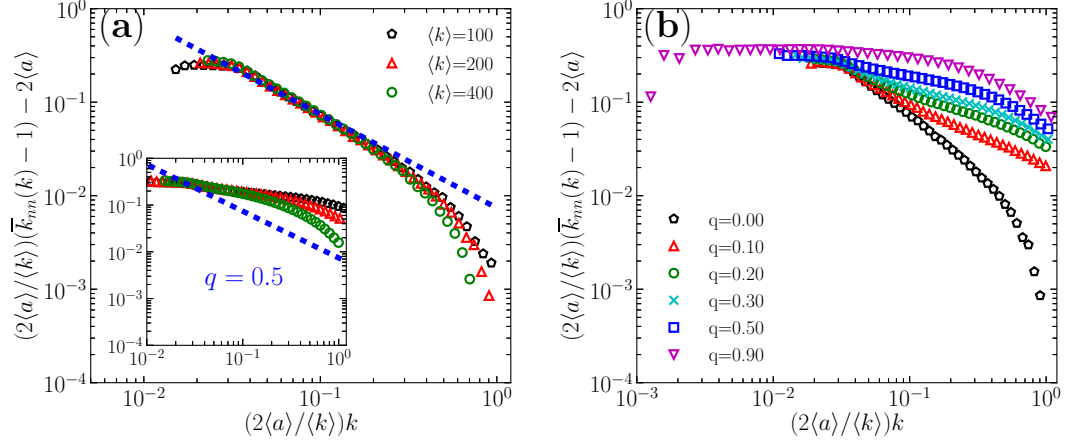


Figura 5.6: Grado medio de primeros vecinos de un nodo de grado  $k$  reescalado, para redes acumuladas en el tiempo obtenidas mediante el modelo **GSG** con pdf de actividad  $F(a) \propto a^{-1.5}$  y población constante  $N = 10^4$ . Los símbolos corresponden al promedio sobre 100 simulaciones numéricas. **(a)** Colapso de las soluciones para  $q = 0$  y  $\langle k \rangle = 100, 200$  y  $400$  (el inset corresponde a  $q = 0.5$ ). La línea de trazos representa la predicción de la función de escala de la ec. 5.33. **(b)**  $\bar{k}_{nn}(k)$  reescalado para distintos valores  $q \geq 0$ .

involucra a  $\bar{k}_{nn}(k)$ :

$$\frac{2\langle a \rangle}{\langle k \rangle} (\bar{k}_{nn}(k) - 1) - 2\langle a \rangle \simeq (\langle a^2 \rangle - \langle a \rangle^2) \left( \frac{2\langle a \rangle}{\langle k \rangle} k \right)^{-1}. \quad (5.33)$$

En la Figura 5.6-(a) se muestra el colapso predicho por la ley de escala de la ec. 5.33 para diferentes valores de  $\langle k \rangle$ , con  $q = \gamma = 0$ . Es muy difícil deducir una expresión similar a la de la ec. 5.33 para  $q > 0$  debido a las correlaciones locales inducidas por el mecanismo **TC**. No obstante, en la Figura 5.6-(b) hemos graficado  $(2\langle a \rangle / \langle k \rangle) (\bar{k}_{nn}(k) - 1) - 2\langle a \rangle$  como función de  $2\langle a \rangle k / \langle k \rangle$  para  $q > 0$ , mostrando que de existir una función de escala (como sugiere el inset de la Figura 5.14-(a) para  $q = 0.5$ ), esta deberá presentar una dependencia con  $q$  de la forma  $(2\langle a \rangle / \langle k \rangle) (\bar{k}_{nn}(k) - 1) - 2\langle a \rangle \simeq F_{scal}^q(2\langle a \rangle k / \langle k \rangle)$ .

## 5.4. Coeficiente de clusterización

Otro indicador de la existencia de correlaciones de grado es el coeficiente de clusterización medio  $\bar{C}(k)$  definido en la ec. 2.9 como:

$$\bar{C}(k) = \frac{1}{N_k} \sum_{i \in \text{Deg}(k)} C_i, \quad (5.34)$$

donde  $\text{Deg}(k)$  es el conjunto de nodos con grado  $k$ , con  $N_k$  su cardinal, y  $C_i$  es el coeficiente de clusterización local definido en la ec. 2.6.

El coeficiente de clusterización se encuentra asociado con la frecuencia de aparición de enlaces transitivos o, visto en forma geométrica, como la probabilidad de ocurrencia de triadas de nodos interconectados (triángulos) en la red. Las redes de origen social usualmente exhiben una ley de escala para  $\bar{C}(k)$

$$\bar{C}(k) \sim k^{-\beta}, \quad (5.35)$$

cuyo exponente satisface  $\beta \lesssim 1$  para muchas de las redes analizadas en la literatura [Klimek 2013]. Esta ley de escala es satisfecha por las redes acumuladas en el tiempo generadas mediante el modelo **GSG**, tal como se muestra en la Figura 5.7. Este comportamiento muestra que las redes generadas mediante **GSG** con población y pdf de actividad constantes no se comportan como redes aleatorias de **ER** como sugería la Figura 5.5-(a), dado que en dicho caso deberían presentar el  $\bar{C}(k)$  constante correspondiente a redes aleatorias, cuya expresión exacta es [Dorogovtsev 2004]

$$\bar{C}(k) = \frac{(\langle k^2 \rangle - \langle k \rangle)^2}{N \langle k \rangle^3}. \quad (5.36)$$

Por último, en la Figura 5.8 mostramos el coeficiente de clusterización promedio de la red, definido como

$$\bar{C} = \frac{1}{N} \sum_{i=1}^N C_i = \sum_k P(k) \bar{C}(k), \quad (5.37)$$

en función de la probabilidad de **TC**  $q$ . Se observa un comportamiento creciente de  $\bar{C}$  con  $q$ , tal como resulta esperable por definición.

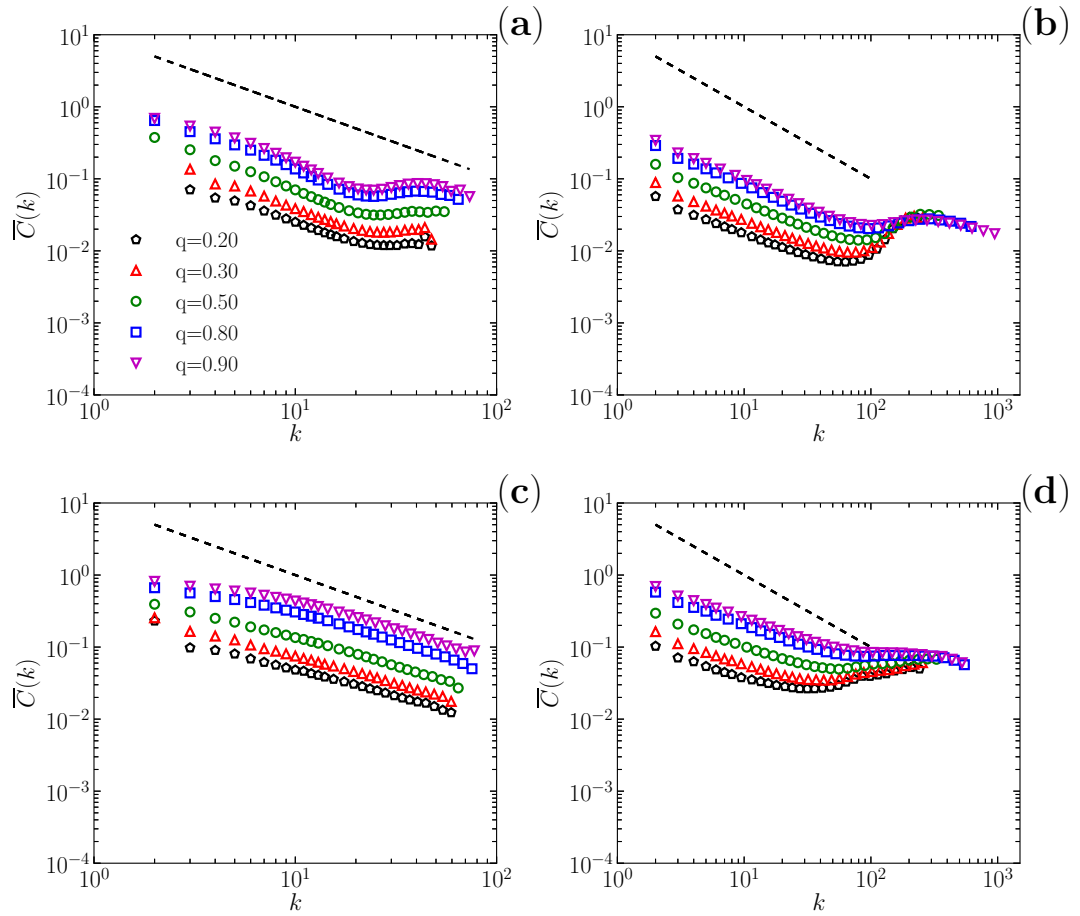


Figura 5.7: Coeficiente de clusterización medio  $\bar{C}(k)$  como función del grado  $k$  para redes acumuladas en el tiempo obtenidas mediante el modelo **GSG**, con población final  $N = 10^5$  y  $\langle k \rangle = 20$  en (a) y (b), mientras que  $N = 10^4$  y  $\langle k \rangle = 20$  para (c) y (d). Los símbolos corresponden a los promedios sobre 100 simulaciones numéricas bajo las siguientes condiciones: (a) Población constante y pdf de actividad  $F(a) = \delta(a - a_0)$ ; (b) población creciente y pdf de actividad  $F(a) = \delta(a - a_0)$ ; (c) población constante y pdf de actividad  $F(a) \propto a^{-1.5}$ ; (d) población creciente y pdf de actividad  $F(a) \propto a^{-1.5}$ . Se muestran como referencia líneas de trazos con pendiente  $-1$  correspondientes a  $\bar{C}(k) \sim k^{-1}$ .

## 5.5. Casos de estudio

En esta sección estudiaremos dos ejemplos de redes sociales con el fin de efectuar una comparación con los resultados del modelo **GSG**. La comparación será semi-cualitativa dado que no pretendemos reproducir con exactitud los resultados de alguna red social en particular (para ello necesitaríamos incorporar más pará-

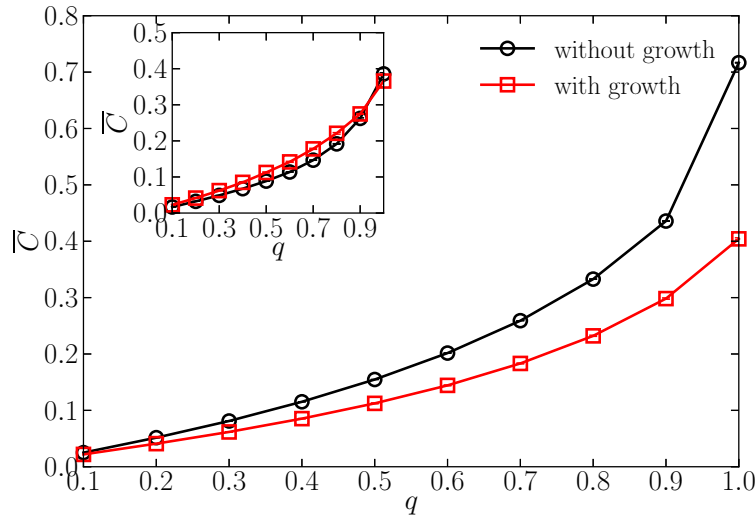


Figura 5.8: Coeficiente de clusterización promedio  $\bar{C}$  para la red en función de la probabilidad  $q$  de enlace **TC**. Todos los resultados corresponden a promedios sobre 100 realizaciones de **GSG** con  $N = 10^4$ ,  $\langle k \rangle = 20$  y  $F(a) \propto a^{-1.5}$ . Inset: resultados para el modelo **GSG** con los parámetros previos pero con actividad constante  $F(a) = \delta(a - a_0)$ .

metros y mecanismos específicos al modelo), sino solo extraer sus características fundamentales.

El primer caso que analizaremos corresponde a un estudio experimental de contactos sociales por proximidad realizado por la colaboración SocioPatterns (<http://www.sociopatterns.org/>). Con el objeto de registrar los contactos sociales entre los asistentes a una reunión social, se disponen emisores de radiofrecuencia en el torso de cada individuos, emitiendo pulsos correspondientes al id del sujeto con resolución temporal de 20 segundos y un alcance de  $\sim 1\text{m}$  [Cattuto 2010]. De este modo, el experimento no solo registra los encuentros sino también su duración. Uno de estos experimentos fue realizado durante la conferencia ACM Hypertext 2009 (**HT**) realizada en Turin, Italia. Si bien se registraron los contactos sociales durante los dos días de la conferencia, nosotros trabajaremos solo con los datos provenientes del primer día. La red **HT** cuenta con  $N = 113$  nodos y  $L = 2196$  enlaces bidireccionales [Isella 2011].

El segundo caso que analizaremos es un subgrafo de la red social Facebook (**FG**) correspondiendo a  $N = 63731$  usuarios (con  $N_{cc} = 63392 \approx 0.995 \times N$  integrando la máxima componente conectada) pertenecientes a la ciudad de New Orleans, Louisiana (EEUU), vinculados por  $L = 817090$  enlaces que representan relaciones de amistad bidireccionales. El subgrafo **FG** fue reportado

en [Viswanath 2009]. Tal como en el caso anterior, este grafo también aporta información temporal a través de la fecha de comienzo de la amistad. Al contrario de lo que ocurre con **HT**, la red **FG** exhibe crecimiento de la población en un contexto sin restricciones espaciales naturales.

### 5.5.1. Propiedades topológicas

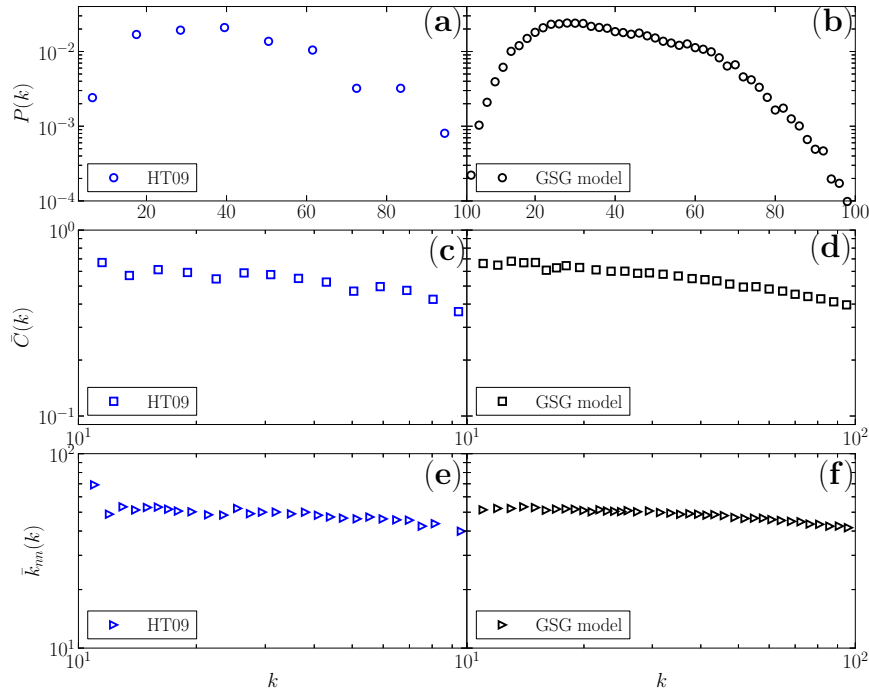


Figura 5.9: Distribución de grado  $P(k)$ , coeficiente de clusterización medio  $\bar{C}(k)$  y grado medio de primeros vecinos  $\bar{k}_{nn}(k)$  para (a),(c),(e) la red acumulada en el tiempo **HT**, y (b),(d),(f) para el modelo **GSG** con población constante  $N = 113$ ,  $\langle k \rangle = 38.9$ , probabilidad de **TC**  $q = 0.8$  y pdf de actividad del tipo ley de potencias con exponente  $\xi = 0.79$ .

En la Figura 5.9 se muestra  $P(k)$ ,  $\bar{C}(k)$  y  $\bar{k}_{nn}(k)$  para la red **HT**. En cuanto a la distribución de grados  $P(k)$ , presenta un comportamiento Poissoniano para  $k$  pequeño, con un decaimiento aproximadamente exponencial compatible con el obtenido para nuestro modelo **GSG** con población constante (ver Figura 5.9-(a)). Por el contrario, la red **FG** muestra una  $P(k)$  con un comportamiento del tipo  $P(k) \sim k^{-\alpha}$ , con  $\alpha \approx 3.4$  para  $k$  grandes (ver Figura 5.10-(a)). Este último comportamiento también resulta compatible con el observado para **GSG**, pero en este caso con población de nodos creciente. De igual modo, tanto



$\bar{C}(k)$  como  $\bar{k}_{nn}(k)$  para las redes **HT** y **FG**, exhiben cualitativamente el mismo comportamiento que nuestro modelo **GSG** para población constante y creciente, respectivamente.

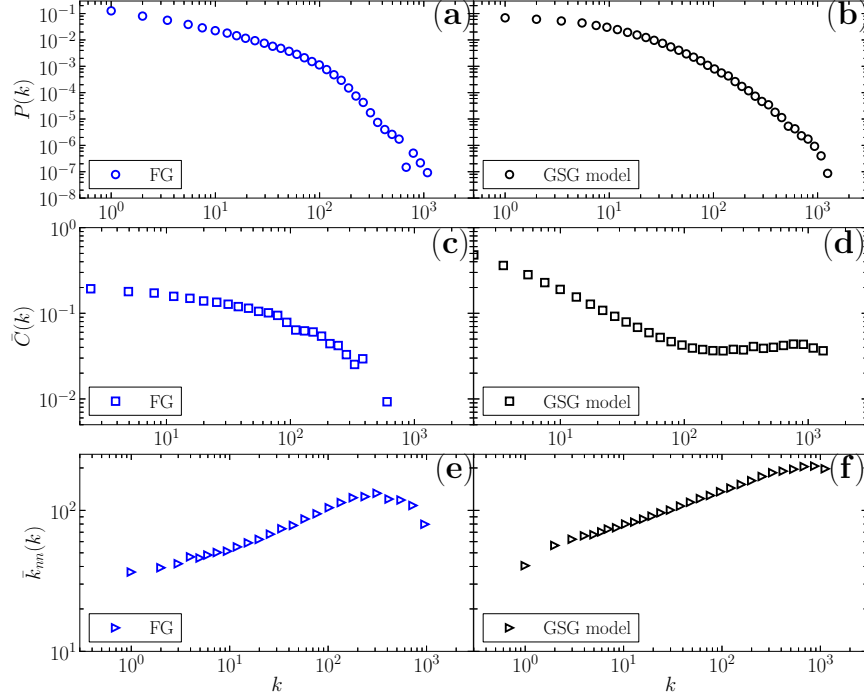


Figura 5.10: Distribución de grado  $P(k)$ , coeficiente de clusterización medio  $\bar{C}(k)$  y grado medio de primeros vecinos  $\bar{k}_{nn}(k)$  para (a),(c),(e) la red acumulada en el tiempo **FG**, y (b),(d),(f) para el modelo **GSG** para población creciente, con  $N = 63731$ ,  $\langle k \rangle = 25.6$ , probabilidad de enlaces **TC**  $q = 0.8$  y pdf de actividad del tipo ley de potencias con exponente  $\xi = 1.5$ .

Para precisar la comparación del párrafo previo, realizamos una simulación del modelo **GSG** con población constante  $N = 113$ ,  $L = 2196$  ( $\langle k \rangle = 38.9$ ), con una pdf de actividad de la forma  $F(a) \sim a^{-\xi}$ , y fijando la probabilidad de cierre triádico  $q = 0.8$  (cuya elección será justificada en breve). De este modo, el único parámetro de ajuste de nuestro modelo resulta el exponente  $\xi$  de  $F(a)$ . Ajustando  $\xi$  por el método de máxima verosimilitud (MLE) usando el modelo **GSG** con población constante sobre los datos correspondientes a  $P(k)$  para la red **HT**, obtuvimos el gráfico de la Figura 5.9-(b). El ajuste se realizó considerando  $\xi \in [0.5, 3]$  con un paso variable  $\delta\xi = 0.01-0.1$ . Los gráficos de las Figuras 5.9-(d) y (f) no fueron ajustados sino que corresponden a los parámetros previamente fijados y al  $\xi$  obtenido del ajuste de  $P(k)$ . El comportamiento del modelo resulta en todos los casos cualitativamente semejante al observado para la red real **HT**.

En cuanto a la red **FG**, procedimos de idéntica forma, aunque en este caso ajustando a través del modelo **GSG** con población creciente, partiendo de una población inicial  $N_0 = 100$  con una tasa de crecimiento de nodos  $\lambda = N/L$ , siendo  $N = 63731$  y  $L = 817090$  la cantidad de nodos y enlaces finales de nuestra red generada, respectivamente (ver Figuras 5.10-(b), (d) y (f)).

### 5.5.2. Mecanismo TC en redes reales

Debido a que tanto la red **HT** como la **FG** cuentan con datos de su evolución temporal con buena resolución, es posible analizar si el mecanismo **TC** tiene o no lugar en dichos casos. Para ello, sea  $\mathbf{l}_t = (i, j)_t$  el enlace que emerge a tiempo  $t$ , uniendo los nodos  $i = \mathbf{l}_t(1)$  y  $j = \mathbf{l}_t(2)$ . Por otra parte, sea  $d(\mathbf{l}_t, t)$  la distancia entre dichos nodos en el instante inmediatamente previo a la aparición de  $\mathbf{l}_t$ . De este modo, aquellos enlaces  $\mathbf{l}_t$  para los cuales  $d(\mathbf{l}_t, t) = 2$ , llamados enlaces transitivos, son el producto de un mecanismo **TC** (cualquiera sea su origen).

Luego, definimos  $N_d(T)$  como el número acumulado de enlaces con  $d(\mathbf{l}_t, t) = d$  para  $t < T$ . Asimismo, definimos  $d(\mathbf{l}_t, t) = 0$  cuando alguno de los nodos  $\mathbf{l}_t(1)$  o  $\mathbf{l}_t(2)$  es nuevo en la red, y  $d(\mathbf{l}_t, t) = \infty$  cuando no existe algún camino de cualquier distancia que una a dichos nodos previo a la concreción de  $\mathbf{l}_t$ . Dado que no se permiten múltiples enlaces entre nodos, consideramos  $N_{d=1}(T) = 0$ . Dicho lo anterior, la distribución de distancias puede definirse formalmente como:

$$P_T(d) = \lim_{L \rightarrow \infty} \frac{N_d(T)}{L} \quad (5.38)$$

siendo  $L = \sum_{i \in \mathbb{N}_0} N_i(T)$  el número total de enlaces.

La distribución de distancias  $P_T(d)$  depende del orden temporal de los enlaces  $\{\mathbf{l}_t\}_{t \leq T}$  y es por este motivo que provee información valiosa acerca del mecanismo de crecimiento de la red. Con el objeto de verificar si es que existe un mecanismo que privilegie la aparición de enlaces transitivos ( $d = 2$ ) para los casos aquí analizados, comparamos la distribución  $P_T(d)$  para el orden temporal real de los enlaces  $\{\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_T\}$ , con la distribución promedio  $\langle P_T(d) \rangle_{rand}$  sobre 100 permutaciones aleatorias. Sea  $\{\mathbf{l}_{\sigma_t}\}_{t=1, \dots, T}$  una permutación aleatoria, donde  $\sigma_t \in \text{Perm}(T)$  pertenece al conjunto de permutaciones de  $T$  índices  $\text{Perm}(T)$ . Entonces, cada  $\{\mathbf{l}_{\sigma_t}\}_{t=1, \dots, T}$  genera exactamente la misma red acumulada a tiempo  $T$  pero con una distribución  $P_T^\sigma(d)$ .

La Figura 5.11 muestra  $P_T(d)$  para las redes **HT** (Figura 5.11-(a)) y **FG** (Figura 5.11-(b)) junto con  $\langle P_T(d) \rangle_{rand}$  en cada caso. En ambos casos se obser-

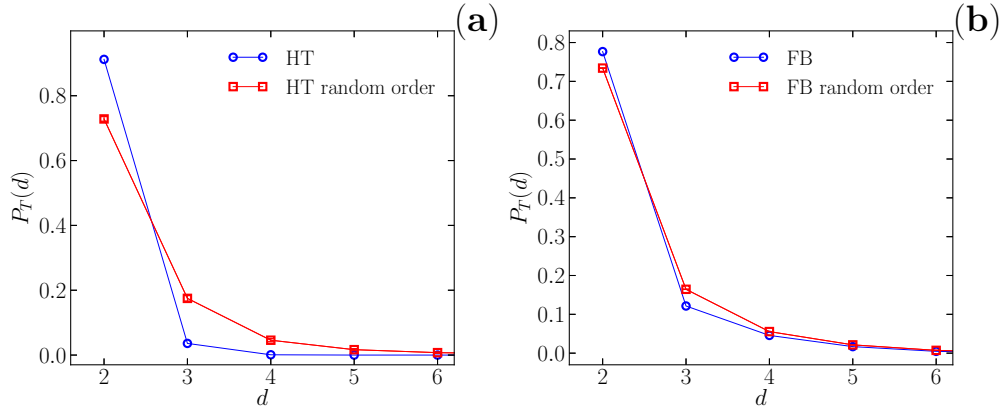


Figura 5.11: Distribución de probabilidad de distancias  $P_T(d)$  (círculos azules) y sus correspondientes valores promedio sobre 100 permutaciones aleatorias de la sucesión real de enlaces ordenada en el tiempo (cuadrados rojos) para: (a) la red de contactos cara-a-cara **HT**, y (b) el subgrafo de amistades online **FG**. Los errores standard resultan más pequeños que los símbolos.

van diferencias significativas entre  $P_T(d)$  y  $\langle P_T(d) \rangle_{rand}$ . En particular,  $P_T(d) < \langle P_T(d) \rangle_{rand}$  para  $d > 2$ , mientras que  $P_T(2) > \langle P_T(2) \rangle_{rand}$ , para ambas redes **HT** y **FG**. La fuerte discrepancia para  $d = 2$  puede cuantificarse a partir del cálculo del  $z$ -score, definido como:

$$z = \frac{P_T(2) - \langle P_T(2) \rangle_{rand}}{\sigma_{rand}}, \quad (5.39)$$

dando por resultado  $z_{HT} = 32$  y  $z_{FG} = 185$ . Por otra parte, los enlaces transitivos resultan claramente predominantes en ambos casos analizados, con  $P_T(2) = 0.91(2)$  para **HT** y  $P_T(2) = 0.77(4)$  para **FG**. Estos hecho dan cuenta de un fuerte efecto de memoria en el mecanismo de crecimiento de la red, que da lugar a una predominio de enlaces **TC** o transitivos.

# Persistencia del sarampión sobre redes sociales

---

## Índice

---

<b>6.1. El modelo</b> . . . . .	<b>98</b>
6.1.1. La red social . . . . .	99
6.1.2. Tasas de transición . . . . .	99
6.1.3. Paso temporal . . . . .	101
6.1.4. Aproximación multinomial . . . . .	102
<b>6.2. Análisis de la persistencia</b> . . . . .	<b>102</b>
6.2.1. Extinciones post-epidémicas con importación permanente de infectados . . . . .	105

---

A mediados del siglo pasado, el matemático inglés Maurice S. Bartlett mostró que, en la era pre-vacuna, el sarampión resultaba endémico oscilatorio en ciudades grandes (con población  $> 250000$  hab.) mientras que presentaba extinciones y rebrotes en ciudades pequeñas. Es así como definió un *tamaño crítico de comuna* (CCS  $250000 - 500000$  hab.) por encima del cual se observaría persistencia del sarampión en poblaciones aisladas [Bartlett 1956, Bartlett 1957]. No obstante, modelos más sofisticados mostraron que el umbral propuesto por Bartlett es fuertemente dependiente de las distribuciones de los tiempos característicos de transición entre los distintos estados [Keeling 1997b, Lloyd 2001].

Por otra parte, también se demostró que el CCS depende de las heterogeneidades en las frecuencias de contactos entre los distintos grupos etarios y de factores estacionales vinculados con el ciclo lectivo [Bolker 1993, Bolker 1995]. Sin embargo, se le ha dedicado muy poca atención al impacto de las heterogeneidades de contactos entre individuos, frecuentemente descritos mediante una red compleja. Entre las escasas excepciones, podemos citar el trabajo de Matt Keeling y colaboradores [Keeling 1997a] acerca del modelo de *aproximación de pares*

(consistente en aproximar funciones de correlación de 3 cuerpos como productos de correlaciones de 2 cuerpos) aplicado a la propagación de una enfermedad infecciosa sobre una red compleja. Por otra parte, también cabe destacar el trabajo de Verdasca y colaboradores [Verdasca 2005], donde los autores estudian los efectos de la propiedad de mundo pequeño en los brotes recurrentes de sarampión. A través de estos enfoques se ha explorado el impacto de la distribución de grados y, a lo sumo, de las correlaciones de dos cuerpos entre los grados de los nodos. Sin embargo, según nuestro leal saber y entender, hasta el momento se desconoce si la existencia de correlaciones locales de orden superior entre los nodos, como las descritas por el coeficiente de clusterización  $\bar{C}$  de la red, tienen algún impacto en el patrón de persistencia del sarampión.

En este capítulo presentamos un modelo SEIR estocástico y basado en el individuo para estudiar la propagación del sarampión en la etapa pre-vacunación, a partir del cual mostramos que la persistencia disminuye al incrementar  $\bar{C}$ . La estructura de correlaciones locales de la red será controlada a partir de la probabilidad de enlaces transitivos  $q$ , introducida en el modelo generador de redes sociales del Capítulo 5.

## 6.1. El modelo

Un modelo compartimental Susceptible-Expuesto-Infectado-Refractario (SEIR) del tipo estocástico resulta el más apropiado para describir la persistencia del sarampión. Utilizamos para ello simulaciones numéricas basadas en el individuo, considerando las redes generadas en el capítulo precedente.

Cada agente de la red puede adoptar uno de los cuatro estados posibles: susceptible, expuesto, infectado o refractario. Un agente susceptible (S) solo puede contraer sarampión de alguno de sus vecinos directos infectados (I), para resultar expuesto (E). Los agentes expuestos no contagian la enfermedad hasta tanto hayan adquirido la carga viral necesaria para ser considerados infectivos (I). Finalmente, los infectivos pasan al estado refractario, donde adquieren inmunidad permanente. Adicionalmente, consideramos una tasa de mortalidad ( $\mu$ ) y otra de natalidad ( $\nu$ ), de modo que ambas resulten iguales, conservando así la población de la red constante. La inclusión de factores demográficos en el problema resulta esencial para renovar la población de susceptibles.

El modelo estocástico basado en el individuo que elegimos para estudiar la persistencia del sarampión es una Cadena de Markov de Tiempo Discreto (DTMC). Esto es, el estado del sistema a tiempo  $t + \delta t$  solo depende del estado a

tiempo  $t$ . La elección más frecuente en los modelos de campo medio es trabajar con una Cadena de Markov de Tiempo Continuo (CTMC). Sin embargo, dicha elección requiere de un tiempo de cómputo prohibitivo para la clase de problema que abordamos aquí, involucrando poblaciones del orden de  $N \sim 10^6$  individuos, cada uno con su tasa de contagio individual dependiente del entorno. Esta implementación daría lugar a tantas tasas de transición como individuos haya en la población. De este modo, la solución exacta de una CTMC a partir del algoritmo de Gillespie [Gillespie 1976] resulta extremadamente lenta para modelos basados en el individuo como el aquí expuesto.

### 6.1.1. La red social

La red social empleada en cada caso es generada mediante el algoritmo del capítulo anterior para población constante. Los nodos y enlaces representan a los individuos y a sus vínculos sociales capaces de transmitir una enfermedad infecciosa, respectivamente. En particular, buscamos generar redes con diferente nivel de correlaciones locales, reguladas por la probabilidad de enlace transitivo  $q$ , con el fin de analizar su impacto en la persistencia del sarampión sobre la red. Trabajamos con redes estáticas de grado medio  $\langle k \rangle$  y poblaciones comprendidas entre  $100000 \leq N \leq 3000000$  individuos.

No estudiaremos heterogeneidades entre los individuos más allá de las provenientes de la red de contactos, con el objeto de circunscribir nuestro análisis al efecto aislado de éstas últimas. En particular, no consideraremos heterogeneidades entre grupos etarios y sus diferentes frecuencias de contacto. Además, trabajaremos con distribución uniforme de actividad para los agentes sociales. Tampoco incorporaremos el factor estacional asociado al incremento de la frecuencia de contactos entre niños pequeños durante el ciclo lectivo.

### 6.1.2. Tasas de transición

Las probabilidades de transición por individuo entre los distintos estados son esencialmente las descritas en el Apéndice A, salvo el agregado de una tasa de mortalidad  $\mu = 4.2 \times 10^{-5}$  compatible con una expectativa de vida de 65 años (considerando como referencia la correspondiente a Inglaterra en la década de 1950 durante las mayores epidemias de sarampión en la era pre-vacunación [<http://www.ons.gov.uk/>]). Por lo explicitado anteriormente, la tasa de nacimiento se considera igual a la de muerte  $\nu = \mu$ . Las distintas probabilidades de transición para un individuo  $i$  durante un intervalo temporal  $\delta t$ , se resumen en la Tabla 6.1.

En particular, la probabilidad de contagio para un nodo/individuo  $i$  susceptible debido a un número  $z(i)$  de vecinos infectados, se define como:

$$P^i(S \rightarrow E) = 1 - (1 - \delta t c / \langle k \rangle)^{z(i)} \quad (6.1)$$

siendo  $c/\langle k \rangle$  la tasa promedio de contagio por enlace para una red con grado medio  $\langle k \rangle$ . El valor del parámetro  $c$  se fija en  $c = 2.8$ , de modo que el número reproductivo básico resulte compatible con el conocido para el sarampión  $12 \leq R_0 \leq 18$  [Anderson 1991].

Cuadro 6.1: Probabilidades de transición por individuos para el modelo epidemiológico.

Transición	Probabilidad
$S \rightarrow E$	$1 - (1 - \delta t c / \langle k \rangle)^{z(i)}$
$E \rightarrow I$	$\sigma \delta t$
$I \rightarrow R$	$\gamma \delta t$
$S, E, I, R \rightarrow$ (muerte)	$\mu \delta t$
$\rightarrow S$ (nacimiento de $S$ )	$\mu \delta t$

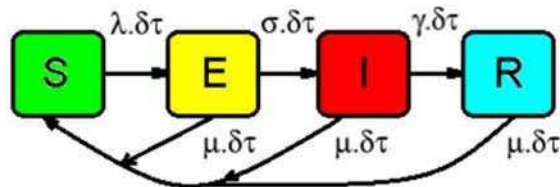


Figura 6.1: Esquema de las probabilidades de transición por individuo para el modelo SEIR.

Cabe destacar que los individuos fallecidos, cualquiera sea su estado, son sustituidos por recién nacidos susceptibles. Asimismo, consideraremos que la probabilidad de muerte o de tener descendencia es independiente del estado epidemiológico de los individuos. El ciclo asociado con nuestro modelo para el sarampión es el representado en la Figura 6.1.

Además de las ya mencionadas tasas de nacimiento y muerte, la restantes tasas empleadas a lo largo de este capítulo se obtienen a partir de los tiempos

de permanencia en el estado expuesto ( $T_E = 8$  días,  $\sigma = 0.125/\text{día}$ ) e infectivo ( $T_I = 5$  días,  $\gamma = 0.2/\text{día}$ ).

Es importante destacar que las distribuciones de tiempos de espera para los estados I y E contempladas en este capítulo serán exponenciales, dada la característica estrictamente Markoviana con tasas constantes del modelo. Hemos destacado la fuerte dependencia entre las distribuciones de tiempo de espera y las propiedades de persistencia en el Capítulo 1 [Keeling 1997b, Keeling 1998, Lloyd 2001, Conlan 2009]. Sin embargo, aquí pretendemos mostrar solo el efecto de las propiedades topológicas de la red, preservando el resto de las variables del problema en su formulación más sencilla.

### 6.1.3. Paso temporal

La elección del paso temporal  $\delta t$  requiere ciertos cuidados. Por un lado, es deseable que durante  $\delta t$  ocurran muchos eventos de distinta especie a fin de disminuir el tiempo de cómputo con respecto al algoritmo de Gillespie (que modela un proceso de Poisson con un único evento por paso). No obstante, si  $\delta t$  es demasiado grande, podrían tener lugar múltiples transiciones para un mismo individuo, no registradas por el esquema propuesto. En cambio, si el paso temporal resultara muy pequeño, se desperdiciaría tiempo de cómputo en pasos durante los cuales no se registraría ningún evento, resultando un algoritmo aún más lento que el de Gillespie. Finalmente, el paso temporal debe satisfacer:

$$\delta t < \min\{T_E, T_I\}, \quad (6.2)$$

a fin de evitar las transiciones múltiples, pero lo suficientemente grande como para comprender muchos eventos. Contemplando dichos aspectos, fijamos  $\delta t = 0.5$  día para todas las simulaciones realizadas.

Dado que la probabilidad de contagio de cada nodo S adquiere un valor particular dependiendo de su entorno, la actualización de los susceptibles debe realizarse a nivel individual. Entonces, en cada paso temporal  $\delta t$ , un nodo S puede contagiarse de los infectados de su entorno y pasar a estado E, puede tener descendencia o morir, ambas con probabilidad  $\mu\delta t$ , o puede continuar en su estado original. En contraste, la actualización sincrónica realizada sobre los nodos E, I y R no requiere tratamiento individual, dado que las tasas de transición entre estos estados son constantes para cada individuo. Es por ello que en estos casos la actualización puede realizarse a nivel de poblaciones con el objeto de acelerar el tiempo de cómputo.



### 6.1.4. Aproximación multinomial

Para ejemplificar el enfoque a nivel de poblaciones, consideremos el caso de los nodos en estado E a tiempo  $t$ . En cada paso temporal, dichos nodos pueden realizar la transición al estado I con una probabilidad  $p_I = \sigma\delta t$ , pueden morir (simultáneamente dando lugar a un recién nacido S) con probabilidad  $p_d = \mu\delta t$ , o pueden permanecer en el estado E con probabilidad  $p_s = 1 - (\sigma + \mu)\delta t$ , todas por individuo. Esto es, los nodos E deben ser repartidos entre 3 estados posibles, cada uno caracterizado por una probabilidad de ocurrencia constante. La forma de concretarlo es a través de la distribución multinomial  $M_{X_1, X_2, X_3}$ , con distribución de probabilidad  $p_M$  dada por:

$$p_M(n_1, n_2, n_3) = \frac{n!}{n_1!n_2!n_3!} p_I^{n_1} p_d^{n_2} p_s^{n_3} \quad (6.3)$$

siendo, para este caso,  $n = n_1 + n_2 + n_3$  el número total de nodos E para el tiempo considerado, y  $\{n_i\}_{i=1,2,3}$  las cantidades de nodos E que resultan infectados, fallecidos e invariables, respectivamente, para dicho paso temporal.

Para implementar la distribución multinomial hicimos uso de su descomposición secuencial en distribuciones binomiales  $B_{X_i}(n, p)$ :

$$M_{X_1, \dots, X_k} = B_{X_1}(n, p_1) \otimes B_{X_2}\left(n - n_1, \frac{p_2}{1 - p_1}\right) \otimes \dots \otimes B_{X_k}\left(n - \sum_{i=1}^{k-1} n_i, \frac{p_k}{1 - \sum_{i=1}^{k-1} p_i}\right). \quad (6.4)$$

Por supuesto, la expresión de la ec. 6.4 no implica en modo alguno que la distribución multinomial pueda descomponerse como producto de binomiales, dado que las variables aleatorias  $X_i$  no son independientes. Este hecho se refleja claramente en la representación secuencial, por la cual la realización de  $X_i$  resulta dependiente de las correspondientes a  $\{X_1, \dots, X_{i-1}\}$ .

## 6.2. Análisis de la persistencia

Nuestro objetivo es analizar la persistencia del sarampión sobre una red social, su dependencia con la población total y, en particular, el impacto de las correlaciones locales de orden superior entre los nodos, usualmente no contempladas en las aproximaciones de campo medio, incluso en aquellas de campo medio heterogéneo que incorporan la estructura de red en su formulación. Para ello, consideramos simulaciones numéricas mediante las redes generadas a partir del modelo **GSG** con población constante  $N \in [10^5, 3 \times 10^6]$ , grado medio

$\langle k \rangle = 40$  y probabilidad de enlaces transitivos  $q$ .

En  $t = 0$  se introducen  $I(0) = 10^{-4} \times N$  infectados elegidos uniformemente al azar de entre la población total  $N$ . El resto de los individuos se consideran susceptibles  $S(0) = N - I(0)$ . Si la población estuviera totalmente aislada, esta fracción inicial de infectados en muchos casos derivaría en una pronta extinción, por lo cual es necesario considerar un transitorio inicial durante el cual se introducen infectados importados. En nuestro caso, siguiendo el trabajo de A. Lloyd [Lloyd 2001], elegimos una tasa constante de infectados importados  $\lambda_{import} = 5/\text{año}$ . La duración del transitorio inicial se fijó en 30 años, período que resulta suficiente para establecer oscilaciones estocásticas recurrentes (ver Figura 6.2).

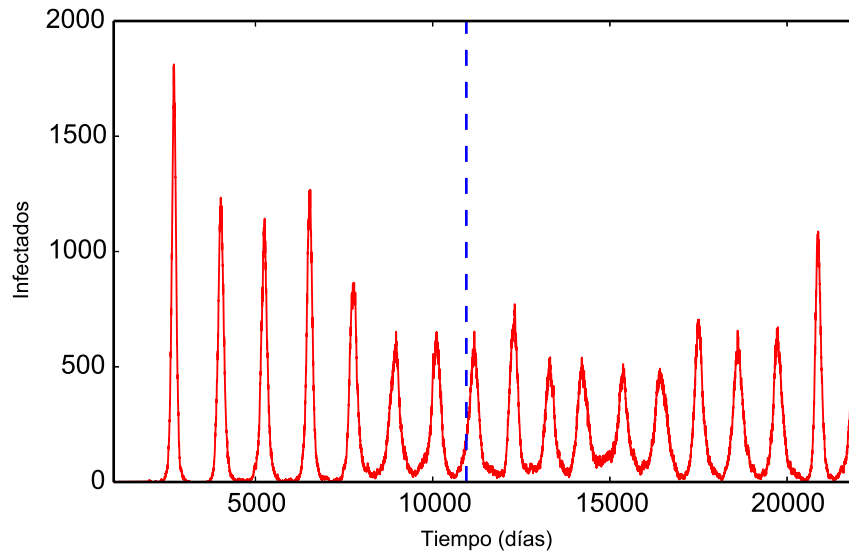


Figura 6.2: Prevalencia del sarampión para una simulación del modelo sobre una red con  $N = 800000$ ,  $\langle k \rangle = 40$  y  $q = 0.7$ . La línea punteada azul indica el tiempo para el cual culmina el transitorio inicial. El ejemplo mostrado corresponde a una realización persistente.

Luego del transitorio, se deja evolucionar la simulación, sin la introducción de infectados externos, durante 30 años adicionales. La Figura 6.2 muestra un ejemplo típico de las oscilaciones estocásticas observadas para la prevalencia del sarampión en las simulaciones numéricas de nuestro modelo. Diremos que el sarampión resulta **persistente** en una realización particular del modelo, **cuando se concluye el segundo período de 30 años sin extinción definitiva del sarampión**.

Para cada población  $N \in [10^5, 3 \times 10^6]$  efectuamos 100 realizaciones del modelo de propagación, Entonces, **definimos la fracción de realizaciones extintas  $P_{extin}$  antes de la culminación del segundo período de 30 años como el indicador de la probabilidad de persistencia del sarampión en la red:**

$$P_{extin} = \frac{\text{realizaciones extintas}}{\text{número total de realizaciones}} \quad (6.5)$$

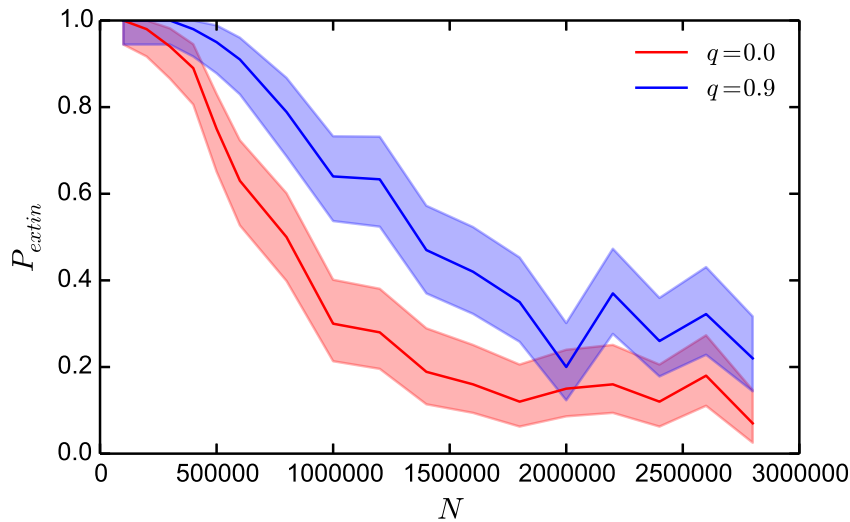


Figura 6.3: Fracción de realizaciones extintas  $P_{extin}$  para el modelo SEIR descrito en el texto principal. Se comparan realizaciones sobre redes con  $q = 0.9$  y  $q = 0.0$ . Los intervalos de confianza (95% CI) considerados son binomiales.

Tal como mencionamos al comienzo del capítulo, es un resultado conocido que para poblaciones lo suficientemente grandes, la tasa de nacimientos y muertes  $\mu$  resulta suficiente para renovar la población de susceptibles, permitiendo así que el sarampión se torne recurrente endémico. Este es precisamente el comportamiento observado en la Figura 6.3, donde presentamos  $P_{extin}$  en función de  $N$ , tanto para redes generadas con  $q = 0.7$  como para redes con  $q = 0.0$ , siendo estas últimas esencialmente redes aleatorias. La Figura 6.3 muestra que  $P_{extin}$  presenta diferencias significativas entre ambos casos analizados. Las redes generadas con  $q = 0.0$  resultan más persistentes ante el sarampión que aquellas con  $q = 0.9$ . Estas últimas presentan fuertes correlaciones de grado inducidas por los enlaces transitivos que, en particular, derivan en un coeficiente de clusterización elevado  $\bar{C}$ , como mostramos en el Capítulo 5.

Observamos en la Figura 6.3 que el CCS (tamaño crítico de comuna) se encuentra en el intervalo  $[5 \times 10^5, 10^6]$  para  $q = 0$ , y en  $[10^6, 2 \times 10^6]$  para  $q = 0.9$ .

Ambos exceden los valores observados para las series temporales de incidencia en el Reino Unido durante las décadas de 1940 y 1950. Esto es esperable dado que en nuestro modelo simplificado no hemos considerado factores estacionales inducidos por el ciclo lectivo, ni heterogeneidades de vínculos entre distintos grupos etarios [Bolker 1995, Keeling 1997b, Keeling 1998], entre otros refinamientos del modelo. De todos modos, el resultado aquí presentado es sumamente interesante debido a que muestra un aspecto desmerecido en los modelos previos: la incidencia en la persistencia de las correlaciones impuestas por la estructura de contactos. El grado medio  $\langle k \rangle$ , la probabilidad de contagio por contacto entre individuos  $S$  e  $I$  y las diferentes tasas de transición entre estados se han mantenido invariantes. Es decir, hemos conservado todos los parámetros que habitualmente se consideran relevantes en los modelos epidemiológicos de campo medio. Esto se justifica dado que muchos de dichos parámetros han sido establecidos independientemente a partir de consideraciones biológicas y no como parámetros de ajuste de un modelo particular, como ocurre para el enfoque metapoblacional con grupos etarios propuesto en [Bolker 1993, Keeling 1997b, Conlan 2009].

La medida de persistencia a través de la  $P_{extin}$  no es la más natural cuando se analizan series temporales reales. El aspecto más crítico es que las poblaciones de las distintas ciudades no se encuentran aisladas de sus vecinas. Es por ello que las series temporales reales no presentan extinciones definitivas sino períodos de ausencia de infectados, que se prolongan hasta que algunos importados ingresan a la ciudad provocando un rebrote. En la siguiente sección definiremos una medida alternativa de persistencia que contempla este hecho.

### 6.2.1. Extinciones post-epidémicas con importación permanente de infectados

Las series temporales para los registros históricos de incidencia de sarampión presentan brotes epidémicos, seguidos de períodos de baja o nula incidencia, dependiendo del tamaño de la población. Para poblaciones grandes, los períodos post-epidémicos presentan fundamentalmente baja, aunque no nula, incidencia. En cambio, en las ciudades más pequeñas, los brotes son seguidos de períodos prolongados con ausencia de infectados en los registros. Entonces, una forma de cuantificar la persistencia del sarampión es a través de la fracción de extinciones posteriores a brotes epidémicos, que denominaremos  $E_{post}$ ,

$$E_{post} = \frac{\text{Extinciones post-epidémicas}}{\text{Total de brotes epidémicos}} \quad (6.6)$$

La definición de  $E_{post}$  necesita de un criterio para establecer cuándo hay un brote epidémico y cuándo una extinción post-epidémica. En nuestro caso, optamos por seguir el criterio utilizado en [Conlan 2009] considerando una extinción post-epidémica cuando transcurren **tres semanas posteriores a un brote sin infectados**.

En lo que respecta a los brotes epidémicos, en [Bartlett 1957] Bartlett optó por considerar un umbral de brote epidémico de 1 nuevo caso de infección por cada 4000 habitantes, para un intervalo temporal de una semana. Sin embargo, para poblaciones pequeñas, este criterio resulta poco restrictivo, al clasificar fluctuaciones estocásticas demográficas como brotes epidémicos. Es por ello que Bartlett tuvo que implementar una modificación ad hoc para poblaciones  $N \lesssim 10^5$ , de tal modo que su inspección visual de las series temporales coincidiera con la clasificación debida al criterio. El objetivo final es que el umbral elegido sea suficiente para que los picos de los reportes semanales de incidencia en las series temporales (reales o simuladas) sean clasificados como brotes epidémicos, y que las pequeñas fluctuaciones sean descartadas. Nuevamente, seguiremos el criterio más sensible impuesto en [Conlan 2009], donde eligieron un umbral de **1 nuevo infectado por cada 2000 habitantes en un período de una semana**.

Es necesario conservar un flujo constante de infectados importados para que  $E_{post}$  pueda definirse. De modo tal que bajo este nuevo enfoque asumiremos  $\lambda_{import} = 5/\text{año}$  constante durante el tiempo total de simulación, en todos los casos. Como antes, en cada simulación se deja transcurrir un transitorio de 30 años y luego se estima  $E_{post}$  sobre los 30 años siguientes. En la Figura 6.4 mostramos  $E_{post}$  en función del tamaño de la población para los mismos parámetros epidemiológicos anteriormente reportados. Nuevamente, comparamos los casos correspondientes a redes generadas a partir del modelo del Capítulo 5 con  $q = 0.0$  y  $q = 0.9$ . La Figura 6.4 evidencia que el caso correspondiente a redes con  $q = 0.0$  presenta menor proporción de extinciones post-epidémicas que el correspondiente a  $q = 0.9$ , en total acuerdo con los resultados de la sección anterior.

Las comparaciones realizadas hasta ahora contrastan los resultados para redes generadas con distinta probabilidad de enlaces transitivos o de cierre triádico  $q$ . Por ello, adjudicamos las diferencias observadas a la existencia de correlaciones de alto orden entre los grados de los nodos. Sin embargo, hemos visto en el capítulo precedente que al modificar  $q$  también se observan diferencias cuantitativas en las distribuciones de grado de las redes obtenidas. Entonces, cabe preguntarnos si las diferencias observadas en la persistencia son realmente debidas al coeficiente de clusterización elevado, o si, en cambio, tienen origen en la modifi-

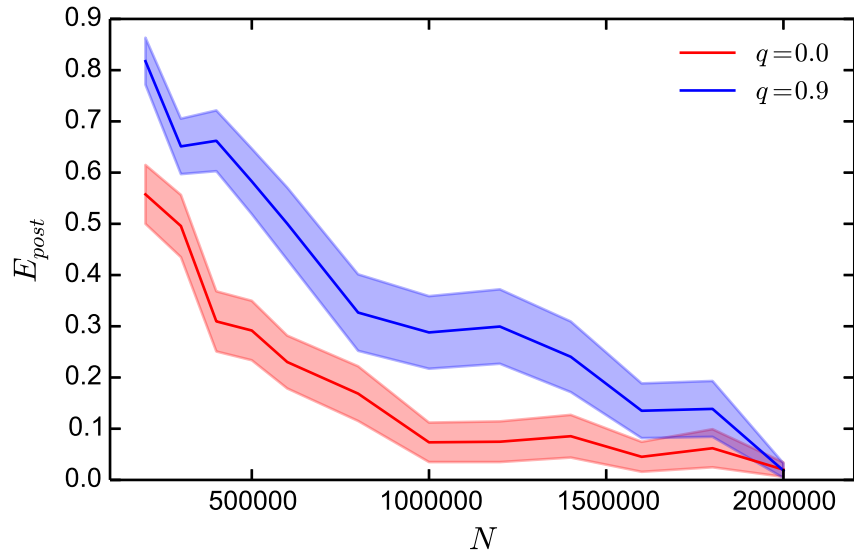


Figura 6.4: Fracción de extinciones post-epidémicas  $E_{post}$  para el modelo descrito en el texto principal. Se comparan realizaciones sobre redes con  $q = 0.9$  y  $q = 0.0$ . Los intervalos de confianza (95 % CI) mostrados corresponden a 1.96 veces el error standard para el promedio sobre 100 simulaciones de 30 años (posteriores al transitorio de otros 30 años de duración).

cación de la distribución de grado. Para dilucidar esta cuestión en la Figura 6.5 comparamos la magnitud de  $E_{post}$  en función de  $N$ , considerando por un lado las redes originales con  $q = 0.9$ , y por el otro, redes aleatorias generadas con idéntica distribución de grados pero construidas mediante el **modelo configuracional** propuesto en [Molloy 1995]. Este último caso corresponde a generar redes auténticamente aleatorias para las cuales los grados de los nodos son exactamente los obtenidos en las redes originales, pero los enlaces son conectados al azar.

En la Figura 6.5 puede observarse que, efectivamente, las discrepancias en cuanto a  $E_{post}$  provienen mayoritariamente del alto coeficiente de clusterización. No obstante, también se observa que la fracción de extinciones post-epidémicas para el modelo configuracional con la distribución de grados correspondiente a  $q = 0.9$ , es aún menor que la obtenida para redes generadas con  $q = 0.0$  en la Figura 6.4. Este hecho podría deberse al efecto de la distribución de grado con decaimiento exponencial más prolongado para  $q = 0.9$  (ver Figura 5.2 del Capítulo 5) y la correspondiente presencia de nodos con grado máximo superior a los obtenidos para  $q = 0.0$ .

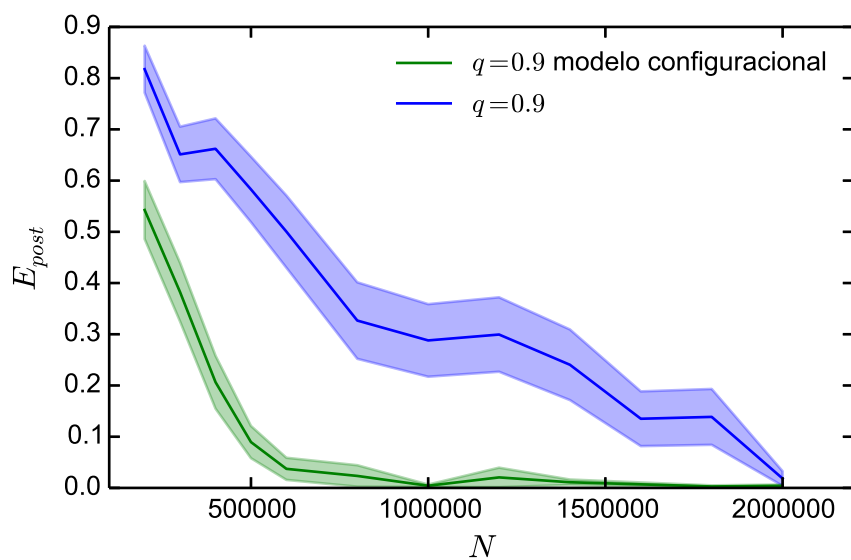


Figura 6.5:  $E_{post}$  como función del tamaño de la población para redes generadas mediante el modelo configuracional. Se compara la fracción de extinciones post-epidémicas conservando la distribución de grado de las redes con  $q = 0.9$  para el modelo introducido en el Capítulo 5. Los intervalos de confianza (95% CI) mostrados corresponden a 1.96 veces el error standard para el promedio sobre 100 simulaciones de 30 años (posteriores al transitorio de otros 30 años de duración).

# Conclusiones Finales y Perspectivas

---

En esta tesis estudiamos dos ejemplos concretos donde el hecho de considerar una estructura de red incide radicalmente en el resultado final del fenómeno de propagación analizado. El primero de ellos corresponde a la propagación de la postura relativa a la vacunación para el caso particular de la vacuna **TV**. En el Capítulo 3 introducimos un modelo de diseminación cultural sobre una red estática sencilla, que conduce a la clusterización de agentes no-vacunadores como resultado de un mecanismo de imitación social mediado por homofilia. A diferencia de [Salathé 2008], nuestro enfoque no introduce una cantidad fija de no-vacunadores en la red, sino que estos surgen naturalmente como consecuencia del mecanismo de formación de opiniones, intervenido por un pequeño grupo de no-vacunadores obstinados (grupo anti-vacunación). Por otra parte, prescindimos de argumentos de teoría de juegos para mostrar que la opinión de vacunación puede propagarse sobre la red y fijarse en dominios o clusters de convergencia cultural por simple afinidad entre sus miembros.

Si bien nuestro enfoque es hipotético, ofrece una explicación plausible para la formación de dominios de no-vacunadores en situaciones de alta cobertura promedio de la vacuna TC, debido a la diseminación de opiniones por el accionar de un pequeño grupo anti-vacunación. El mecanismo de imitación social utilizado ha sido observado a través de experiencias concretas vinculadas con la adopción de comportamientos saludables [Centola 2011, Christakis 2013]. Asimismo, diversos estudios sociales concluyen que las opiniones adversas respecto de la vacunación encuentran un ambiente propicio en comunidades donde prevalece una convergencia cultural reflejada en ciertas orientaciones religiosas, una identificación con tendencias naturistas extremas, o el escepticismo exacerbado en contra de las políticas públicas [May 2003, Jacobson 2007].

En el Capítulo 4 presentamos una extensión del modelo previo, al que denominamos **modelo generalizado**. Consideramos enlaces heterogéneos diferenciados en dos categorías: personales y no-personales. Aunque ambos están involucrados en el mecanismo de diseminación de opiniones, solo los primeros son capaces



de transmitir una enfermedad infecciosa como el sarampión. Aún dentro de este esquema optimista (dado que disminuye la cantidad de enlaces involucrados en la transmisión de la enfermedad) se observan cualitativamente los mismos resultados que con el modelo original. Adicionalmente, incorporamos un mecanismo de adaptación estructural en la red social, que acompaña y complementa el mecanismo de adaptación por imitación. El objetivo es modelar los dos mecanismos identificados como responsables de la clusterización de individuos afines: la imitación social y la elección de amistades por homofilia [McPherson 2001]. La adaptación por reconexión de enlaces por homofilia da lugar a una transición desde el consenso no-vacunador hacia el vacunador menos abrupta que en el caso anterior, con una región crítica más extensa. Por otra parte, las características topológicas de los clusters máximos obtenidos se diferencian con respecto a los de idéntico tamaño correspondientes a red estática. Tales diferencias topológicas se plasman en las características de propagación, como se evidencia en la Figura 4.5.

Otra alternativa del modelo generalizado corresponde a la inclusión de campañas de información en favor de la vacunación. Las campañas masivas se introducen mediante un campo externo  $\phi$  que actúa a modo de sesgo contra la transición de vacunador hacia no-vacunador. La característica satisfactoria de dichas campañas es que no promueven la vacunación compulsiva, ni siquiera buscan convencer a los no-vacunadores. Muy por el contrario, las campañas aquí modeladas proponen un enfoque conservador que solo busca evitar que los vacunadores cambien su opinión. Mostramos el efector fuertemente mitigador de  $\phi$  tanto en la versión estática como en la adaptativa de nuestro modelo generalizado. Sin embargo, el efecto de  $\phi$  resulta más intenso sobre los clusters del modelo adaptativo debido a su mayor fragilidad estructural. Nuevamente se evidencia la importancia del estudio de las características topológicas del grafo, en este caso, de no-vacunadores.

Probablemente la conclusión más importante del modelo de los Capítulos 3 y 4 sea el impacto de la clusterización de no-vacunados sobre la red, inducida por el mecanismo de imitación social mediado por homofilia, en la inmunidad de grupo. Los modelos de campo medio predicen que una cobertura de la vacuna **TC** del 95 % sería suficiente para prevenir brotes de sarampión, protegiendo así a los que no han sido vacunados por razones de fuerza mayor. Nosotros mostramos, a través de simulaciones numéricas, que este umbral podría ser insuficiente si se considera la estructura de contactos sociales en conjunto con un proceso de diseminación de opiniones. No obstante, aunque los efectos de la clusterización de no-vacunados predichos por nuestro modelo han sido verificados en numero-

sas oportunidades [May 2003, Salathé 2011], aún restan experiencias controladas que ofrezcan evidencia científica, tales como las recientemente realizadas para el caso de la gripe [Barclay 2014]. Este conocimiento concreto permitiría establecer nuevas políticas de vacunación y campañas de prevención focalizadas en la erradicación definitiva del sarampión. Los resultados presentados en los Capítulos 3 y 4 han sido enviados para su publicación [Medus 2014b].

Como trabajo futuro, el modelo de diseminación de opiniones propuesto en esta tesis aún requiere de un análisis de sensibilidad más exhaustivo respecto de los diferentes parámetros involucrados. Por otra parte, así como se han incorporado grupos anti-vacunación, también resultaría interesante explorar el efecto de promotores de la vacunación que gocen de prestigio, confianza y autoridad frente a la sociedad, es decir, los médicos. Otro aspecto que proyectamos abordar en el futuro inmediato es el efecto de una epidemia inminente de sarampión como un factor que induce a la vacunación por temor [Fu 2011, Bauch 2012].

Hemos visto que los aspectos topológicos de las redes sociales resultan relevantes en problemas de diseminación de opiniones y enfermedades infecciosas. Es por ello que diversos trabajos se han abocado al estudio específico tanto de sus propiedades estáticas como dinámicas. Las propiedades topológicas estáticas más observadas en las redes de origen social son el alto coeficiente de clusterización promedio  $\bar{C}$ , comparable con el de algunas redes regulares, y distancia media entre nodos  $\bar{l}$  del orden de la correspondiente a redes aleatorias ( $\bar{l} \sim \ln N$ ). Sin embargo, recientes trabajos dedicados a la dinámica microscópica de los enlaces sociales, han propuesto modelos de redes variables en el tiempo, dando lugar a redes acumuladas que no contemplan estas propiedades estáticas antes mencionadas [Perra 2012, Starnini 2013a, Starnini 2013b]. Si bien estos capturan correctamente las propiedades dinámicas de los enlaces (frecuencia, duración, etc.), no satisfacen las propiedades estáticas de las redes acumuladas en el tiempo que han sido observadas y modeladas tanto en la actualidad como en el pasado reciente.

En el Capítulo 5 propusimos un modelo de crecimiento de redes que incorpora la dinámica de formación de enlaces a través del concepto de **actividad social** [Perra 2012], de tal modo que las redes acumuladas en el tiempo generadas satisfagan las propiedades topológicas usuales. Para ello incorporamos un aspecto hasta el momento soslayado en los modelos de redes variables en el tiempo: el carácter no-markoviano de los agentes sociales. Esto es, consideramos efectos de memoria de largo plazo en los agentes. En contraposición, los modelos que estudian la dinámica de los enlaces sociales frecuentemente asumen agentes markovianos que se comportan como caminantes aleatorios, estableciendo vínculos

con otros elegidos al azar [Starnini 2013a].

Mostramos que el mecanismo de cierre triádico (TC) del modelo descrito en el Capítulo 5, logra introducir las correlaciones de alto orden usualmente observadas en las redes de origen social acumuladas en el tiempo. Propusimos dos variantes del modelo para población de nodos constante y creciente. Describimos analíticamente las distribuciones de grado en ambos casos y estudiamos el comportamiento asintótico asociado. A su vez, mostramos que las características topológicas de las redes generadas mediante nuestro modelo, se corresponden con las observadas en dos casos de estudio: un subgrafo de la red social Facebook [Viswanath 2009] y una red de contactos cara-a-cara obtenida del proyecto SocioPatterns [Cattuto 2010, Isella 2011]. Aunque el modelo no está centrado en ninguna dinámica particular de formación y borrado de enlaces, da lugar a un marco teórico dentro del cuál ésta podría introducirse.

Adicionalmente, demostramos que el mecanismo de crecimiento de redes aquí introducido contiene al *modelo de variables ocultas* de [Boguna 2003] como caso particular. Los resultados del Capítulo 5 fueron publicados en [Medus 2014a].

Finalmente, en el Capítulo 6 utilizamos las redes generadas en el capítulo precedente para estudiar la incidencia de la transitividad inducida por el mecanismo de cierre triádico en la persistencia del sarampión sobre la red. Desarrollamos un algoritmo estocástico de tiempo discreto, basado en el individuo y del tipo SEIR para estudiar la propagación del sarampión. Como parte de este algoritmo original, utilizamos la aproximación multinomial a nivel de poblaciones con el fin de acelerar el tiempo de cómputo. Encontramos que la persistencia disminuye al incrementar la probabilidad de cierre triádico  $q$ . Este hecho se verifica para los dos criterios de persistencia introducidos.

Los resultados del Capítulo 6 muestran que el conocimiento específico de la topología de una red social puede tener un impacto notable en el estudio de la persistencia del sarampión, así como también de otras enfermedades infecciosas transmitidas por contacto. Estos resultados serán enviados para su publicación a la brevedad. Como trabajo a futuro, resta considerar un modelo más detallado que incorpore una dinámica de surgimiento y borrado tanto de enlaces como de nodos. Por otra parte, el modelo también podría incorporar cohortes enlazadas preferencialmente por homofilia, esto es, la probabilidad de enlace entre dos individuos de la misma cohorte podría resultar superior a la correspondiente para otros dos pertenecientes a cohortes distintas. Asimismo, podemos introducir distribuciones de tiempos de espera más realista del tipo Gamma y estudiar su importancia relativa con respecto al impacto de los efectos de la topología de la

red. Hemos comenzado a trabajar sobre algunas de estas líneas.



# Modelo epidemiológico

---

El sarampión es una enfermedad viral infecciosa causada por un paramixovirus, siendo el ser humano su único huésped conocido. Se trata de una enfermedad altamente contagiosa, fundamentalmente transmitida de persona a persona por vía aérea a través de las microgotas de Flügge emitidas al hablar o estornudar. Afecta principalmente a niños pequeños en edad de primera escolaridad y es una de las principales causas de mortalidad infantil <sup>1</sup>. La vacuna TV, administrada globalmente desde fines de la década de 1980, produce inmunidad contra el sarampión, las paperas y la rubeola.

## A.1. Modelo compartimental estocástico

A lo largo de la presente tesis utilizamos un modelo epidemiológico estocástico compartimental del tipo SEIR (Susceptible-Expuesto-Infectado-Refractario) para el sarampión. Solo aquellos individuos infectados son capaces de transmitir la enfermedad, mientras que los individuos recuperados adquieren inmunidad permanente (estado refractario). Consideramos una estructura de contactos entre agentes descrita por medio de una red compleja  $G(N, L)$ , compuesta de  $N$  nodos y  $L$  enlaces entre los mismos. Los efectos demográficos se consideran despreciables.

Las probabilidades de transición entre estados por individuo y para un intervalo temporal  $\delta t$  se definen en la Tabla A.1. Merece especial atención la transición de estado susceptible a refractario ( $P(S \rightarrow E)$ ). Si llamamos  $z(i)$  a la cantidad de vecinos infectados para el agente  $i$  de la red y  $\langle k_p \rangle$  al grado medio de enlaces personales por agente, entonces la probabilidad de contagio del agente  $i$  debida a un contacto en el intervalo temporal  $\delta t$  con uno de sus vecinos infectados, viene dada por  $\delta t c / \langle k_p \rangle$  (donde  $c / \langle k_p \rangle$  corresponde a la tasa de contagio por enlace). Luego, la probabilidad de que el agente  $i$  contraiga la enfermedad en  $\delta t$  producto

---

<sup>1</sup><http://www.who.int/mediacentre/factsheets/fs286/en/>

de sus vecinos infecciosos es:

$$P^i(S \rightarrow E) = 1 - (1 - \delta t c / \langle k_p \rangle)^{z(i)}. \quad (\text{A.1})$$

Cuadro A.1: Probabilidades de transición por individuo para el modelo SEIR utilizado.

Transición	Probabilidad
$S \rightarrow E$	$1 - (1 - \delta t c / \langle k_p \rangle)^{z(i)}$
$E \rightarrow I$	$\sigma \delta t$
$I \rightarrow R$	$\gamma \delta t$

En este trabajo consideramos  $c = 2.8$ , valor compatible con el número reproductivo básico del sarampión ( $12 \leq R_0 \leq 18$ ) para los modelos deterministas [Anderson 1991]. Por otra parte, consideramos que el tiempo característico de la transición de expuesto a infectado es  $T_E = 8$  días y el de infectado a refractario  $T_I = 5$  días [Lloyd 2001, Conlan 2009]. De este modo, las correspondientes tasas de transición se definen como  $\sigma = 1/T_E = 0.125$  y  $\gamma = 1/T_I = 0.2$ .

Definimos como  $S_t$ ,  $E_t$ ,  $I_t$  y  $R_t$  las poblaciones de agentes susceptibles, expuestos, infectados y refractarios a tiempo  $t$ , respectivamente. El modelo SEIR estocástico con red compleja propuesto representa un proceso de Markov de paso discreto  $\delta t$ . Luego, el algoritmo actualiza sincrónicamente las poblaciones de cada estado para un paso temporal  $\delta t = 1/2$  día, del siguiente modo:

1. Cada agente susceptible  $i \in S_t$  se vuelve Expuesto con la probabilidad dada en la ec. A.1, solo considerando aquellos agentes infectados  $j$  pertenecientes a su entorno ( $j \in I_t$ ). La población de susceptibles se actualiza para el siguiente paso temporal:  $S_{t+\delta t}$ .
2. Cada agente expuesto  $i \in E_t$  se vuelve Infectado con probabilidad  $\sigma \delta t$ . La población de agentes expuestos se actualiza para el siguiente paso temporal:  $E_{t+\delta t}$ .
3. Cada agente infectado  $i \in I_t$  se recupera con probabilidad  $\gamma \delta t$ . La población de agentes infectados se actualiza para el siguiente paso temporal:  $I_{t+\delta t}$ .
4. Se actualiza el tiempo  $t = t + \delta t$ , retornando al paso 1 hasta que  $E_t + I_t = 0$ .

Notamos que el algoritmo anterior no actualiza la población de refractarios dado que al mantenerse la población total  $N$  constante, esta puede derivarse a partir de las restantes.





# Breve introducción al modelo de variables ocultas

---

## B.1. Modelo de variables ocultas

La clase de modelos con variables ocultas consideran una propiedad intrínseca de los nodos que determina completamente las propiedades de la red [Boguna 2003]. El modelo comienza con  $N$  nodos desconectados, cada uno asociado con una variable oculta  $h$ , obtenida mediante una distribución de probabilidad  $\rho(h)$ . Posteriormente, los enlaces se establecen a partir de una probabilidad de conexión simétrica  $r(h_i, h_j)$ . De este modo, tanto  $\rho(h)$  como  $r(h_i, h_j)$  determinan completamente las propiedades topológicas de una *red aleatoria Markoviana*. Dichas propiedades topológicas son obtenidas en función de la variable oculta  $h$  y, luego, un propagador  $g(k|h)$  provee la transformación desde  $h$  hacia el grado usual  $k$ . El propagador  $g(k|h)$  representa la probabilidad condicional de que un nodo con variable oculta  $h$  resulte finalmente con grado  $k$ . De esta forma, la distribución de grado puede ser escrita como:

$$P(k) = \sum_h g(k|h)\rho(h). \quad (\text{B.1})$$

De igual modo es también posible obtener el grado medio de los vecinos inmediatos de un nodo con grado  $k$  como [Boguna 2003]

$$\bar{k}_{nn}(k) = 1 + \frac{1}{P(k)} \sum_h g(k|h)\rho(h)\bar{k}_{nn}(h), \quad (\text{B.2})$$

donde  $\bar{k}_{nn}(h)$  se define como

$$\bar{k}_{nn}(h) = \frac{N}{\bar{k}(h)} \sum_{h'} \rho(h')\bar{k}(h')r(h, h'), \quad (\text{B.3})$$

con  $\bar{k}(h')$  el grado medio de un  $h'$ -nodo, dado por

$$\bar{k}(h) = N \sum_{h'} \rho(h') r(h, h'). \quad (\text{B.4})$$

Adicionalmente, el coeficiente de clusterización medio como función del grado  $k$  puede ser escrito como

$$\bar{C}(k) = \frac{1}{P(k)} \sum_h \rho(h) g(k|h) \bar{C}(h) \quad (\text{B.5})$$

donde  $\bar{C}(h)$  se define

$$\bar{C}(h) = \sum_{h', h''} P(h'|h) r(h', h'') P(h''|h), \quad (\text{B.6})$$

con  $P(h'|h)$  la probabilidad condicional de conexión entre un  $h$ -nodo y un  $h'$ -nodo, dada por:

$$P(h'|h) = \frac{N \rho(h') r(h, h')}{\bar{k}(h)}. \quad (\text{B.7})$$

En el modelo **GSG** introducido en el Capítulo 5, la actividad  $a$  juega el rol de variable oculta, con la pdf de actividad  $F(a)$  tomando el lugar de la distribución  $\rho(h)$  en una formulación continua. Sin embargo, las redes agregadas en el tiempo obtenidas a través del modelo **GSG** resultan no-Markovianas debido a las correlaciones locales inducidas por el mecanismo de cierre triádico **TC** (excepto para el caso con  $q = 0$ ). Asimismo, en el Capítulo 5 hemos obtenido  $\bar{N}_{k|a}$  como un equivalente del propagador correspondiente al modelo de variables ocultas para cualquier  $q \geq 0$ . Sin embargo, no nos fue posible derivar las correlaciones de grado de alto orden solo a partir de conocida la probabilidad de conexión  $r(a, a')$  como ocurre para el modelo de variables ocultas. Esto es así debido a que el mecanismo de conexión **TC** introduce correlaciones de grado más allá de las impuestas por el mecanismo inducido por la actividad social de los agentes. No obstante, pudimos describir todas las propiedades topológicas de las redes una vez conocidas  $F(a)$  y  $r(a, a')$  para el caso particular con  $q = 0$ .

Por último, daremos los lineamientos generales para una derivación de la función de escala de la ec. 5.33. En este caso particular tenemos  $q = 0$  y, por esta razón, podemos obtener una expresión aproximada para la probabilidad  $r_\ell(i, j)$  de que dos nodos  $i$  y  $j$  con actividades  $a_i$  y  $a_j$ , respectivamente, resulten conectados luego de  $\ell$  enlaces agregados. Sea  $s_\ell(i, j) = 1 - r_\ell(i, j)$  la probabilidad

de que no se establezca una conexión entre aquellos dos nodos cuando  $\ell$  enlaces fueron agregados. Así, podemos escribir  $s_\ell(i, j)$  como [Starnini 2013b]:

$$s_\ell(i, j) = \sum_{z_i, z_j} P_\ell(z_i) P_\ell(z_j) \left[ \prod_{n=1}^{z_i} \left( 1 - \frac{1}{N-n} \right) \right] \left[ \prod_{m=1}^{z_j} \left( 1 - \frac{1}{N-m} \right) \right], \quad (\text{B.8})$$

con  $P_\ell(z_i)$  ( $P_\ell(z_j)$ ) la probabilidad de que  $z_i$  ( $z_j$ ) enlaces emerjan de  $i$  ( $j$ ) luego de  $\ell$  enlaces agregados, dada por:

$$P_\ell(z_i) = \binom{\ell}{z_i} \left( \frac{a_i}{N\langle a \rangle} \right)^{z_i} \left( 1 - \frac{a_i}{N\langle a \rangle} \right)^{\ell - z_i}, \quad (\text{B.9})$$

siendo  $a_i/(N\langle a \rangle)$  la probabilidad de elegir el nodo  $i$  con actividad  $a_i$  de entre toda la población. Considerando  $z_i, z_j \ll N$  en la ec B.8, tenemos entonces

$$\prod_{n=1}^{z_i, j} \left( 1 - \frac{1}{N-n} \right) \simeq \left( 1 - \frac{1}{N} \right)^{z_i, j}. \quad (\text{B.10})$$

Luego, substituyendo la ec. B.9 en B.8, junto con la aproximación de la ec. B.10, obtenemos  $r_\ell(a_i, a_j)$  con

$$r_\ell(a_i, a_j) = 1 - s_\ell(a_i, a_j) \simeq 1 - \left( 1 - \frac{a_i}{N^2\langle a \rangle} \right)^\ell \left( 1 - \frac{a_j}{N^2\langle a \rangle} \right)^\ell, \quad (\text{B.11})$$

que puede ser aproximada por

$$r_\ell(a_i, a_j) \simeq 1 - \exp \left( -\frac{\ell}{N^2\langle a \rangle} (a_i + a_j) \right). \quad (\text{B.12})$$

A través de la correspondencia  $\rho(h) \rightarrow F(a)$ ,  $r(h, h') \rightarrow r_\ell(a, a')$ ,  $g(k|h) \rightarrow \bar{N}_{k|a}$  y reemplazando la expresión de la ec. B.12 en las ecs. B.3 y B.4, arribamos a la función de escala de la ecuación 5.33 siguiendo el procedimiento descrito en [Starnini 2013b].



# Bibliografía

- [Albert 1999] R. Albert, H. Jeong and A.L. Barabasi. *The Diameter of the World Wide Web*. Nature, vol. 401, pages 130–131, 1999. (ver pág 22.)
- [Albert 2002] R. Albert and A.-L. Barabási. *Statistical mechanics of complex networks*. Rev. Mod. Phys., vol. 74, no. 1, pages 47–97, January 2002. (ver pág 21.)
- [Anderson 1991] R. M. Anderson and R. M. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, USA, 1991. (ver pág 12, 100 y 116.)
- [Andre 2008] F E Andre, R Booy, H L Bock, J Clemens, S K Datta, T J John, B W Lee, S Lolekha, H Peltola, T A Ruff, M Santosham and H J Schmitt. *Vaccination greatly reduces disease, disability, death and inequity worldwide*. Bull. World Health Organ., vol. 86, no. 2, pages 140–6, February 2008. (ver pág 2.)
- [Axelrod 1997] R. Axelrod. *The Dissemination of Culture: A Model with Local Convergence and Global Polarization*. J. Conflict Resolut., vol. 41, no. 2, pages 203–226, 1997. (ver pág 7, 33, 36 y 50.)
- [Bailey 1954] Norman TJ Bailey. *A statistical method of estimating the periods of incubation and infection of an infectious disease*. 1954. (ver pág 12.)
- [Bailey 1956] Norman TJ Bailey. *On estimating the latent and infectious periods of measles: I. Families with two susceptibles only*. Biometrika, pages 15–22, 1956. (ver pág 12.)
- [Bailey 1964] Norman TJ Bailey. *Some stochastic models for small epidemics in large populations*. Applied Statistics, pages 9–19, 1964. (ver pág 17.)
- [Bailey 1975] Norman TJ Bailey *et al.* *The mathematical theory of infectious diseases and its applications*. Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE., 1975. (ver pág 16 y 17.)
- [Barabasi 1999a] A. L. Barabasi and R. Albert. *Emergence of scaling in random networks*. Science, vol. 286, pages 509–512, 1999. (ver pág 21.)

- [Barabasi 1999b] A.L. Barabasi and R. Albert. *Emergence of scaling in random networks*. Science, vol. 286, no. 5439, page 509, 1999. (ver pág 10, 31, 74, 83 y 85.)
- [Barabási 2001] Albert L. Barabási. *The physics of the Web*. PhysicsWeb.ORG, online journal, July 2001. (ver pág 20.)
- [Barclay 2014] Victoria C Barclay, Timo Smieszek, Jianping He, Guohong Cao, Jeanette J Rainey, Hongjiang Gao, Amra Uzicanin and Marcel Salathé. *Positive network assortativity of influenza vaccination at a high school: implications for outbreak risk and herd immunity*. PLoS ONE, vol. 9, no. 2, page e87042, 2014. (ver pág 111.)
- [Barrat 2004] A. Barrat, M. Barthélemy, R. Pastor-Satorras and A. Vespignani. *The Architecture of Complex Weighted Networks*. Proceedings of the National Academy of Sciences, vol. 101, no. 11, pages 3747–3752, 2004. (ver pág 25.)
- [Bartlett 1956] MS Bartlett. *Deterministic and stochastic models for recurrent epidemics*. In Proceedings of the third Berkeley symposium on mathematical statistics and probability, volume 4, page 109. University of California Press Berkeley, 1956. (ver pág 97.)
- [Bartlett 1957] Maurice S Bartlett. *Measles periodicity and community size*. Journal of the Royal Statistical Society. Series A (General), pages 48–70, 1957. (ver pág 14, 15, 97 y 106.)
- [Bauch 2003] Chris T Bauch, Alison P Galvani and David JD Earn. *Group interest versus self-interest in smallpox vaccination policy*. Proceedings of the National Academy of Sciences, vol. 100, no. 18, pages 10564–10567, 2003. (ver pág 4.)
- [Bauch 2004] Chris T Bauch and David JD Earn. *Vaccination and the theory of games*. Proceedings of the National Academy of Sciences of the United States of America, vol. 101, no. 36, pages 13391–13394, 2004. (ver pág 4.)
- [Bauch 2005] Chris T Bauch. *Imitation dynamics predict vaccinating behaviour*. Proc. Biol. Sci., vol. 272, no. 1573, pages 1669–75, August 2005. (ver pág 4.)

- [Bauch 2012] Chris T Bauch and Samit Bhattacharyya. *Evolutionary game theory and social learning can determine how vaccine scares unfold*. PLoS computational biology, vol. 8, no. 4, page e1002452, 2012. (ver pág 111.)
- [Bearman 2004] Peter S. Bearman, James Moody and Katherine Stove. *Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks*. American Journal of Sociology, vol. 110, no. 1, pages 44–91, 2004. (ver pág 20.)
- [Biss 2014] Eula Biss. *On Immunity: An Inoculation*. Macmillan, 2014. (ver pág 9.)
- [Blondel 2008] V.D. Blondel, J.L. Guillaume, R. Lambiotte and E.L.J.S. Meck. *Fast unfolding of communities in large networks*. J. Stat. Mech, page P10008, 2008. (ver pág 70.)
- [Boguna 2003] Marian Boguna and Romualdo Pastor-Satorras. *Class of correlated random networks with hidden variables*. Physical Review E, vol. 68, page 036112, 2003. (ver pág 79, 112 y 119.)
- [Boguñá 2004] Marián Boguñá, Romualdo Pastor-Satorras, Albert Díaz-Guilera and Alex Arenas. *Models of social networks based on social distance attachment*. Phys. Rev. E, vol. 70, no. 5 Pt 2, page 056122, November 2004. (ver pág 23.)
- [Bolker 1993] B M Bolker and B T Grenfell. *Chaos and biological complexity in measles dynamics*. Proc. Biol. Sci., vol. 251, no. 1330, pages 75–81, January 1993. (ver pág 13, 15, 16, 97 y 105.)
- [Bolker 1995] Benjamin Bolker and Bryan Grenfell. *Space, persistence and dynamics of measles epidemics*. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, vol. 348, no. 1325, pages 309–320, 1995. (ver pág 16, 17, 97 y 105.)
- [Bond 2012] Robert M Bond, Christopher J Fariss, Jason J Jones, Adam D I Kramer, Cameron Marlow, Jaime E Settle and James H Fowler. *A 61-million-person experiment in social influence and political mobilization*. Nature, vol. 489, no. 7415, pages 295–8, September 2012. (ver pág 11 y 52.)



- [Brown 2012] Katrina F Brown, Susannah J Long, Mary Ramsay, Michael J Hudson, John Green, Charles A Vincent, J Simon Kroll, Graham Fraser and Nick Sevdalis. *UK parents' decision-making about measles–mumps–rubella (MMR) vaccine 10 years after the MMR-autism controversy: A qualitative analysis*. *Vaccine*, vol. 30, no. 10, pages 1855–1864, 2012. (ver pág 3.)
- [Burgess 2006] David C Burgess, Margaret A Burgess and Julie Leask. *The MMR vaccination and autism controversy in United Kingdom 1998–2005: Inevitable community outrage or a failure of risk communication?* *Vaccine*, vol. 24, no. 18, pages 3921–3928, 2006. (ver pág 3.)
- [Cardillo 2013] Alessio Cardillo, Catalina Reyes-Suárez, Fernando Naranjo and Jesús Gómez-Gardeñes. *The Evolutionary Vaccination Dilemma in Complex Networks*. *CoRR*, vol. abs/1303.3502, 2013. (ver pág 5, 35 y 36.)
- [Castellano 2000] C Castellano, M Marsili and A Vespignani. *Nonequilibrium Phase Transition in a Model for Social Influence*. *Phys Rev Lett*, vol. 85, no. 16, pages 3536–3539, 2000. (ver pág 8.)
- [Cattuto 2010] Ciro Cattuto, Wouter Van den Broeck, Alain Barrat, Vittoria Colizza, Jean-François Pinton and Alessandro Vespignani. *Dynamics of person-to-person interactions from distributed RFID sensor networks*. *PloS one*, vol. 5, no. 7, page e11596, 2010. (ver pág 2, 11, 92 y 112.)
- [Centola 2010] Damon Centola. *The Spread of Behavior in an Online Social Network Experiment*. *Science*, vol. 329, no. 5996, pages 1194–1197, 2010. (ver pág 2.)
- [Centola 2011] Damon Centola. *An experimental study of homophily in the adoption of health behavior*. *Science*, vol. 334, no. 6060, pages 1269–72, December 2011. (ver pág 52 y 109.)
- [Cho 2011] Eunjoon Cho, Seth A Myers and Jure Leskovec. *Friendship and mobility: user movement in location-based social networks*. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1082–1090. ACM, 2011. (ver pág 2.)
- [Christakis 2013] Nicholas A Christakis and James H Fowler. *Social contagion theory: examining dynamic social networks and human behavior*. *Statistics in medicine*, vol. 32, no. 4, pages 556–577, 2013. (ver pág 109.)

- [Cohen 2003] R. Cohen and S. Havlin. *Scale-Free Networks Are Ultrasmall*. Phys. Rev. Lett., vol. 90, page 058701, 2003. (ver pág 22.)
- [Colomer-de Simon 2012] Pol Colomer-de Simon and Marián Boguñá. *Clustering of random scale-free networks*. Phys. Rev. E, vol. 86, no. 2 Pt 2, page 026120, August 2012. (ver pág 23.)
- [Conlan 2009] Andrew JK Conlan, Pejman Rohani, Alun L Lloyd, Matthew Keeling and Bryan T Grenfell. *Resolving the impact of waiting time distributions on the persistence of measles*. Journal of The Royal Society Interface, page rsif20090284, 2009. (ver pág 12, 17, 101, 105, 106 y 116.)
- [Conover 2011] Michael Conover, Jacob Ratkiewicz, Matthew Francisco, Bruno Gonçalves, Filippo Menczer and Alessandro Flammini. *Political polarization on twitter*. In ICWSM, 2011. (ver pág 11 y 52.)
- [Crandall 2010] David J. Crandall, Lars Backstrom, Dan Cosley, Siddharth Suri, Daniel Huttenlocher and Jon Kleinberg. *Inferring social ties from geographic coincidences*. Proceedings of the National Academy of Sciences, vol. 107, no. 52, pages 22436–22441, 2010. (ver pág 2.)
- [Davidsen 2002] J Davidsen, H Ebel and S Bornholdt. *Emergence of a Small World from Local Interactions: Modeling Acquaintance Networks*. Phys Rev Lett, vol. 88, no. 12, page 128701, 2002. (ver pág 31 y 78.)
- [Domínguez 2008] Angela Domínguez, Nuria Torner, Irene Barrabeig, Ariadna Rovira, Cristina Rius, Joan Cayla, Elsa Plasencia, Sofia Minguell, M. Rosa Sala, Ana Martínez, Josep Costa, Mar Mosquera and Carmen Cabezas. *Large Outbreak of Measles in a Community with High Vaccination Coverage: Implications for the Vaccination Schedule*. Clinical Infectious Diseases, vol. 47, no. 9, pages 1143–1149, November 2008. (ver pág 4.)
- [d’Onofrio 2011] Alberto d’Onofrio, Piero Manfredi and Piero Poletti. *The impact of vaccine side effects on the natural history of immunization programmes: an imitation-game approach*. J. Theor. Biol., vol. 273, no. 1, pages 63–71, March 2011. (ver pág 4.)
- [d’Onofrio 2012] Alberto d’Onofrio, Manfredi Piero and Poletti Piero. *The Interplay of Public Intervention and Private Choices in Determining the Outcome of Vaccination Programmes*. PLoS ONE, vol. 7, no. 10, page e45653, 10 2012. (ver pág 66.)

- [Dorogovtsev 2004] S N Dorogovtsev. *Clustering of correlated networks*. Phys Rev E Stat Nonlin Soft Matter Phys, vol. 69, no. 2 Pt 2, page 027104, February 2004. (ver pág 90.)
- [Dorso 2010] Claudio Oscar Dorso and A. D. Medus. *Community Detection in Networks*. I. J. Bifurcation and Chaos, vol. 20, no. 2, pages 361–367, 2010. (ver pág 28.)
- [Eagle 2006] Nathan Eagle and Alex Pentland. *Reality mining: sensing complex social systems*. Personal and Ubiquitous Computing, vol. 10, no. 4, pages 255–268, 2006. (ver pág 2 y 11.)
- [Eagle 2009] Nathan Eagle, Alex (Sandy) Pentland and David Lazer. *Inferring friendship network structure by using mobile phone data*. Proceedings of the National Academy of Sciences, vol. 106, no. 36, pages 15274–15278, 2009. (ver pág 2 y 11.)
- [Erdős 1959] Paul Erdős and Alfréd Rényi. *On Random Graphs*. Publications Mathematicae, vol. 6, page 290, 1959. (ver pág 21 y 28.)
- [Erdős 1960] Paul Erdős and Alfréd Rényi. *On the Evolution of Random Graphs*. Publication of the Mathematical Institute of the Hungarian Academy of Science, pages 17–61, 1960. (ver pág 21.)
- [Fronczak 2004] Agata Fronczak, Piotr Fronczak and Janusz A Hołyst. *Average path length in random networks*. Phys. Rev. E, vol. 70, no. 5 Pt 2, page 056110, November 2004. (ver pág 21 y 22.)
- [Fu 2011] Feng Fu, Daniel I. Rosenbloom, Long Wang and Martin A. Nowak. *Imitation dynamics of vaccination behaviour on social networks*. Proceedings: Biological Sciences, vol. 278, no. 1702, pages 42–49, 2011. (ver pág 35, 36 y 111.)
- [Gastañaduy 2014] Paul A Gastañaduy, Susan B Redd, Amy Parker Fiebelkorn, Jennifer S Rota, Paul A Rota, William J Bellini, Jane F Seward and Gregory S Wallace. *Measles—United States, January 1–May 23, 2014*. MMWR. Morbidity and mortality weekly report, vol. 63, no. 22, page 496, 2014. (ver pág 3.)
- [Gillespie 1976] Daniel T. Gillespie. *A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions*. J. Comput. Phys., vol. 22, page 403, 1976. (ver pág 99.)

- [Gonçalves 2011] Bruno Gonçalves, Perra Nicola and Vespignani Alessandro. *Modeling Users' Activity on Twitter Networks: Validation of Dunbar's Number*. PLoS ONE, vol. 6, no. 8, page e22656, 08 2011. (ver pág 11.)
- [Gonzalez 2008] M. Gonzalez, C. Hidalgo and A. Barabasi. *Understanding individual human mobility patterns*. Nature, no. 453, pages 779–782, June 2008. (ver pág 2 y 11.)
- [Hagberg 2008] Aric A. Hagberg, Daniel A. Schult and Pieter J. Swart. *Exploring network structure, dynamics, and function using NetworkX*. In Proceedings of the 7th Python in Science Conference (SciPy2008), pages 11–15, Pasadena, CA USA, August 2008. (ver pág 70.)
- [Hahne 2010] Susan Hahne, Margreet J M te Wierik, Liesbeth Mollema, Eva van Velzen, Eric de Coster, Corien Swaan, Hester de Melker and Robert van Binnendijk. *Measles outbreak, the Netherlands, 2008*. Emerging Infect. Dis., vol. 16, no. 3, pages 567–9, March 2010. (ver pág 4.)
- [Holme 2002] Petter Holme and Beom J. Kim. *Growing scale-free networks with tunable clustering*. Physical Review E, vol. 65, no. 2, page 026107, 2002. (ver pág 31, 78 y 80.)
- [HPA ] (ver pág 3.)
- [Isella 2011] Lorenzo Isella, Juliette Stehlé, Alain Barrat, Ciro Cattuto, Jean-François Pinton and Wouter Van den Broeck. *What's in a crowd? Analysis of face-to-face behavioral networks*. Journal of Theoretical Biology, vol. 271, no. 1, pages 166–180, 2011. (ver pág 2, 11, 74, 92 y 112.)
- [Jacobson 2007] Robert M Jacobson, Paul V Targonski and Gregory A Poland. *A taxonomy of reasoning flaws in the anti-vaccine movement*. Vaccine, vol. 25, no. 16, pages 3146–3152, 2007. (ver pág 109.)
- [Karinthy 1929] Frigyes Karinthy. *Chain-links*. Everything is the Other Way, 1929. (ver pág 29.)
- [Keeling 1997a] M. J. Keeling, D. A. Rand and A. J. Morris. *Correlation Models for Childhood Epidemics*. Proceedings: Biological Sciences, vol. 264, no. 1385, pages 1149–1156, August 1997. (ver pág 97.)

- [Keeling 1997b] MJ Keeling and BT Grenfell. *Disease extinction and community size: modeling the persistence of measles*. Science, vol. 275, no. 5296, pages 65–67, 1997. (ver pág 12, 16, 17, 97, 101 y 105.)
- [Keeling 1998] MJ Keeling and BT Grenfell. *Effect of variability in infection period on the persistence and spatial spread of infectious diseases*. Mathematical biosciences, vol. 147, no. 2, pages 207–226, 1998. (ver pág 16, 17, 101 y 105.)
- [Klemm 2003] Konstantin Klemm, Víctor M Eguíluz, Raúl Toral and Maxi San Miguel. *Nonequilibrium transitions in complex networks: a model of social interaction*. Phys Rev E Stat Nonlin Soft Matter Phys, vol. 67, no. 2 Pt 2, page 026120, February 2003. (ver pág 8 y 40.)
- [Klimek 2013] Peter Klimek and Stefan Thurner. *Triadic closure dynamics drives scaling-laws in social multiplex networks*. New J. Phys., vol. 15, page 063008, June 2013. (ver pág 23, 24 y 90.)
- [Knuth 1993] D. E. Knuth. The Stanford GraphBase: a platform for combinatorial computing. ACM Press New York, NY, USA, 1993. (ver pág 27.)
- [Kossinets 2006] Gueorgi Kossinets and Duncan J. Watts. *Empirical Analysis of an Evolving Social Network*. Science, vol. 311, no. 5757, pages 88–90, January 2006. (ver pág 23, 77 y 78.)
- [Krapivsky 2000] P. L. Krapivsky, S. Redner and F. Leyvraz. *Connectivity of Growing Random Networks*. Physical Review Letters, vol. 85, no. 21, pages 4629–4632, 2000. (ver pág 79 y 84.)
- [Krapivsky 2001] P L Krapivsky, G J Rodgers and S Redner. *Degree distributions of growing networks*. Phys. Rev. Lett., vol. 86, no. 23, pages 5401–4, June 2001. (ver pág 79.)
- [Lazarsfeld 1954] Paul F Lazarsfeld, Robert K Merton et al. *Friendship as a social process: A substantive and methodological analysis*. Freedom and control in modern society, vol. 18, no. 1, pages 18–66, 1954. (ver pág 7 y 35.)
- [Lazer 2009] David Lazer, Alex Sandy Pentland, Lada Adamic, Sinan Aral, Albert Laszlo Barabasi, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann et al. *Life in the network: the*

- coming age of computational social science*. Science (New York, NY), vol. 323, no. 5915, page 721, 2009. (ver pág 2 y 11.)
- [Lerman 2010] Kristina Lerman and Rumi Ghosh. *Information Contagion: An Empirical Study of the Spread of News on Digg and Twitter Social Networks*. ICWSM, vol. 10, pages 90–97, 2010. (ver pág 11 y 52.)
- [Liben-Nowell 2007] David Liben-Nowell and Jon M. Kleinberg. *The link-prediction problem for social networks*. JASIST, vol. 58, no. 7, pages 1019–1031, 2007. (ver pág 78.)
- [Lloyd 2001] Alun L Lloyd. *Realistic distributions of infectious periods in epidemic models: changing patterns of persistence and dynamics*. Theoretical population biology, vol. 60, no. 1, pages 59–71, 2001. (ver pág 16, 97, 101, 103 y 116.)
- [London 1973] Wayne P London and James A Yorke. *Recurrent outbreaks of measles, chickenpox and mumps I. Seasonal variation in contact rates*. American journal of epidemiology, vol. 98, no. 6, pages 453–468, 1973. (ver pág 15.)
- [Maslov 2002] S. Maslov and K. Sneppen. *Specificity and stability in topology of protein networks*. Science, vol. 296, no. 5569, page 910, 2002. (ver pág 25.)
- [Maurice 2009] John M Maurice and Sheila Davey. *State of the World’s Vaccines and Immunization*. World Health Organization, 2009. (ver pág 2.)
- [May 2003] Thomas May and Ross D Silverman. *‘Clustering of exemptions’ as a collective action threat to herd immunity*. Vaccine, vol. 21, no. 11, pages 1048–1051, 2003. (ver pág 109 y 111.)
- [Mbah 2012] Martial L Ndeffo Mbah, Jingzhou Liu, Chris T Bauch, Yonas I Tekel, Jan Medlock, Lauren Ancel Meyers and Alison P Galvani. *The impact of imitation on vaccination behavior in social contact networks*. PLoS Comput. Biol., vol. 8, no. 4, page e1002469, 2012. (ver pág 5, 6 y 36.)
- [McPherson 2001] Miller McPherson, Lynn Smith-Lovin and James M. Cook. *Birds of a Feather: Homophily in Social Networks*. Annu. Rev. Sociol., vol. 27, no. 1, pages 415–444, 2001. (ver pág 7, 35 y 110.)

- [Medus 2005] A. Medus, G. Acuña and C. O. Dorso. *Detection of community structures in networks via global optimization*. Physica A, vol. 358, no. 2-4, pages 593–604, December 2005. (ver pág 10, 24, 27 y 28.)
- [Medus 2009] A D Medus and C O Dorso. *Alternative approach to community detection in networks*. Phys. Rev. E, vol. 79, no. 6 Pt 2, page 066111, June 2009. (ver pág 28.)
- [Medus 2014a] AD Medus and CO Dorso. *Memory effects induce structure in social networks with activity-driven agents*. Journal of Statistical Mechanics: Theory and Experiment, vol. 9, page 009, 2014. (ver pág 112.)
- [Medus 2014b] Andrés D Medus and Claudio O Dorso. *Vaccination and public trust: a model for the dissemination of vaccination behavior with external intervention*. arXiv preprint arXiv:1412.0583, 2014. (ver pág 66 y 111.)
- [Miguel 2005] M.S. Miguel, V.M. Eguiluz, R. Toral and K. Klemm. *Binary and multivariate stochastic models of consensus formation*. Computing in Science Engineering, vol. 7, no. 6, pages 67–73, Nov 2005. (ver pág 9.)
- [Milgram 1967] Stanley Milgram. *The Small World Problem*. Psychology Today, vol. 67, no. 1, pages 61–67, 1967. (ver pág 10 y 29.)
- [Molloy 1995] M. Molloy and B. Reed. *A critical point for random graphs with a given degree sequence*, 1995. (ver pág 29 y 107.)
- [Molloy 1998] M. Molloy and B. Reed. *The size of the giant component of a random graph with a given degree sequence*. Combin. Probab. Comput., vol. 7, page 295, 1998. (ver pág 29.)
- [Muchnik 2013] Lev Muchnik, Sen Pei, Lucas C Parra, Saulo D S Reis, José S Andrade, Shlomo Havlin and Hernán A Makse. *Origins of power-law degree distribution in the heterogeneity of human activity in social networks*. Sci Rep, vol. 3, page 1783, 2013. (ver pág 32, 75 y 81.)
- [Newman 2001a] M. E. J. Newman. *The structure of scientific collaboration networks*. PNAS, vol. 98, no. 2, pages 404–409, 2001. (ver pág 30.)
- [Newman 2001b] M. E. J. Newman, S. H. Strogatz and D. J. Watts. *Random graphs with arbitrary degree distributions and their applications*. Phys. Rev. E, vol. 64, no. 2, page 026118, July 2001. (ver pág 29.)

- [Newman 2002a] M. E. J. Newman. *Assortative Mixing in Networks*. Phys. Rev. Lett., vol. 89, page 208701, 2002. (ver pág 25, 26, 77 y 87.)
- [Newman 2002b] M E J Newman, D J Watts and S H Strogatz. *Random graph models of social networks*. Proc. Natl. Acad. Sci. U.S.A., vol. 99 Suppl 1, pages 2566–72, February 2002. (ver pág 10 y 29.)
- [Newman 2003a] M. E. J. Newman. *Mixing patterns in networks*. Phys. Rev. E, vol. 67, page 026126, 2003. (ver pág 25, 26 y 87.)
- [Newman 2003b] M E J Newman and Juyong Park. *Why social networks are different from other types of networks*. Phys Rev E Stat Nonlin Soft Matter Phys, vol. 68, no. 3 Pt 2, page 036122, September 2003. (ver pág 77.)
- [Newman 2003c] Mark E. J. Newman. Random graphs as models of networks, pages 35–68. Wiley-VCH Verlag GmbH & Co. KGaA, 2003. (ver pág 24.)
- [Newman 2003d] M.E.J. Newman. *The structure and function of complex networks*. SIAM Review, vol. 45, no. 2, pages 167–256, 2003. (ver pág 21 y 30.)
- [Newman 2004] M. E. J. Newman and M. Girvan. *Finding and evaluating community structure in networks*. Phys. Rev. E, vol. 69, no. 2, page 026113, February 2004. (ver pág 10, 24, 26 y 28.)
- [Papadopoulos 2012] Fragkiskos Papadopoulos, Maksim Kitsak, M Ángeles Serrano, Marián Boguñá and Dmitri Krioukov. *Popularity versus similarity in growing networks*. Nature, vol. 489, no. 7417, pages 537–40, September 2012. (ver pág 11.)
- [Parker 2006] Amy A Parker, Wayne Staggs, Gustavo H Dayan, Ismael R Ortega-Sánchez, Paul A Rota, Luis Lowe, Patricia Boardman, Robert Teclaw, Charlene Graves and Charles W LeBaron. *Implications of a 2005 measles outbreak in Indiana for sustained elimination of measles in the United States*. New England Journal of Medicine, vol. 355, no. 5, pages 447–455, 2006. (ver pág 4.)
- [Pastor-Satorras 2001a] Romualdo Pastor-Satorras, Alexei Vázquez and Alessandro Vespignani. *Dynamical and correlation properties of the Internet*. Phys. Rev. Lett., vol. 87, no. 25, page 258701, 2001. (ver pág 24 y 87.)



- [Pastor-Satorras 2001b] Romualdo Pastor-Satorras and Alessandro Vespignani. *Epidemic Spreading in Scale-Free Networks*. Phys. Rev. Lett., vol. 86, pages 3200–3203, Apr 2001. (ver pág 32.)
- [Pearce 2008] Anna Pearce, Catherine Law, David Elliman, Tim J Cole and Helen Bedford. *Factors associated with uptake of measles, mumps, and rubella vaccine (MMR) and use of single antigen vaccines in a contemporary UK cohort: prospective cohort study*. bmj, vol. 336, no. 7647, pages 754–757, 2008. (ver pág 63.)
- [Perra 2012] N Perra, B Gonçalves, R Pastor-Satorras and A Vespignani. *Activity driven modeling of time varying networks*. Sci Rep, vol. 2, page 469, 2012. (ver pág 11, 32, 74, 75, 81, 82 y 111.)
- [PLAN 2012] STRATEGIC PLAN. *GLOBAL MEASLES AND RUBELLA WHO*. 2012. (ver pág 62 y 63.)
- [Plotkin 2004] Susan L Plotkin and Stanley A Plotkin. *A short history of vaccination*. Vaccines, vol. 5, pages 1–16, 2004. (ver pág 2.)
- [Radicchi 2004] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto and D. Parisi. *Defining and identifying communities in networks*. Proceedings of the National Academy of Sciences, vol. 101, no. 9, page 2658, 2004. (ver pág 26.)
- [Ravasz 2003] E. Ravasz and A. L. Barabasi. *Hierarchical organization in complex networks*. Physical Review E, vol. 67, page 026112, 2003. (ver pág 22 y 24.)
- [Saint-Victor 2013] Diane S Saint-Victor and Saad B Omer. *Vaccine refusal and the endgame: walking the last mile first*. Philosophical Transactions of the Royal Society B: Biological Sciences, vol. 368, no. 1623, page 20120148, 2013. (ver pág 63.)
- [Salathé 2008] Marcel Salathé and Sebastian Bonhoeffer. *The effect of opinion clustering on disease outbreaks*. J R Soc Interface, vol. 5, no. 29, pages 1505–8, December 2008. (ver pág 5, 6, 35 y 109.)
- [Salathé 2011] Marcel Salathé and Shashank Khandelwal. *Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control*. PLoS Comput. Biol., vol. 7, no. 10, page e1002199, October 2011. (ver pág 111.)

- [Salmon 2006] Daniel A Salmon, Stephen P Teret, C Raina MacIntyre, David Salisbury, Margaret A Burgess and Neal A Halsey. *Compulsory vaccination and conscientious or philosophical exemptions: past, present, and future*. The Lancet, vol. 367, no. 9508, pages 436–442, 2006. (ver pág 63.)
- [Schenzle 1984] Dieter Schenzle. *An age-structured model of pre-and post-vaccination measles transmission*. Mathematical Medicine and Biology, vol. 1, no. 2, pages 169–191, 1984. (ver pág 16.)
- [Serrano 2005] M. Angeles Serrano and Marian Boguna. *Tuning clustering in random networks with arbitrary degree distributions*. Phys. Rev. E, vol. 72, page 036133, July 2005. (ver pág 24.)
- [Sheldon 2000] Tony Sheldon. *Netherlands faces measles epidemic*. BMJ: British Medical Journal, vol. 320, no. 7227, page 76, 2000. (ver pág 4.)
- [Simon 1955] Herbert A Simon. *On a class of skew distribution functions*. Biometrika, pages 425–440, 1955. (ver pág 10 y 31.)
- [Song 2010] Chaoming Song, Zehui Qu, Nicholas Blumm and Albert-László Barabási. *Limits of Predictability in Human Mobility*. Science, vol. 327, no. 5968, pages 1018–1021, 2010. (ver pág 2 y 11.)
- [Starnini 2013a] Michele Starnini, Andrea Baronchelli and Romualdo Pastor-Satorras. *Modeling human dynamics of face-to-face interaction networks*. Phys. Rev. Lett., vol. 110, no. 16, page 168701, April 2013. (ver pág 11, 74, 111 y 112.)
- [Starnini 2013b] Michele Starnini and Romualdo Pastor-Satorras. *Topological properties of a time-integrated activity-driven network*. Phys. Rev. E, vol. 87, page 062807, Jun 2013. (ver pág 111 y 121.)
- [Sugerman 2010] David E Sugerman, Albert E Barskey, Maryann G Delea, Ismael R Ortega-Sanchez, Daoling Bi, Kimberly J Ralston, Paul A Rota, Karen Waters-Montijo and Charles W Lebaron. *Measles outbreak in a highly vaccinated population, San Diego, 2008: role of the intentionally undervaccinated*. Pediatrics, vol. 125, no. 4, pages 747–55, April 2010. (ver pág 4.)
- [Szell 2010a] Michael Szell, Renaud Lambiotte and Stefan Thurner. *Multirelational organization of large-scale social networks in an online world*. Proc.

- Natl. Acad. Sci. U.S.A., vol. 107, no. 31, pages 13636–41, August 2010. (ver pág 2 y 23.)
- [Szell 2010b] Michael Szell and Stefan Thurner. *Measuring social dynamics in a massive multiplayer online game*. Social Networks, vol. 32, no. 4, pages 313–329, 2010. (ver pág 23.)
- [Taylor 1999] Brent Taylor, Elizabeth Miller, CPaddy Farrington, Maria-Christina Petropoulos, Isabelle Favot-Mayaud, Jun Li and Pauline A Waight. *Autism and measles, mumps, and rubella vaccine: no epidemiological evidence for a causal association*. The Lancet, vol. 353, no. 9169, pages 2026–2029, 1999. (ver pág 3.)
- [Toivonen 2006] Riitta Toivonen, Jukka-Pekka Onnela, Jari Saramäki, Jörkki Hyvönen and Kimmo Kaski. *A model for social networks*. Physica A: Statistical Mechanics and its Applications, vol. 371, no. 2, pages 851–860, 2006. (ver pág 78.)
- [Ugander 2011] Johan Ugander, Brian Karrer, Lars Backstrom and Cameron Marlow. *The Anatomy of the Facebook Social Graph*, 2011. cite arxiv:1111.4503Comment: 17 pages, 9 figures, 1 table. (ver pág 2 y 11.)
- [Vázquez 2002] A. Vázquez, R. Pastor-Satorras and A. Vespignani. *Large-scale topological and dynamical properties of the Internet*. Phys. Rev. E, vol. 65, no. 6, page 66130, 2002. (ver pág 24.)
- [Vázquez 2003] Alexei Vázquez. *Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations*. Physical Review E, vol. 67, no. 5, page 056104, 2003. (ver pág 78 y 80.)
- [Verdasca 2005] J Verdasca, MM Telo da Gama, A Nunes, NR Bernardino, JM Pacheco and MC Gomes. *Recurrent epidemics in small world networks*. Journal of Theoretical Biology, vol. 233, no. 4, pages 553–561, 2005. (ver pág 17 y 98.)
- [Viswanath 2009] Bimal Viswanath, Alan Mislove, Meeyoung Cha and P. Krishna Gummadi. *On the evolution of user interaction in Facebook*. In Jon Crowcroft and Balachander Krishnamurthy, editors, WOSN, pages 37–42. ACM, 2009. (ver pág 11, 93 y 112.)

- [Wakefield 1998] Andrew J Wakefield, Simon H Murch, Andrew Anthony, John Linnell, DM Casson, Mohsin Malik, Mark Berelowitz, Amar P Dhillon, Michael A Thomson, Peter Harvey *et al.* *RETRACTED: Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children.* The Lancet, vol. 351, no. 9103, pages 637–641, 1998. (ver pág 3 y 9.)
- [Wal 2013] *Outbreak of Measles in Wales Nov 2012 – July 2013: Report of the agencies which responded to the outbreak.* Rapport technique, October 2013. (ver pág 3.)
- [Watts 1998] D.J. Watts and S.H. Strogatz. *Collective dynamics of 'small-world' networks.* Nature, no. 393, pages 440–442, 1998. (ver pág 10 y 30.)
- [WHO 2012] WHO. GLOBAL MEASLES AND RUBELLA - STRATEGIC PLAN. World Health Organization, 2012. (ver pág 4 y 48.)
- [Xia 2013] Shang Xia and Jiming Liu. *A computational approach to characterizing the impact of social influence on individuals' vaccination decision making.* PLoS ONE, vol. 8, no. 4, page e60373, 2013. (ver pág 5, 35 y 36.)
- [Yang 2010] Jaewon Yang and Jure Leskovec. *Modeling information diffusion in implicit networks.* In Data Mining (ICDM), 2010 IEEE 10th International Conference on, pages 599–608. IEEE, 2010. (ver pág 11.)
- [Yule 1925] G Udny Yule. *A mathematical theory of evolution, based on the conclusions of Dr. JC Willis, FRS.* Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character, pages 21–87, 1925. (ver pág 10 y 31.)
- [Zachary 1977] W. W. Zachary. *An information flow model for conflict and fission in small groups.* Journal of Anthropological Research, vol. 33, pages 452–473, 1977. (ver pág 27.)