PARTIMOS EN BREVE

MUCHAS GRACIAS

Dadas todas las circunstancias y de manera extraordinaria, el curso se cerrará así:

Las tareas tienen una ponderación de 20%. El otro 80% será una de las otras evaluaciones donde existen dos opciones:

1) Examen 1 o Examen 2 o promedio talleres [el más alto]

2) Evaluación colegiada talleres.

https://forms.gle/9fiHKSrnvYnAnwae7

Network Science

dataScience UDD

NICO NORTHWESTERN INSTITUTE ON COMPLEX SYSTEMS

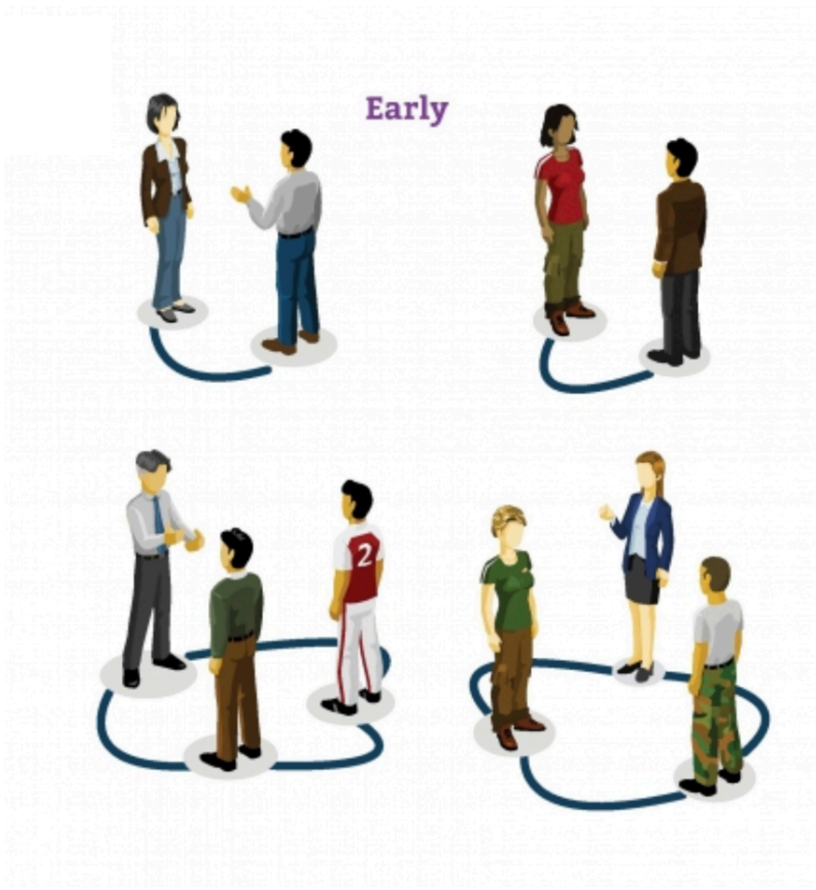
Cristian Candia-Castro Vallejos, Ph.D.

cristiancandia@udd.cl

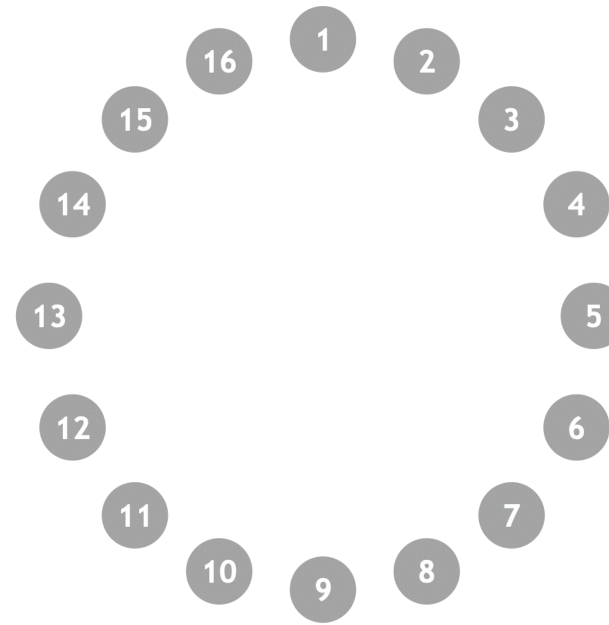
Director Magister en Data Science UDD
Profesor Investigador, Facultad de Ingeniería, UDD
External Faculty Northwestern Institute on Complex Systems, Kellogg School of Management, Northwestern University

Introducción

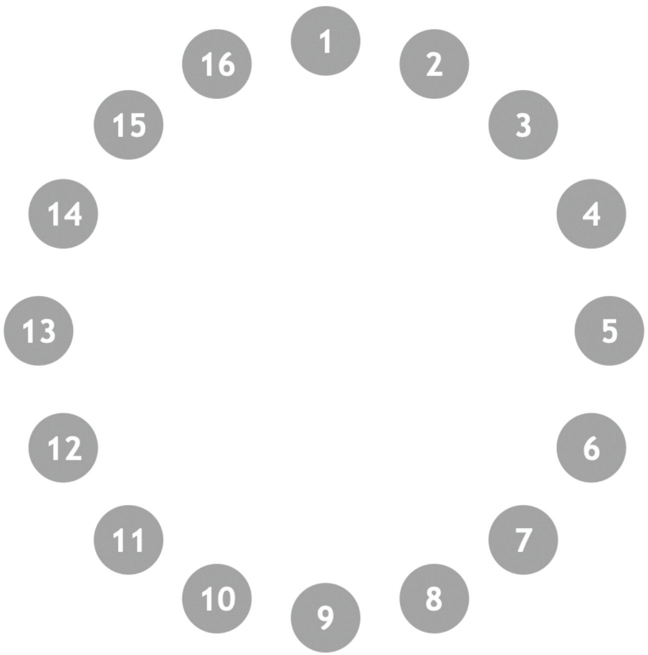
MODELO DE REDES ALEATORIAS



the party algorithm

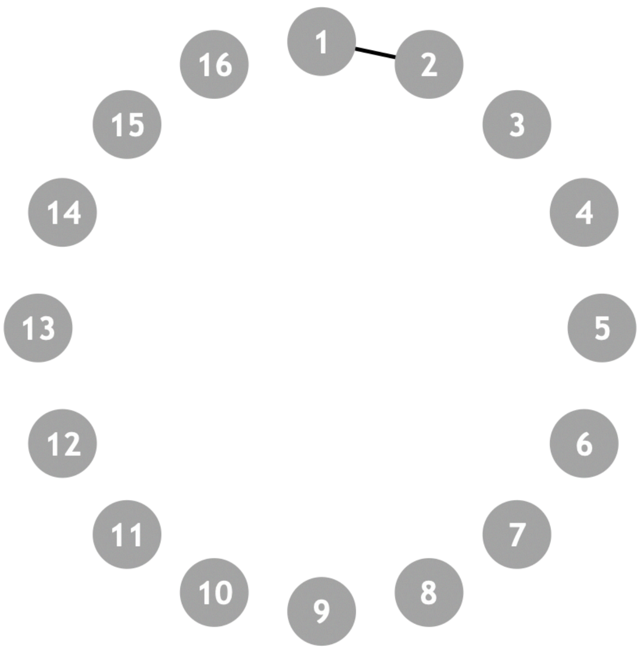


the party algorithm



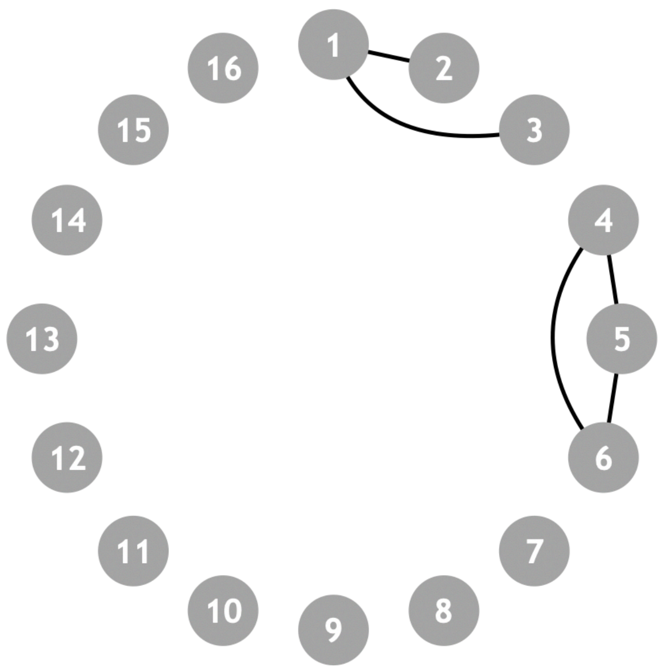
Hagamos un seguimiento de quién habla con quién

the party algorithm



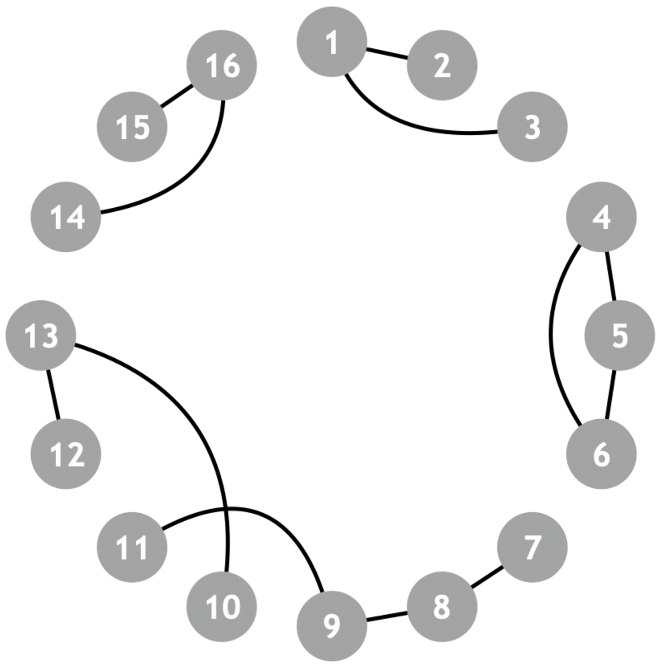
Hagamos un seguimiento de quién habla con quién

the party algorithm



Hagamos un seguimiento de quién habla con quién

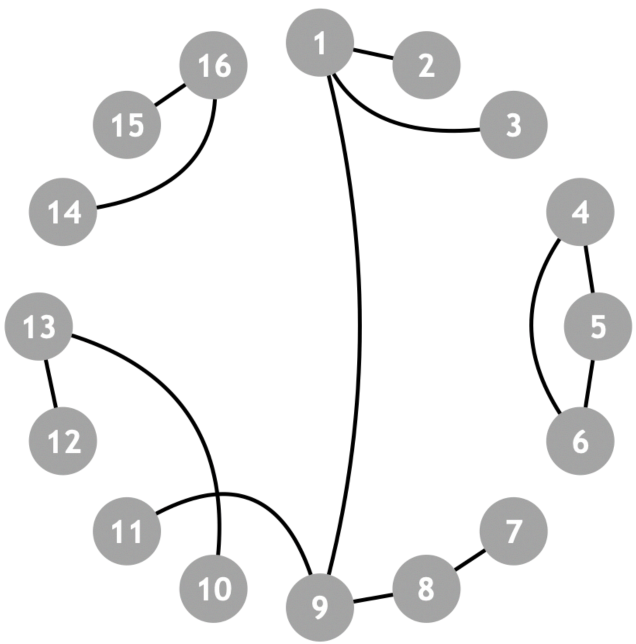
the party algorithm



Hagamos un seguimiento de quién habla con quién

Al principio, la gente habla principalmente con otros que están sentados cerca.

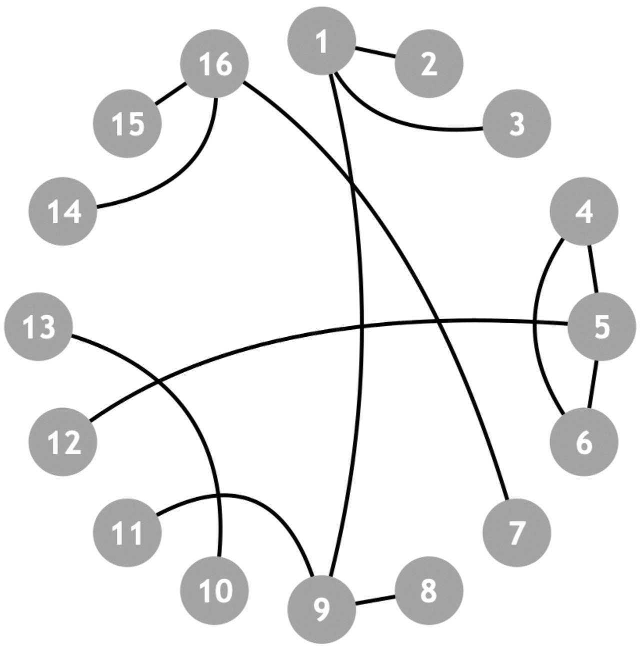
the party algorithm



Hagamos un seguimiento de quién habla con quién

Sin embargo, como la cena avanza hacia las bebidas nocturnas y los invitados se ponen de pie, vemos la reorganización de sus relaciones y se forman nuevos vínculos

the party algorithm

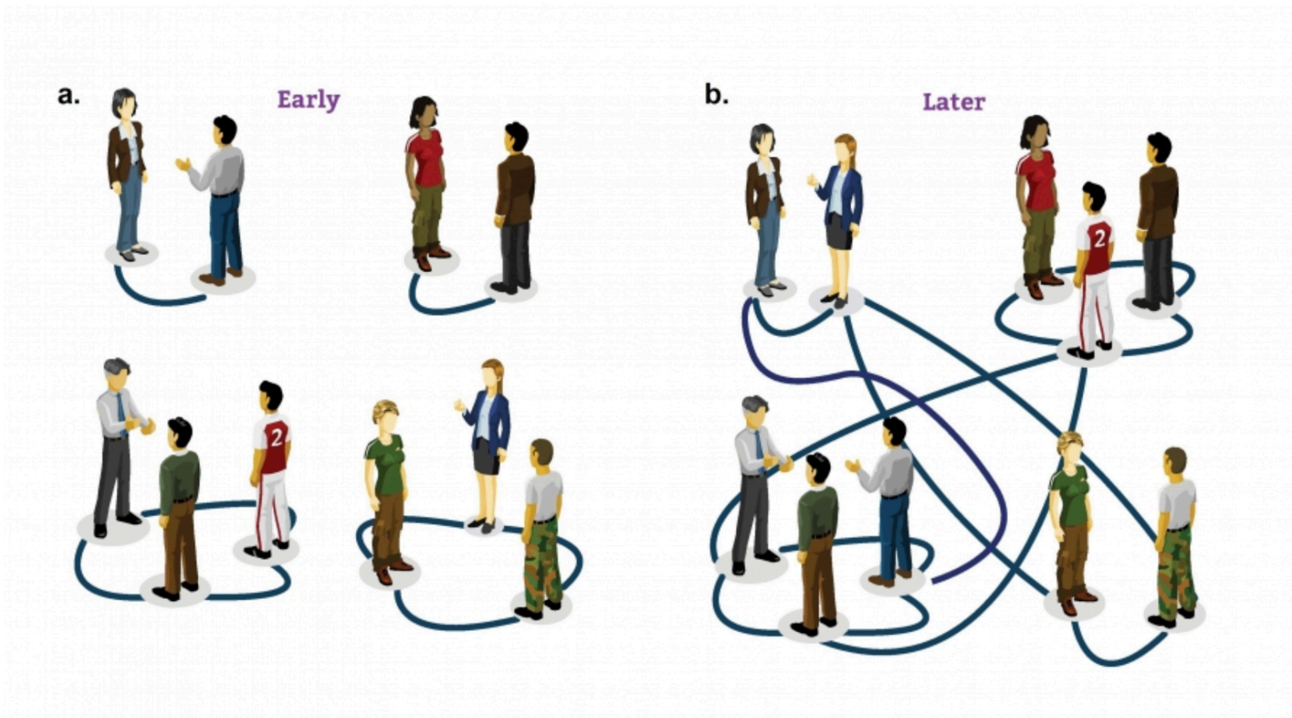


Estas redes no siguen ninguna regla física como en los ejemplos anteriores.

En cambio, son algo aleatorias.
¿No te parece?

Pregunta: ¿Podemos construir un modelo que intente capturar las propiedades de estas redes?

Un algoritmo simple para redes sociales

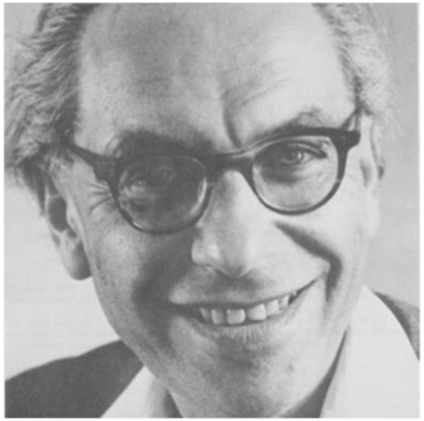


Invitados a una cena se reunen al azar y establecen relaciones sociales.

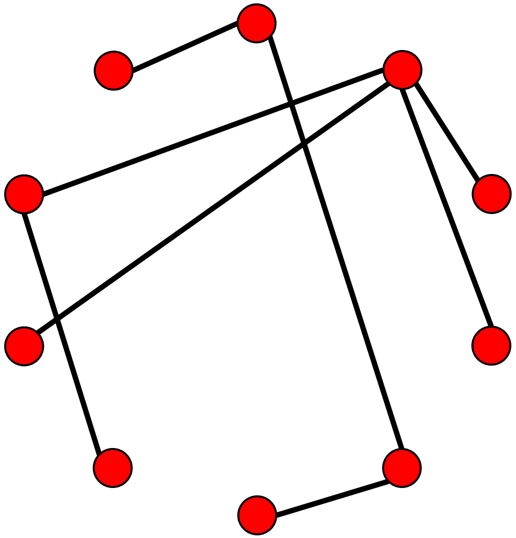
El modelo de redes aleatorias

MODELO DE REDES ALEATORIAS

Pául Erdös
(1913-1996)



Alfréd Rényi
(1921-1970)



Erdös-Rényi model (1960)

Se conectan con probabilidad “p”

p=1/6 N=10

<k> ~ 1.5

MODELO DE REDES ALEATORIAS

Definición:

Un **gafo aleatorio** es un grafo de **N** nodos donde cada par de nodos está conectado con una probabilidad **p.**

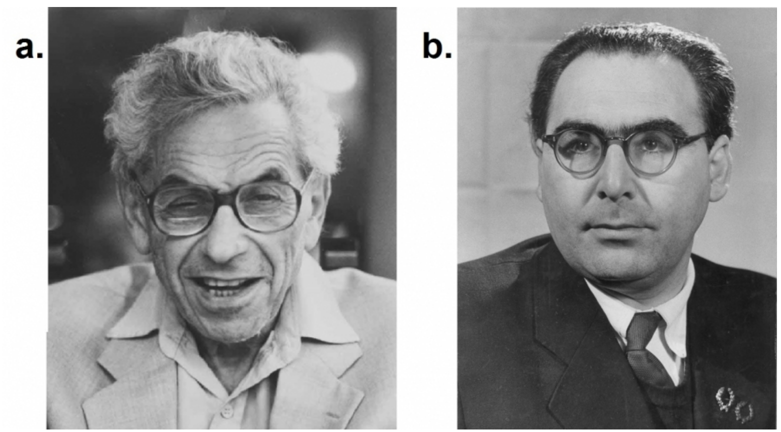
Modelo G(N,L)

N nodos etiquetados están conectados con **L** enlaces colocados aleatoriamente. Erdös y Rényi usaron esta definición en sus papers sobre redes aleatorias.

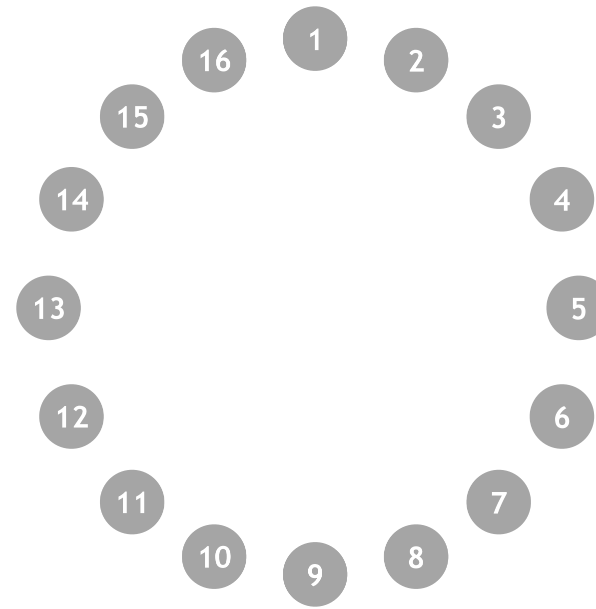
Modelo G(N,p)

Cada par de N nodos etiquetados está conectado con una probabilidad **p,** modelo introducido por Gilbert.

Random Networks



Pál Erdős Alfréd Rényi



Erdős-Rényi algorithm

condición inicial Comenzando con N nodos desconectados (L = 0 enlaces)

paso iterativo Para cada par de individuos, establezca un vínculo con probabilidad p

Random Networks

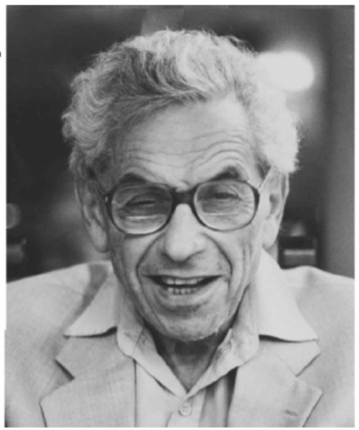
Conectar 1 con 2?

is Random() < p

Si

No

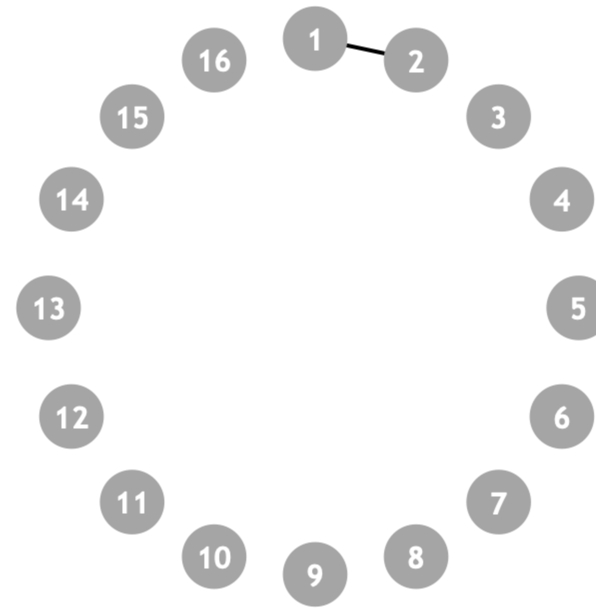
a.



Pál Erdős

b.

Alfréd Rényi



Erdős-Rényi algorithm

condición inicial Comenzando con N nodos desconectados (L = 0 enlaces)

paso iterativo Para cada par de individuos, establezca un vínculo con probabilidad p

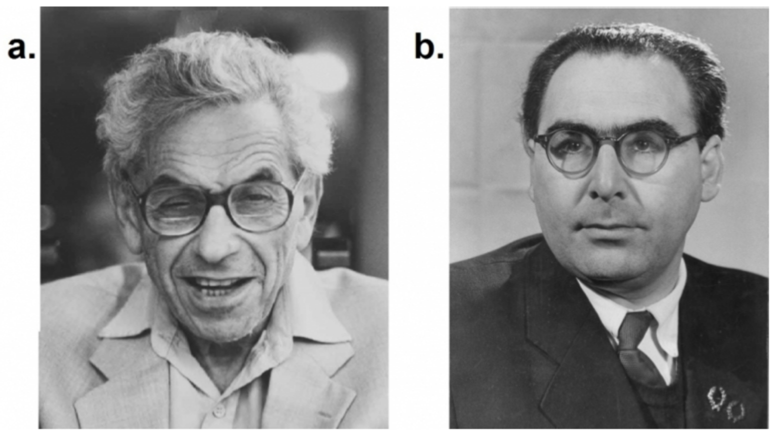
Random Networks

Conectar 1 con 3?

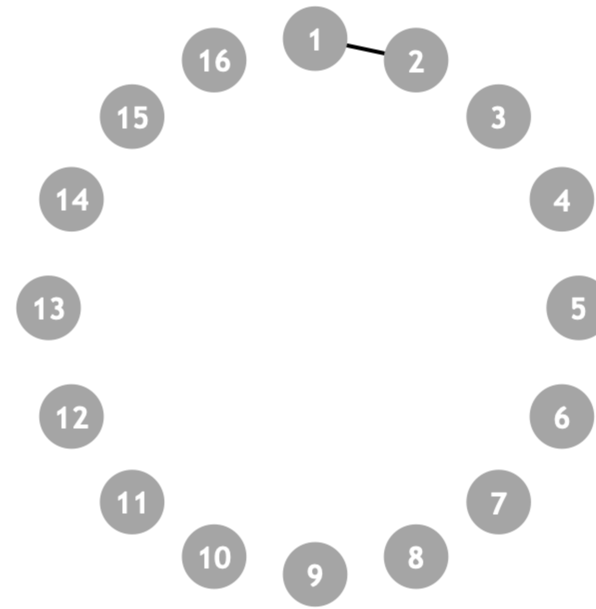
is Random() < p

Sí

No



Pál Erdős   Alfréd Rényi



Erdős-Rényi algorithm

condición inicial Comenzando con N nodos desconectados (L = 0 enlaces)

paso iterativo Para cada par de individuos, establezca un vínculo con probabilidad p

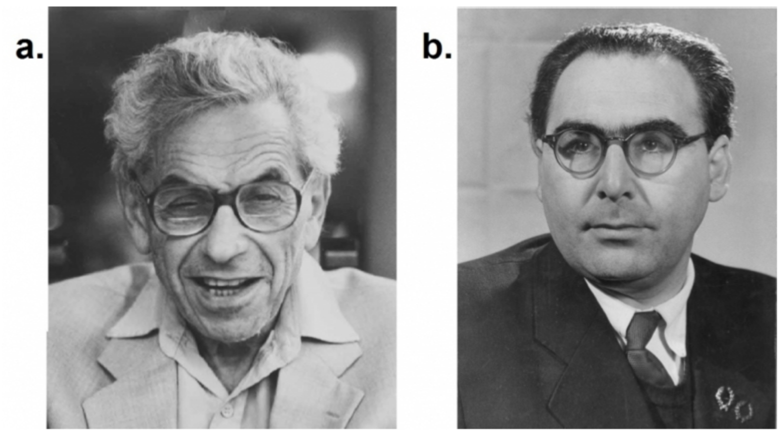
Random Networks

Conectar 1 con 4?

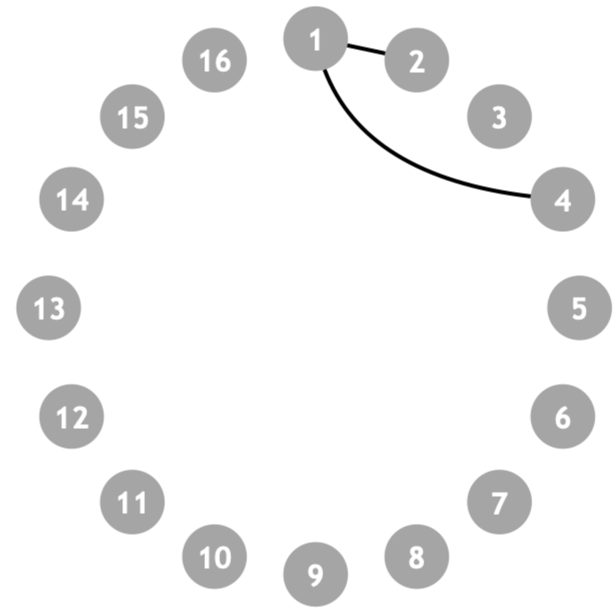
is Random() < p

Si

No



Pál Erdős Alfréd Rényi



Erdős-Rényi algorithm

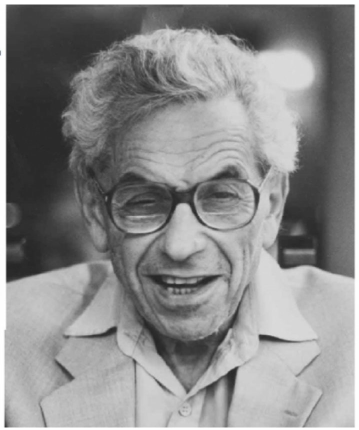
condición inicial Comenzando con N nodos desconectados (L = 0 enlaces)

paso iterativo Para cada par de individuos, establezca un vínculo con probabilidad p

Random Networks

Repite para todas las Posibles conexiones de 1

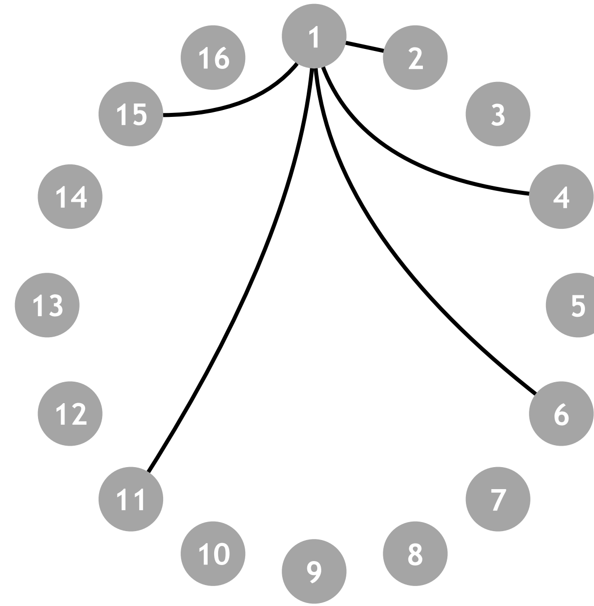
a.



Pál Erdős

b.

Alfréd Rényi



Erdős-Rényi algorithm

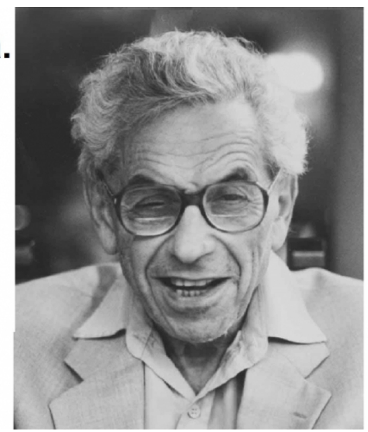
condición inicial Comenzando con N nodos desconectados (L = 0 enlaces)

paso iterativo Para cada par de individuos, establezca un vínculo con probabilidad p

Random Networks

Repite para todos
Los posibles enlaces

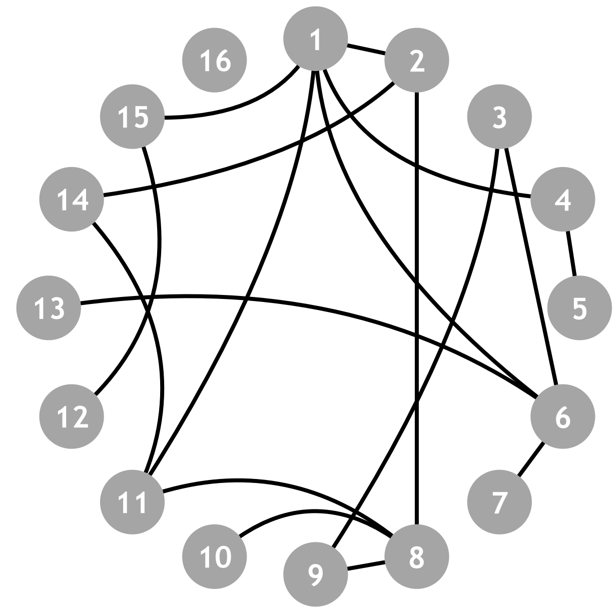
a.



Pál Erdős

b.

Alfréd Rényi



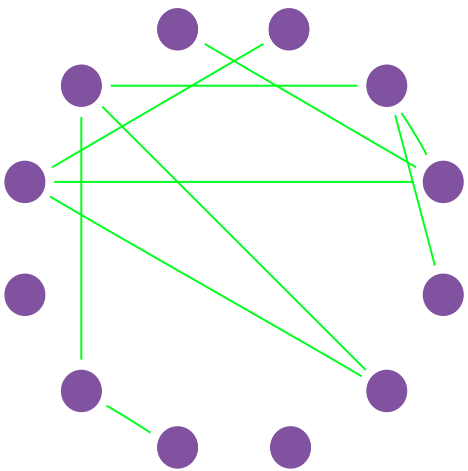
Erdős-Rényi algorithm

condición inicial Comenzando con N nodos desconectados (L = 0 enlaces)

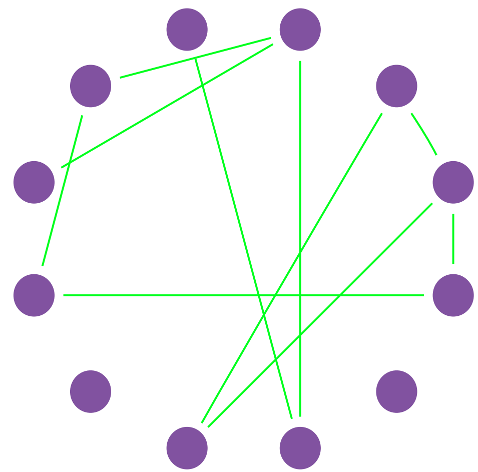
paso iterativo Para cada par de individuos, establezca un vínculo con probabilidad p

MODELO DE REDES ALEATORIAS

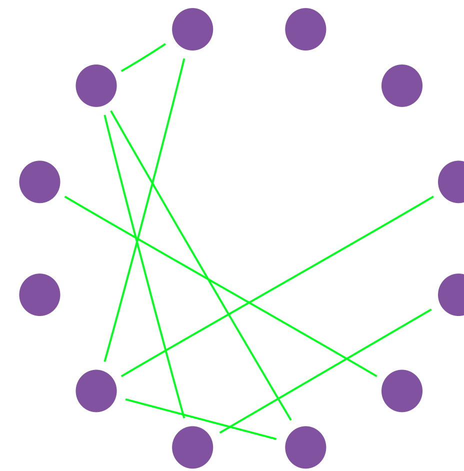
p=1/6
N=12



L=8



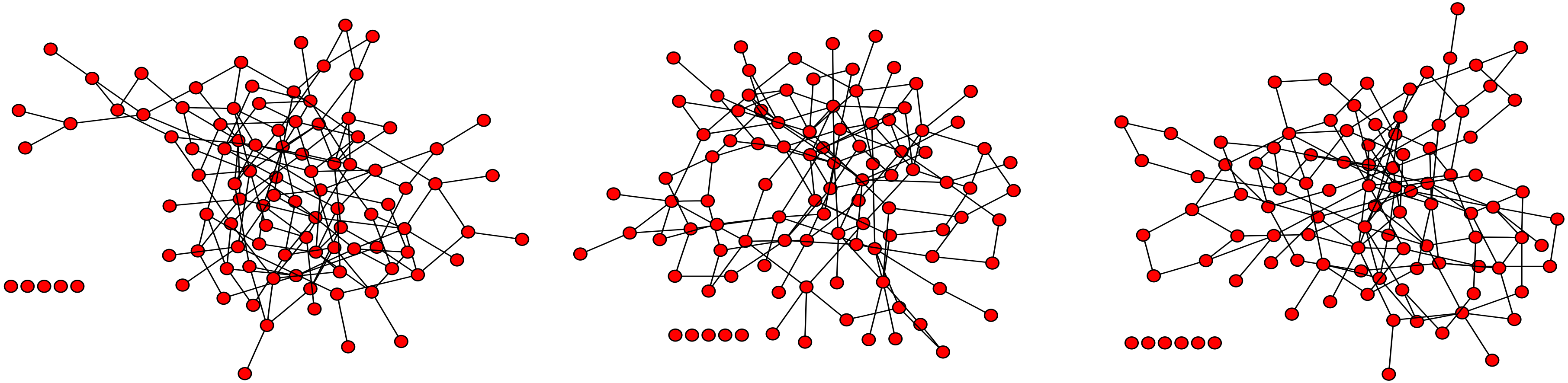
L=10



L=7

MODELO DE REDES ALEATORIAS

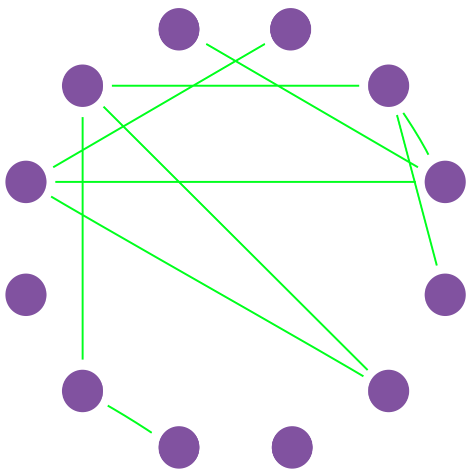
p=0.03
N=100



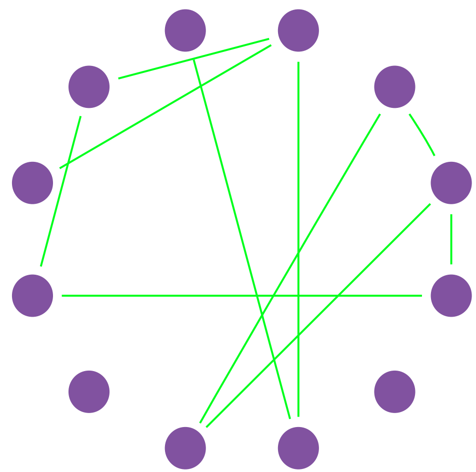
El número de enlaces es variable.

MODELO DE REDES ALEATORIAS

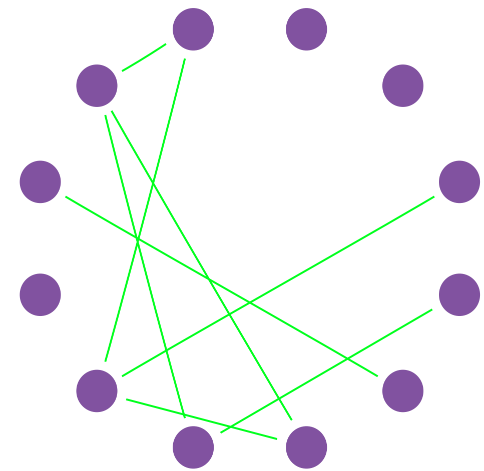
p=1/6
N=12



L=8



L=10



L=7

Número de enlaces en una red aleatoria

P(L): la probabilidad de tener exactamente **L** enlaces en una red de **N** nodos y probabilidad **p**:

El número máximo de enlaces en una red de **N** nodos.

$$P(L) = \underbrace{\begin{pmatrix} \begin{pmatrix} N \\ 2 \end{pmatrix} \\ L \end{pmatrix}}_{\text{Número de formas diferentes en que podemos elegir los enlaces L entre todos los enlaces potenciales.}} p^L (1-p)^{\frac{N(N-1)}{2}-L}$$

Distribución binomial...

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1},$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$

Tutorial Matemático. Distribución Binomial: La línea base

$$p_x = \binom{N}{x} p^x (1-p)^{N-x}$$

Dist. Binomial de "x"

Ejemplo:
N=Enlaces totales potenciales
x=Enlaces de la red
p= probabilidad

$$\langle x \rangle = \sum_{x=0}^{N} x p_x = Np$$

$$< x^2 >= p(1-p)N + p^2 N^2$$

$$\sigma_x = (< x^2 > - < x >^2)^{\frac{1}{2}} \qquad = [p(1-p)N]^{1/2}$$

http://keral2008.blogspot.com/2008/10/derivation-of-mean-and-variance-of.html

MODELO DE REDES ALEATORIAS

P(L): la probabilidad de tener una red de **L** enlaces exactamente

$$P(L) = \binom{\binom{N}{2}}{L} p^L (1-p)^{\frac{N(N-1)}{2}-L}$$

•El número promedio de links <L> en un grafo aleatorio

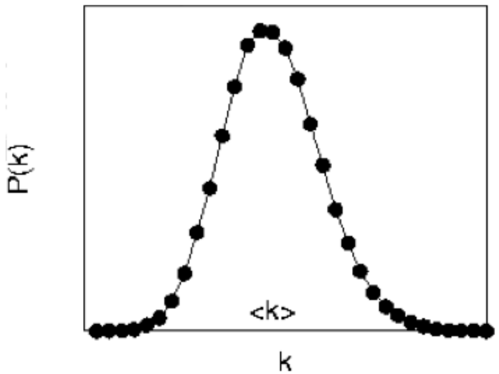
$$< L >= \sum_{L=0}^{\frac{N(N-1)}{2}} LP(L) = p\frac{N(N-1)}{2} \longrightarrow < k >= 2L/N = p(N-1)$$

•La varianza

$$\sigma^2 = p(1-p)\frac{N(N-1)}{2}$$

Distribución de Grado

DISTRIBUCION DE GRADO DE UNA RED ALEATORIA



$$P(k) = \binom{N-1}{k} p^k (1-p)^{(N-1)-k}$$

Selecciona k nodos de un conjunto de N-1 nodos

Probabilidad de tener *k* enlaces

Probabilidiad de perder N-1-k enlaces

$$< k >= p(N-1) \qquad \sigma_k^2 = p(1-p)(N-1)$$

$$\frac{\sigma_k}{< k >} = \left[\frac{1-p}{p}\frac{1}{(N-1)}\right]^{1/2} \approx \frac{1}{(N-1)^{1/2}}$$

A medida que la red aumenta de tamaño, la distribución se vuelve cada vez más estrecha: estamos cada vez más seguros de que el grado de un nodo se encuentra cerca de <k> (valores homogeneos).

DISTRIBUCIÓN DE GRADO DE POISSON

$$P(k) = \binom{N-1}{k} p^k (1-p)^{(N-1)-k}$$

$$<k> = p(N-1)$$

$$p = \frac{<k>}{(N-1)}$$

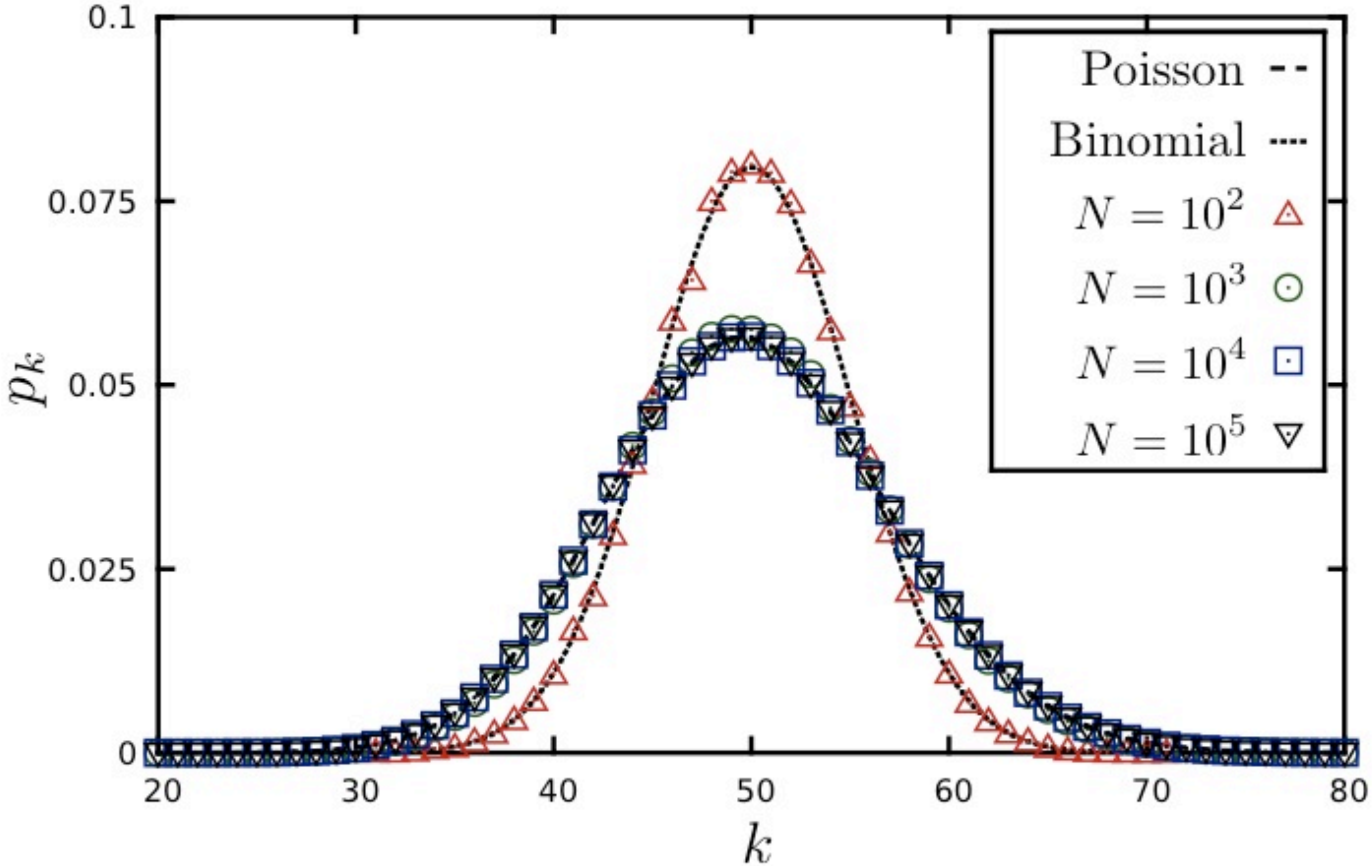
Para *N* grandes y *k* pequeños, llegamos a la distribución de Poisson:

$$P(k) = e^{-<k>} \frac{<k>^k}{k!}$$

DISTRIBUCION DE GRADO DE UNA RED ALEATORIA

<k>=50

$$P(k) = e^{-<k>} \frac{<k>^k}{k!}$$



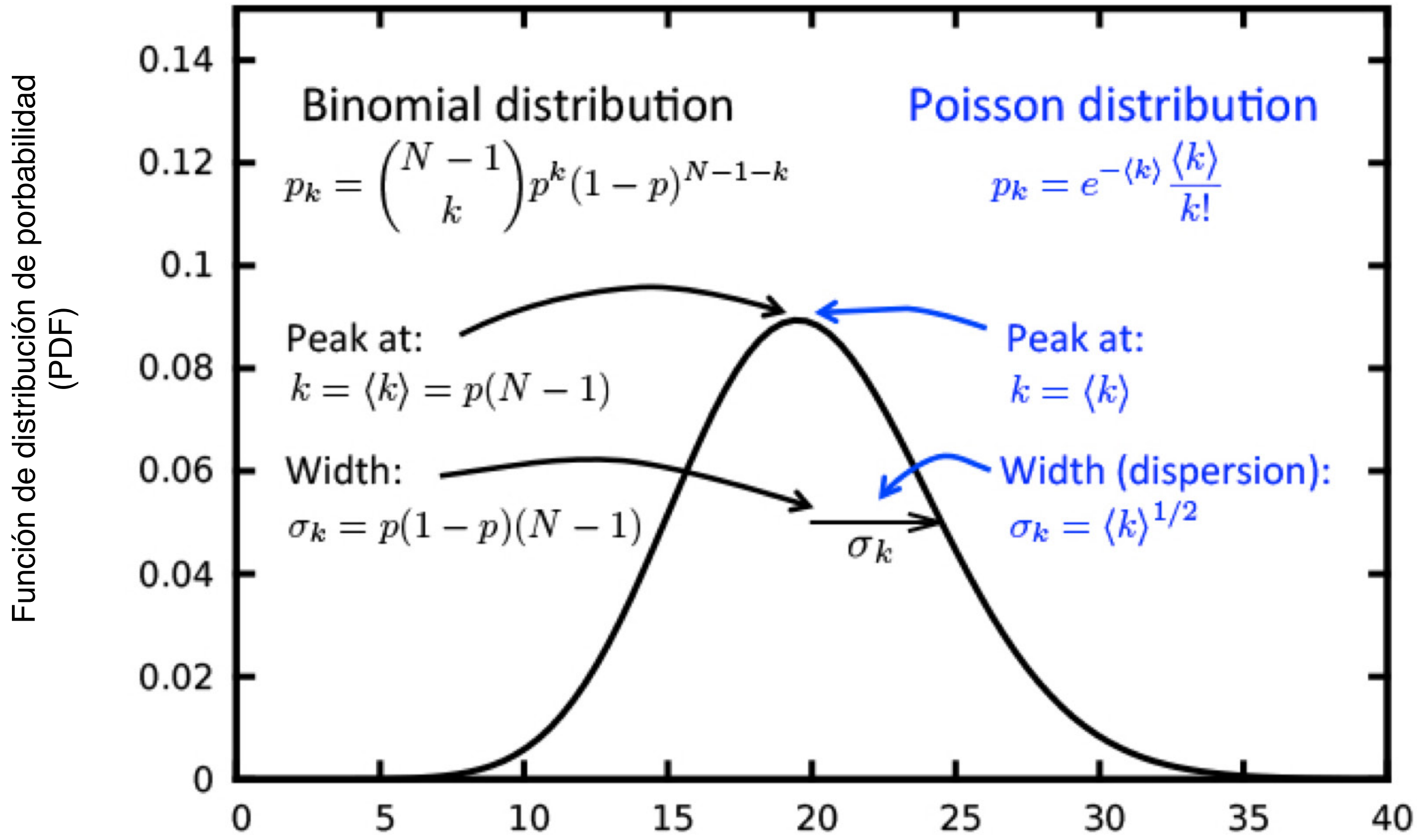
k

P(k): k<<N
Indistinguibles

DISTRIBUCION DE GRADO DE UNA RED ALEATORIA

Resultado exacto
- Distribución Binomial -

Límite para N grande
- Distribución de Poisson -



Las redes reales **no** son Poisson

NO HAY "OUTLIERS" EN UNA SOCIEDAD ALEATORIA

$$P(k) = e^{-<k>} \frac{<k>^k}{k!}$$

→El individuo más conectado tiene grado k_{max}~1,185
→El individuo menos conectado tiene grado k_{min} ~ 816

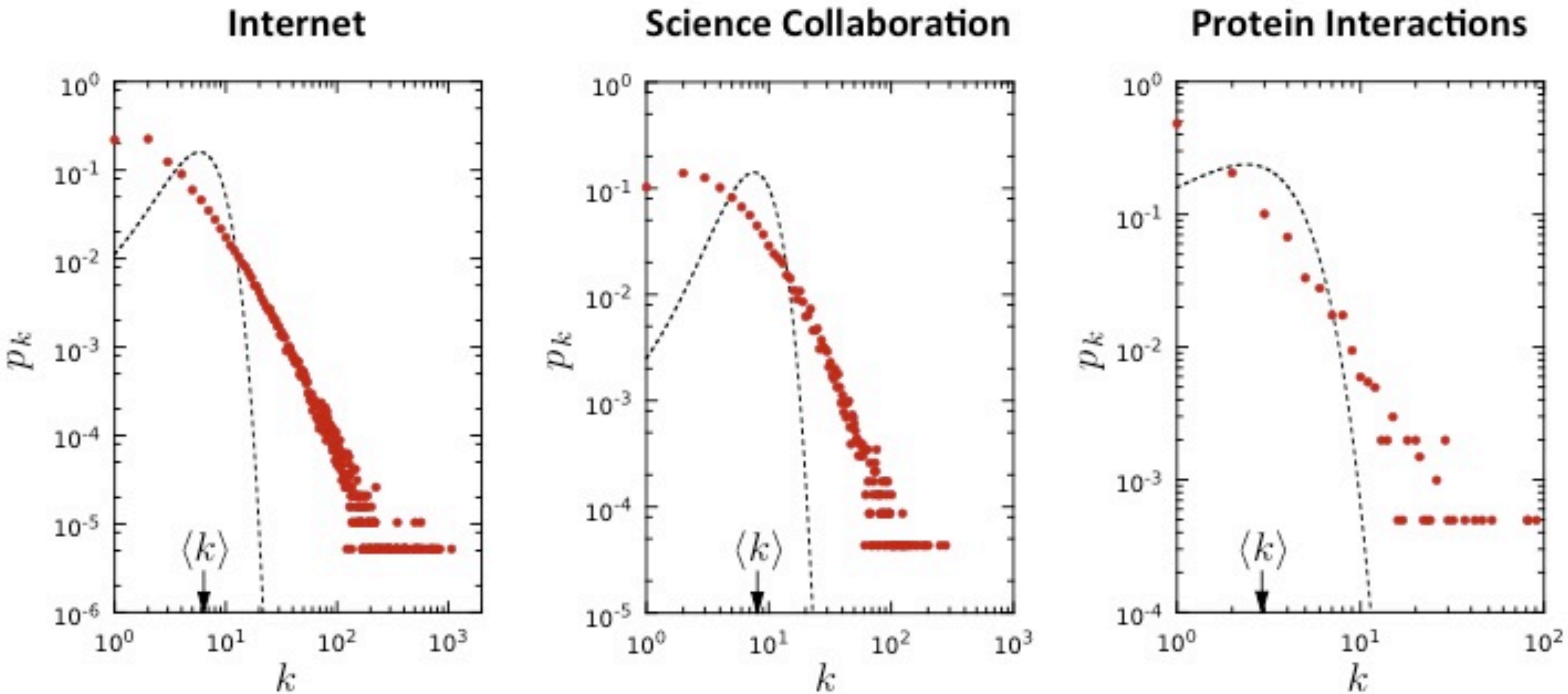
La probabilidad de encontrar un individuo con grado k>2,000 is 10^{-27}. Por lo tanto, la posibilidad de encontrar a un individuo con 2,000 conocidos es tan pequeña que tales nodos son virtualmente inexistentes en una sociedad aleatoria.

→una sociedad aleatoria consistiría principalmente en individuos promedio, con todos con aproximadamente el mismo número de amigos.

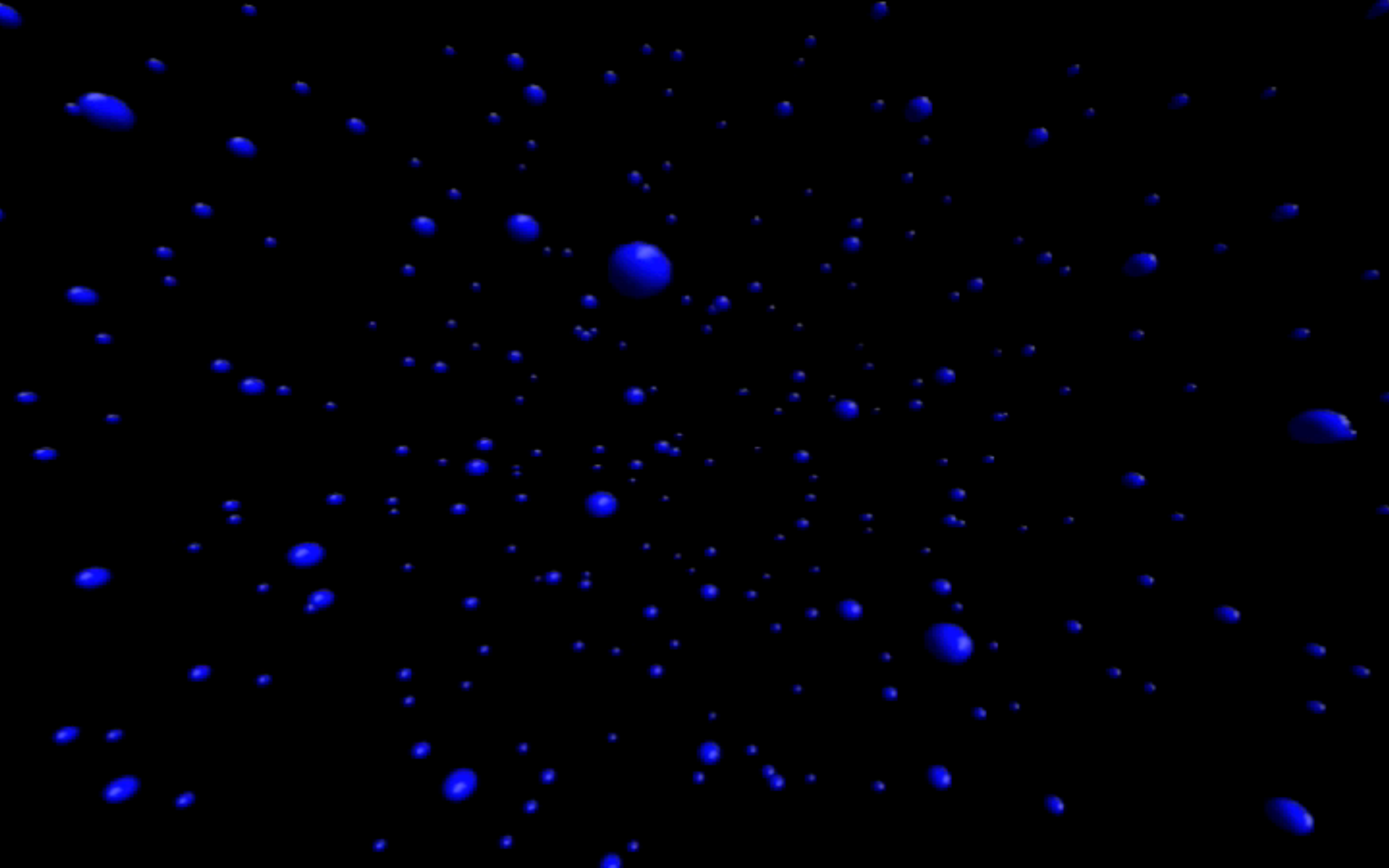
→Carecería de valores atípicos, individuos que son muy populares o solitarios.

ENFRENTANDO LA REALIDAD: Distribución en grado de redes reales.

$$P(k) = e^{-<k>} \frac{<k>^k}{k!}$$

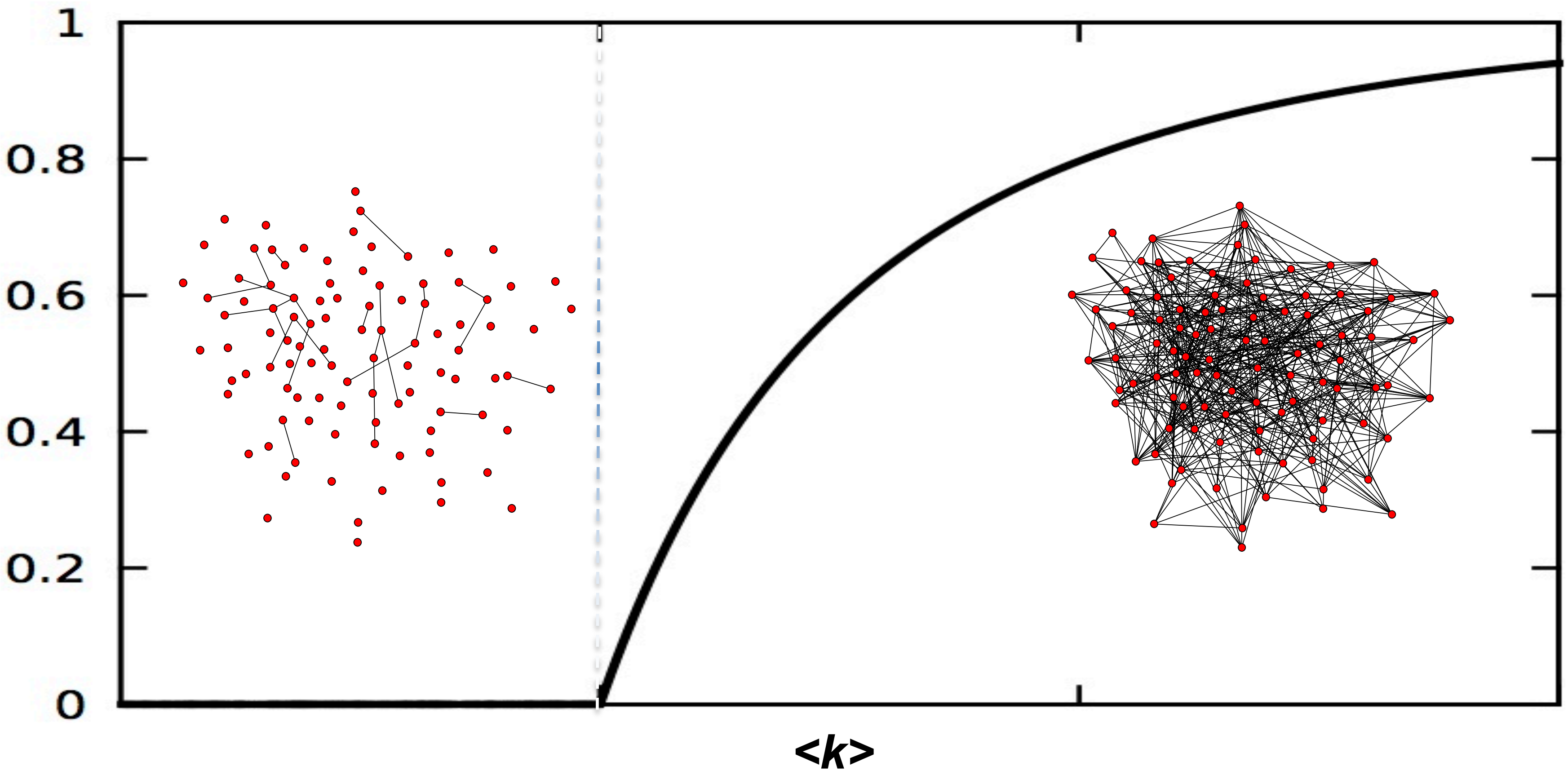


La evolución de una red aleatoria.



Evolución de una red aleatoria.

Nodos desconectados ➔ **Red.**



¿Cómo se produce esta transición?

Evolución de una red aleatoria.

Nodos desconectados → **Red.**

$<k_c>=1$ (Erdos y Renyi, 1959)

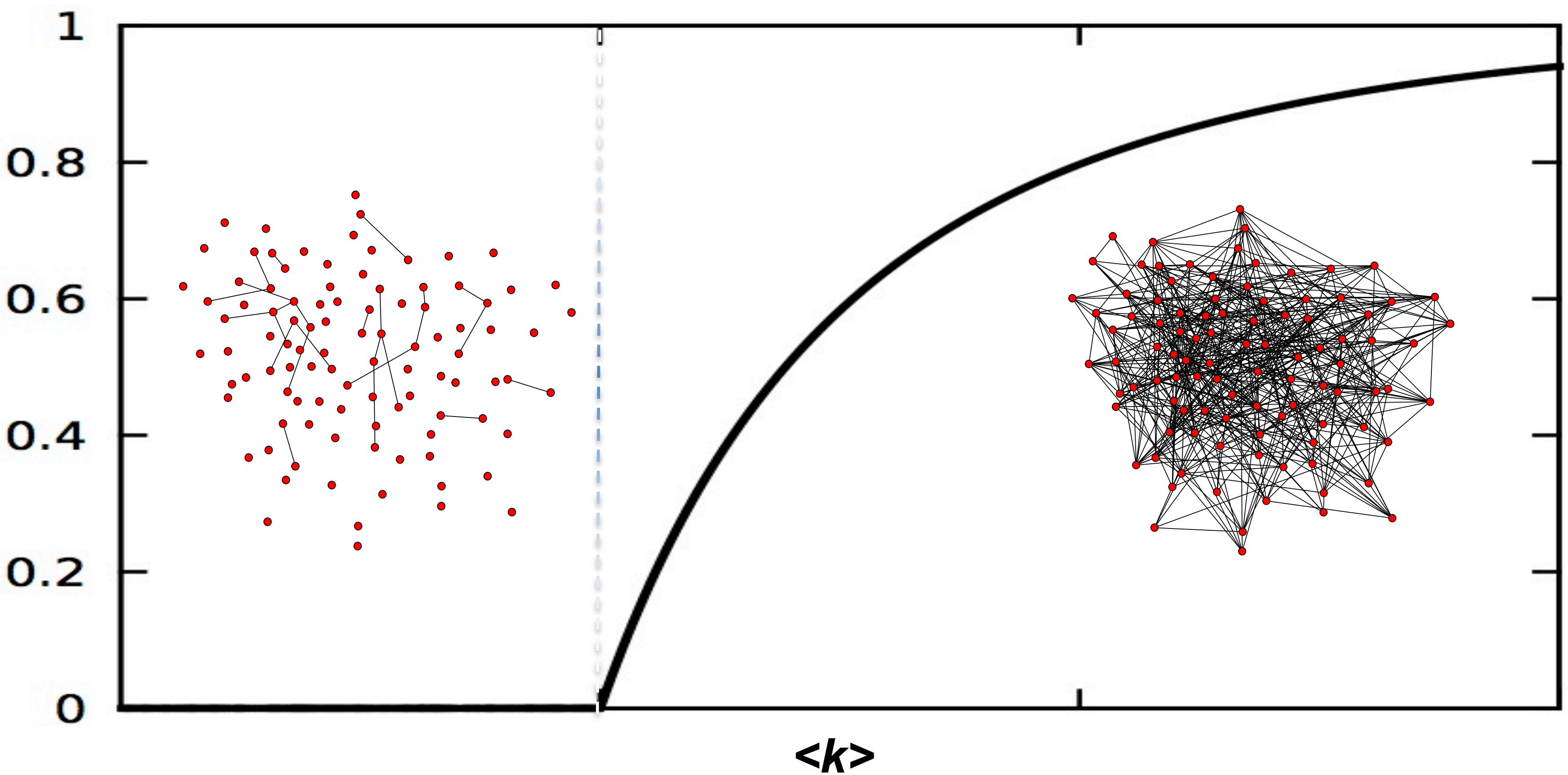
El hecho de que al menos un enlace por nodo sea ***necesario*** para tener un componente gigante no es inesperado. De hecho, para que exista un componente gigante, cada uno de sus nodos debe estar vinculado al menos a otro nodo.

Es algo inesperado, sin embargo, que un enlace es **suficiente** para el surgimiento de un componente gigante.

Es igualmente interesante que la aparición del componente gigante no sea gradual, pero sigue lo que los físicos llaman una **transición de fase** de segundo orden en $<k> = 1$.

Evolución de una red aleatoria.

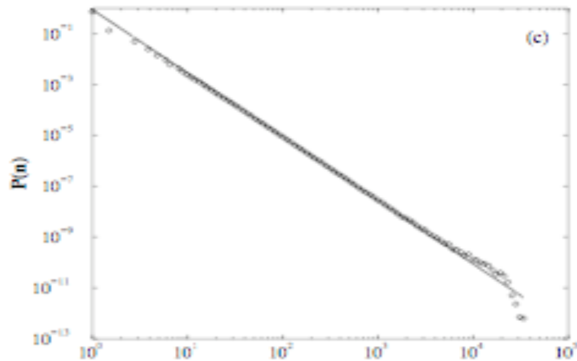
Nodos desconectados ➔ **Red.**



¿Cómo se produce esta transición?

DISTRIBUCION DEL TAMAÑO DEL GRUPO (CLUSTER)

Probabilidad de que un nodo seleccionado al azar pertenezca a un grupo de tamaño *s*:



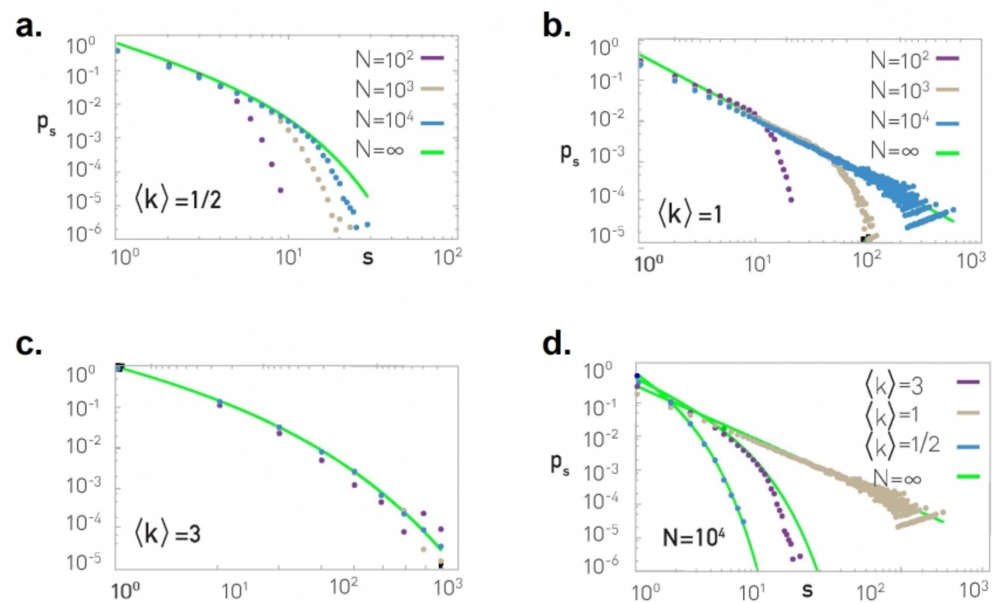
$$p(s) = \frac{e^{-<k>s}(<k>s)^{s-1}}{s!}$$

$$\langle k\rangle^{s-1} = \exp[(s-1)\ln\langle k\rangle]$$

$$p(s) = \frac{s^{s-1}}{s!} e^{-\langle k\rangle s+(s-1)\ln\langle k\rangle}$$

$$s! = \sqrt{2\pi s}\left(\frac{s}{e}\right)^{s}$$

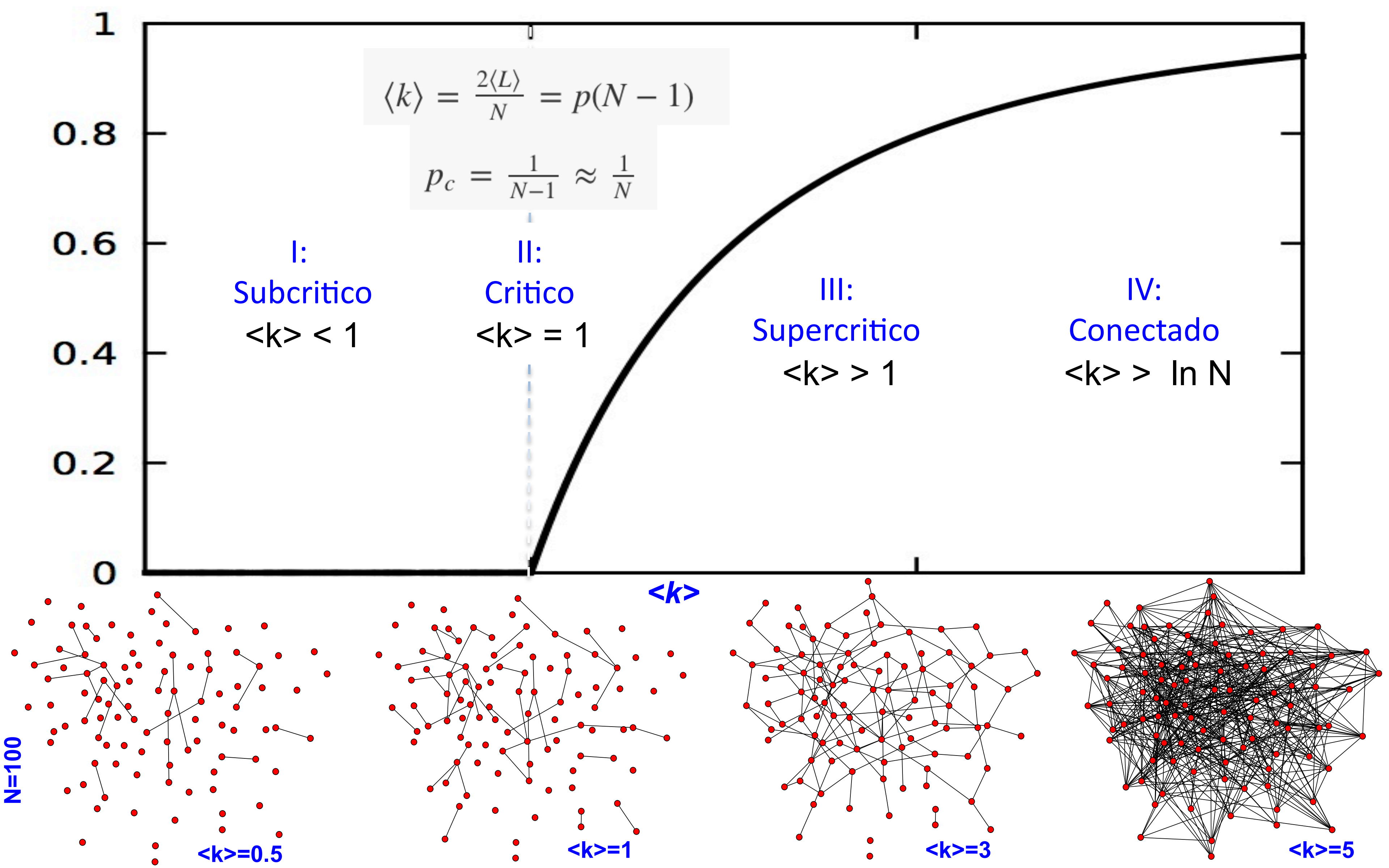
$$p(s) \sim s^{-3/2} e^{-(\langle k\rangle-1)s+(s-1)\ln\langle k\rangle}$$

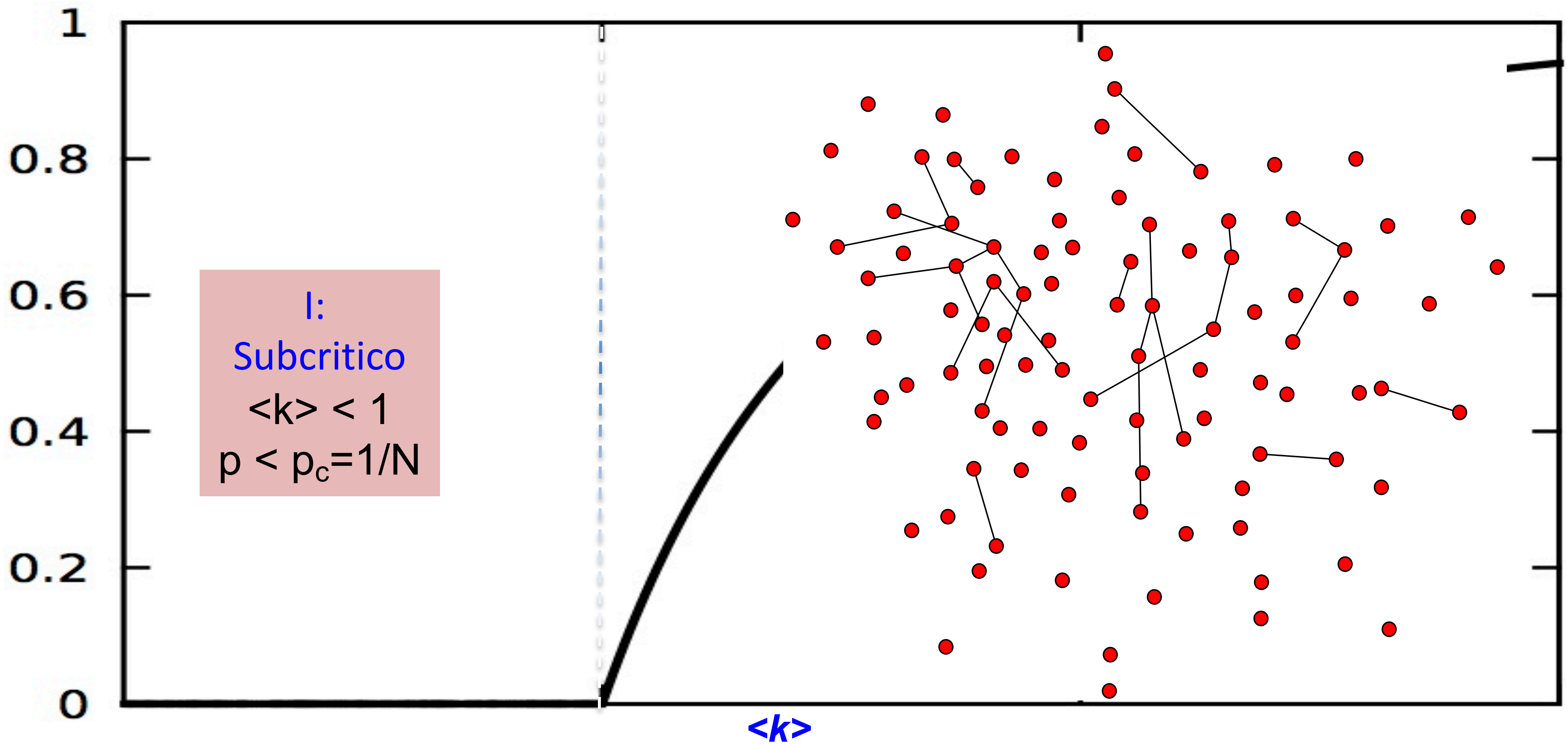


At the critical point <k>=1

$$p(s) \sim s^{-3/2}$$

Derivation in Newman, 2010

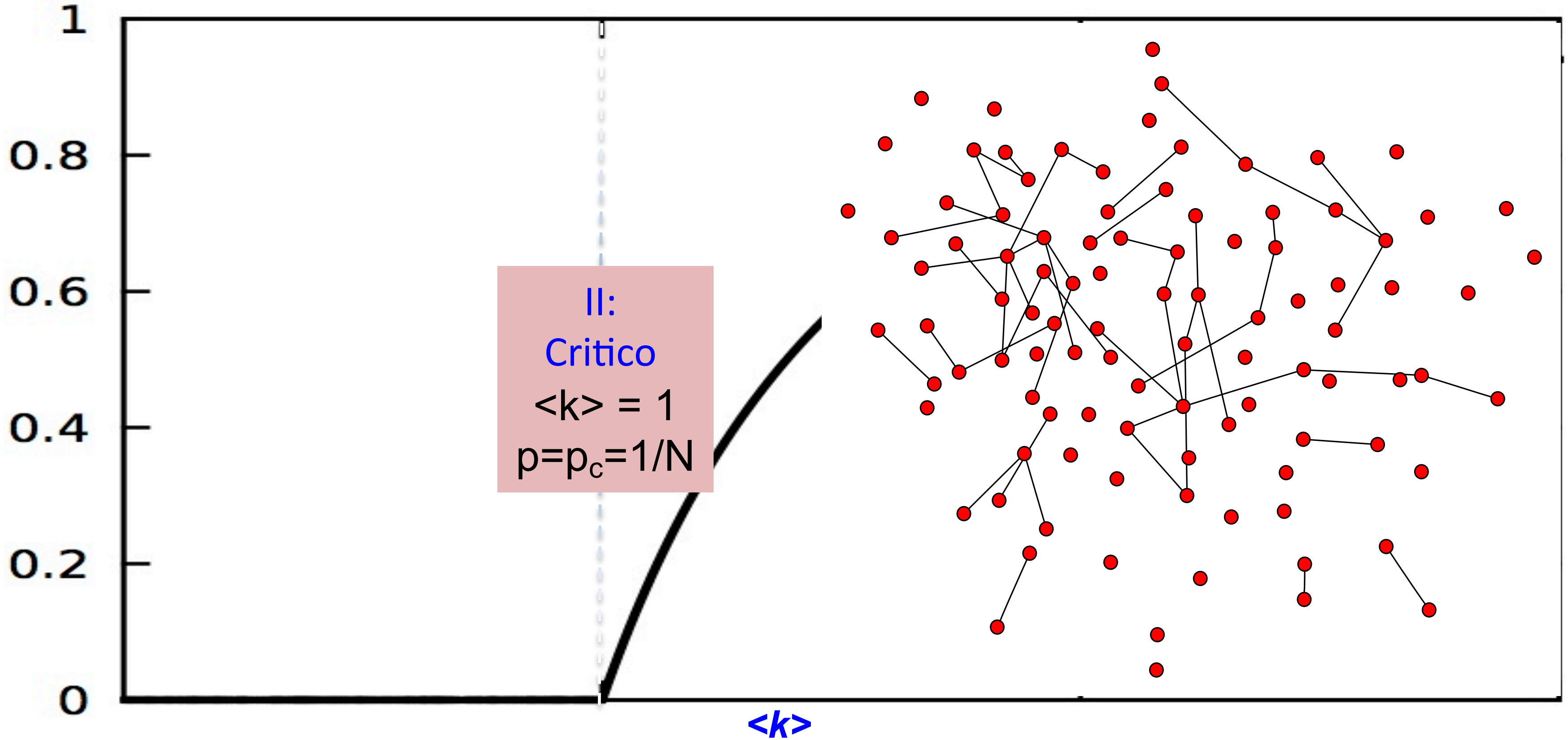

$\langle k \rangle = \frac{2\langle L \rangle}{N} = p(N-1)$
$p_c = \frac{1}{N-1} \approx \frac{1}{N}$
I:
Subcritico
<k> < 1
II:
Critico
<k> = 1
III:
Supercritico
<k> > 1
IV:
Conectado
<k> > ln N
1
0.8
0.6
0.4
0.2
0
<k>
N=100
<k>=0.5
<k>=1
<k>=3
<k>=5



No hay componente gigante

$$p(s) \sim s^{-3/2} e^{-(\langle k \rangle -1)s + (s-1)\ln\langle k \rangle}$$

N-L clusters aislados, la distribución del tamaño del cluster es exponencial
El cluster mas grande es un árbol, su tamaño es ~ *ln N*



Una sola componente gigante: $N_G \sim N^{2/3}$

→Contiene una decreciente fracción de todos los nodos, $N_G/N \sim N^{-1/3}$

→Componentes pequeños son árboles, GC tiene loops

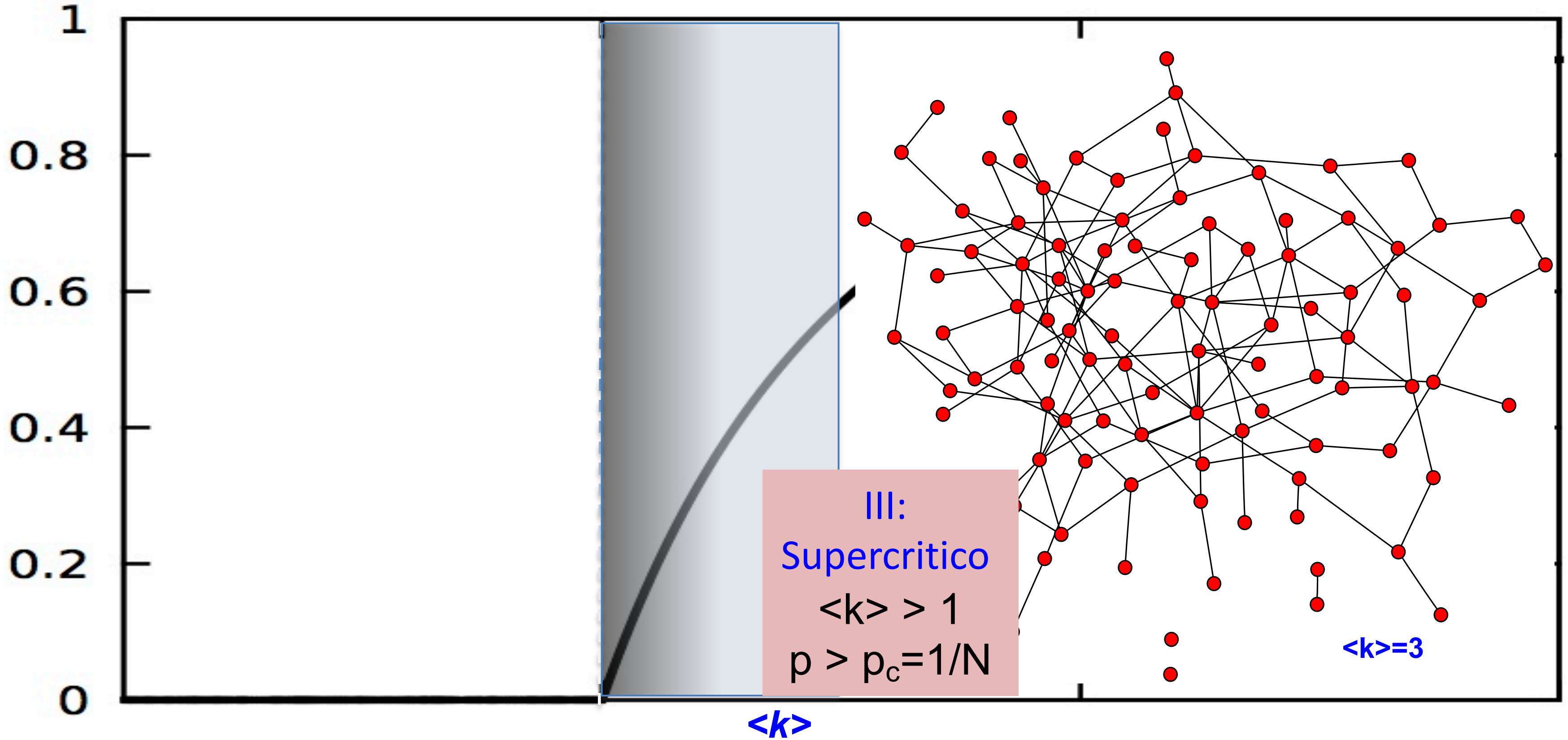
Distribución del tamaño de cluster: $p(s) \sim s^{-3/2}$

(Power-law/ley de potencia)

Un salto en el tamaño del cluster:

N=1,000 → ln N~ 6.9; $N^{2/3} \sim 95$

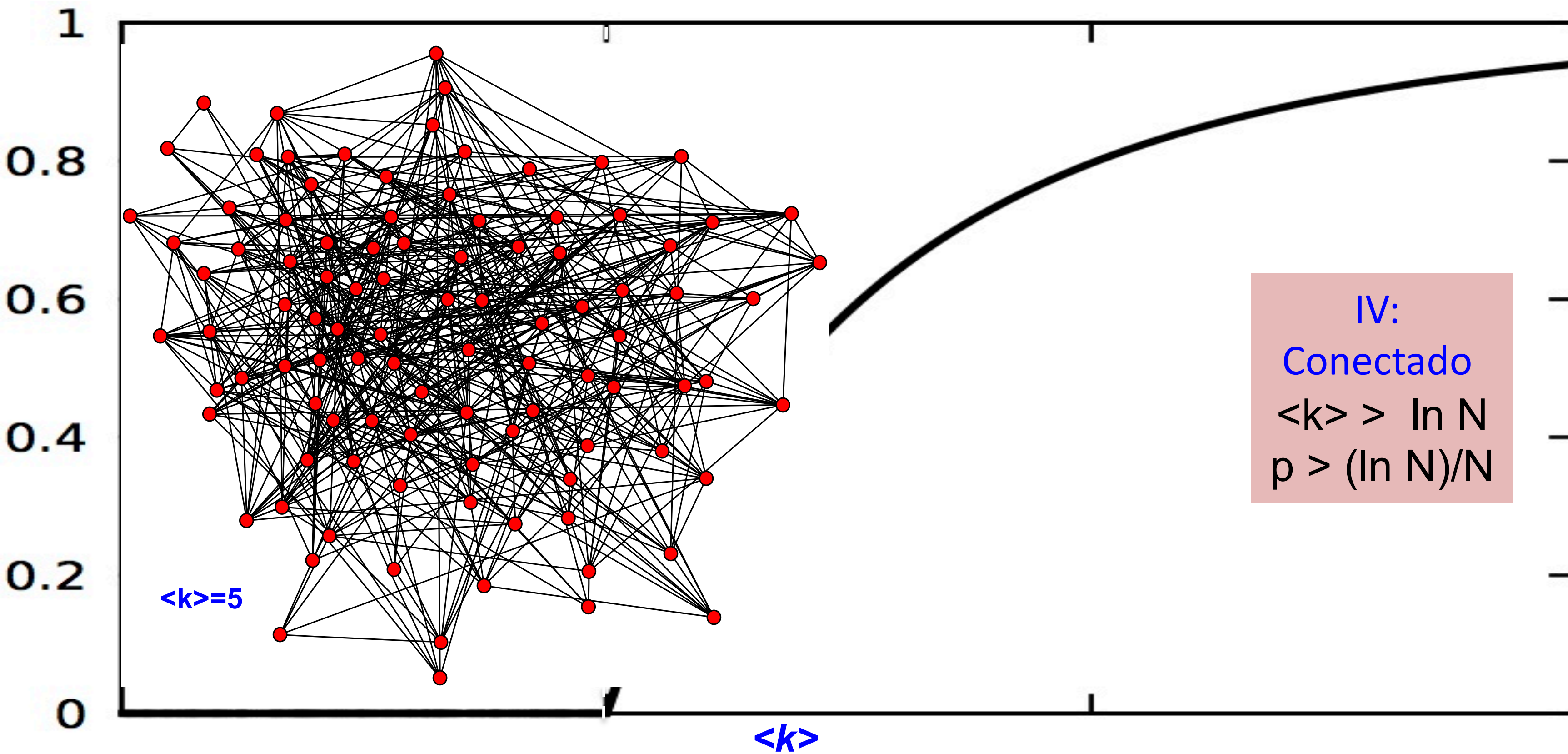
N=7 10^9 → ln N~ 22; $N^{2/3} \sim 3{,}659{,}250$



Una sola componente gigante: $N_G \sim (p-p_c)N$

→GC tiene loops.

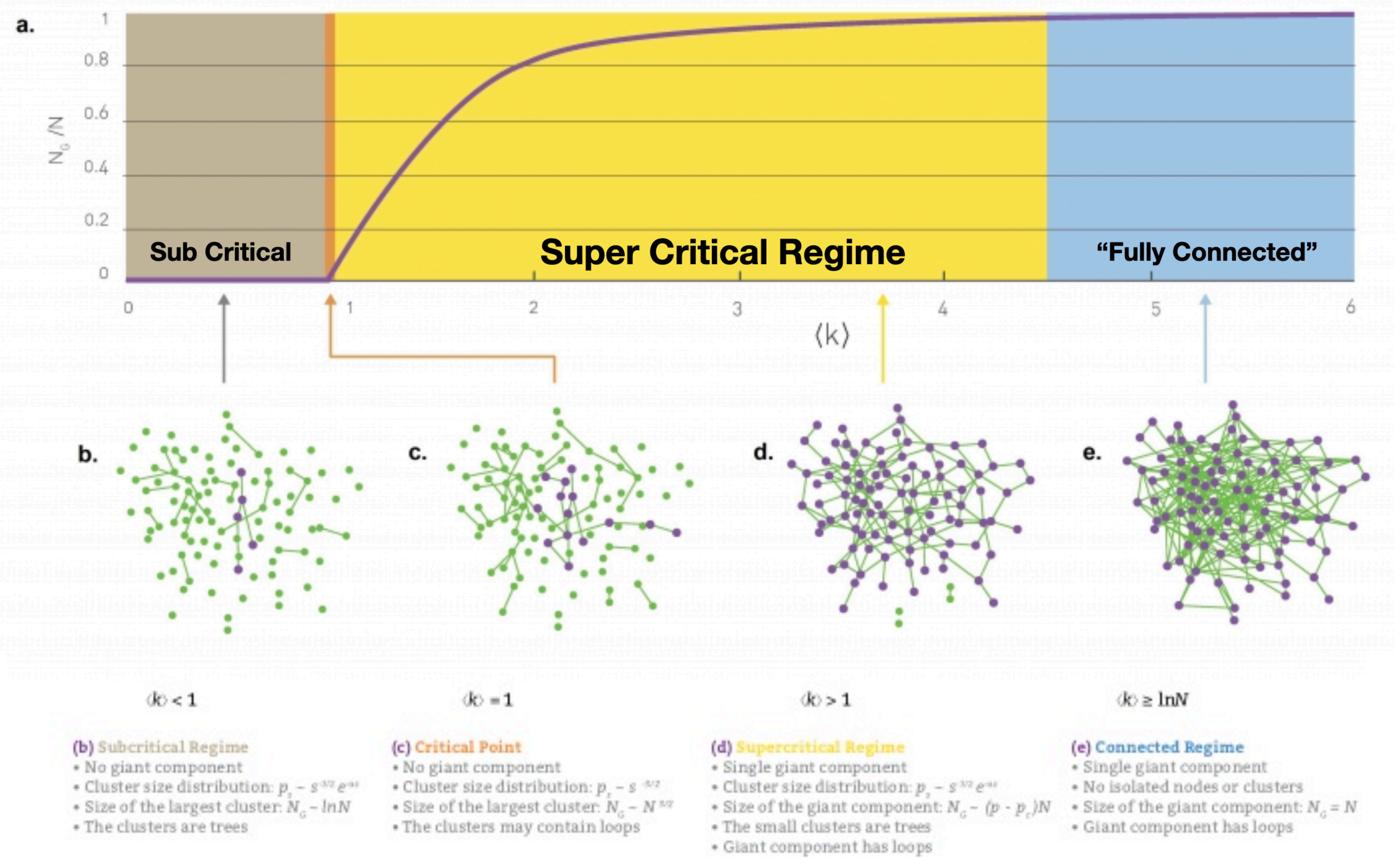
Distribución de tamaños de clusters: exponencial $p(s) \sim s^{-3/2} e^{-(\langle k \rangle - 1)s + (s-1)\ln\langle k \rangle}$



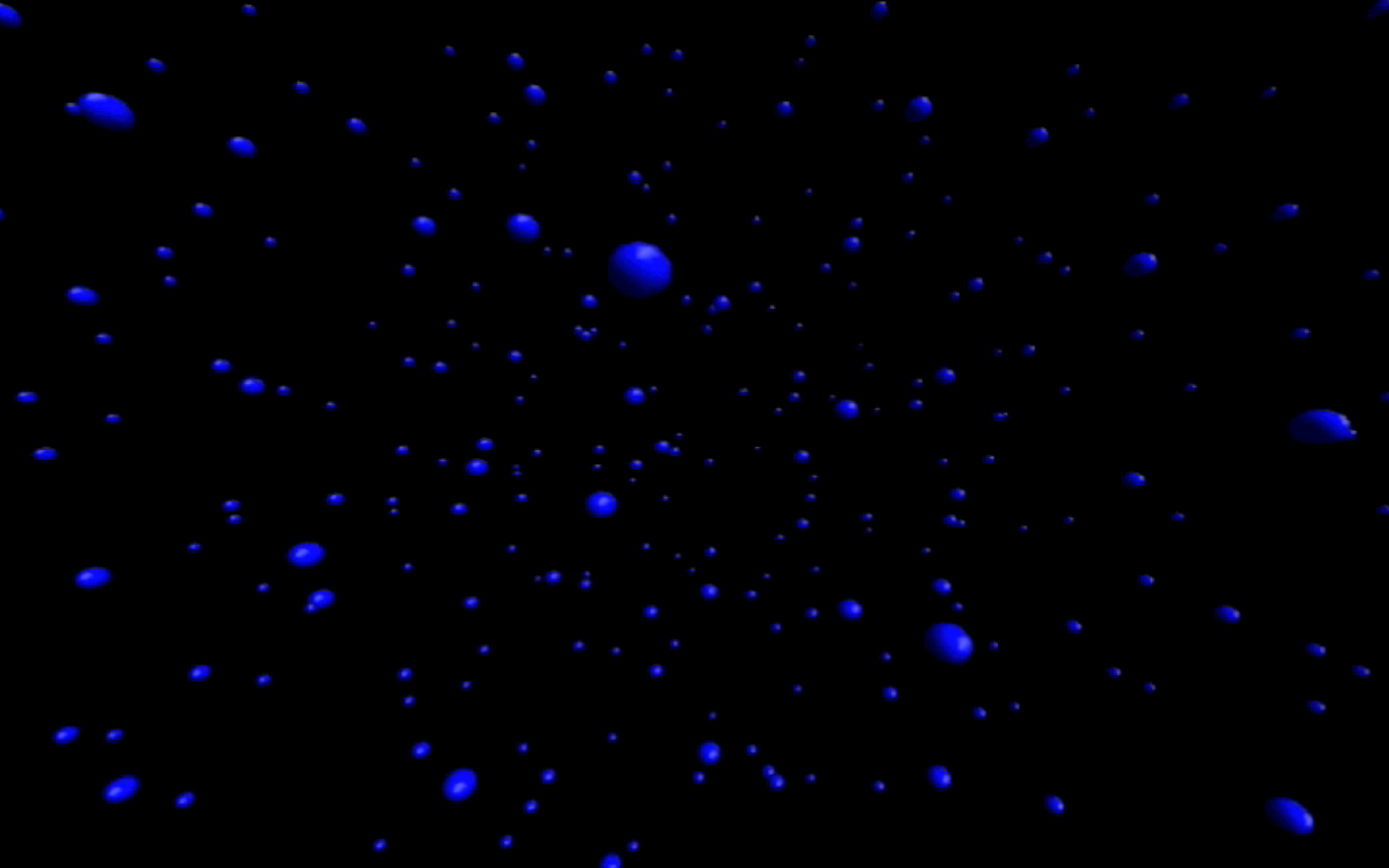
Solo un cluster: N_G=N

→GC es densa.

Distribución de tamaño de clusters: Ninguna

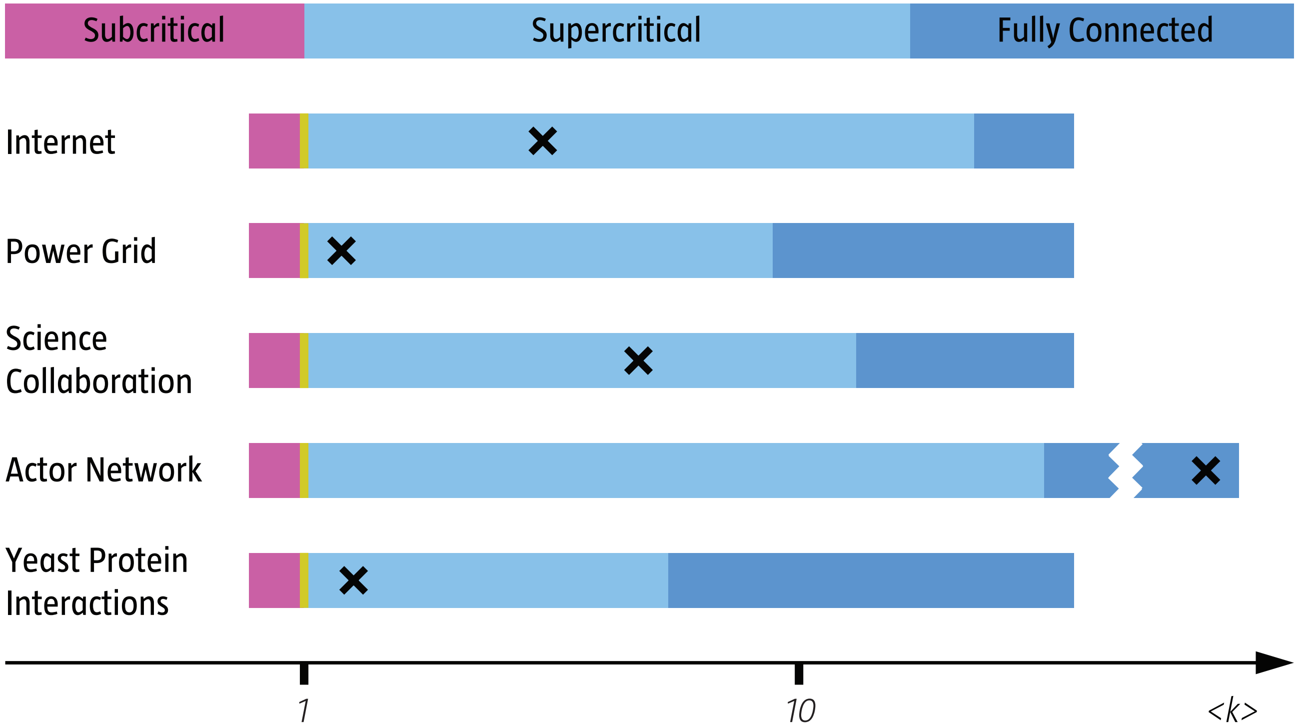


a.
1
0.8
0.6
0.4
0.2
0
N_G/N
Sub Critical
Super Critical Regime
"Fully Connected"
0
1
2
3
4
5
6
⟨k⟩
b.
c.
d.
e.
⟨k⟩ < 1
⟨k⟩ = 1
⟨k⟩ > 1
⟨k⟩ ≥ lnN
(b) Subcritical Regime
• No giant component
• Cluster size distribution: $p_s \sim s^{-3/2} e^{-as}$
• Size of the largest cluster: $N_G \sim \ln N$
• The clusters are trees
(c) Critical Point
• No giant component
• Cluster size distribution: $p_s \sim s^{-3/2}$
• Size of the largest cluster: $N_G \sim N^{2/3}$
• The clusters may contain loops
(d) Supercritical Regime
• Single giant component
• Cluster size distribution: $p_s \sim s^{-3/2} e^{-as}$
• Size of the giant component: $N_G \sim (p - p_c)N$
• The small clusters are trees
• Giant component has loops
(e) Connected Regime
• Single giant component
• No isolated nodes or clusters
• Size of the giant component: $N_G = N$
• Giant component has loops



Las redes reales son supercríticas.

Sección 7



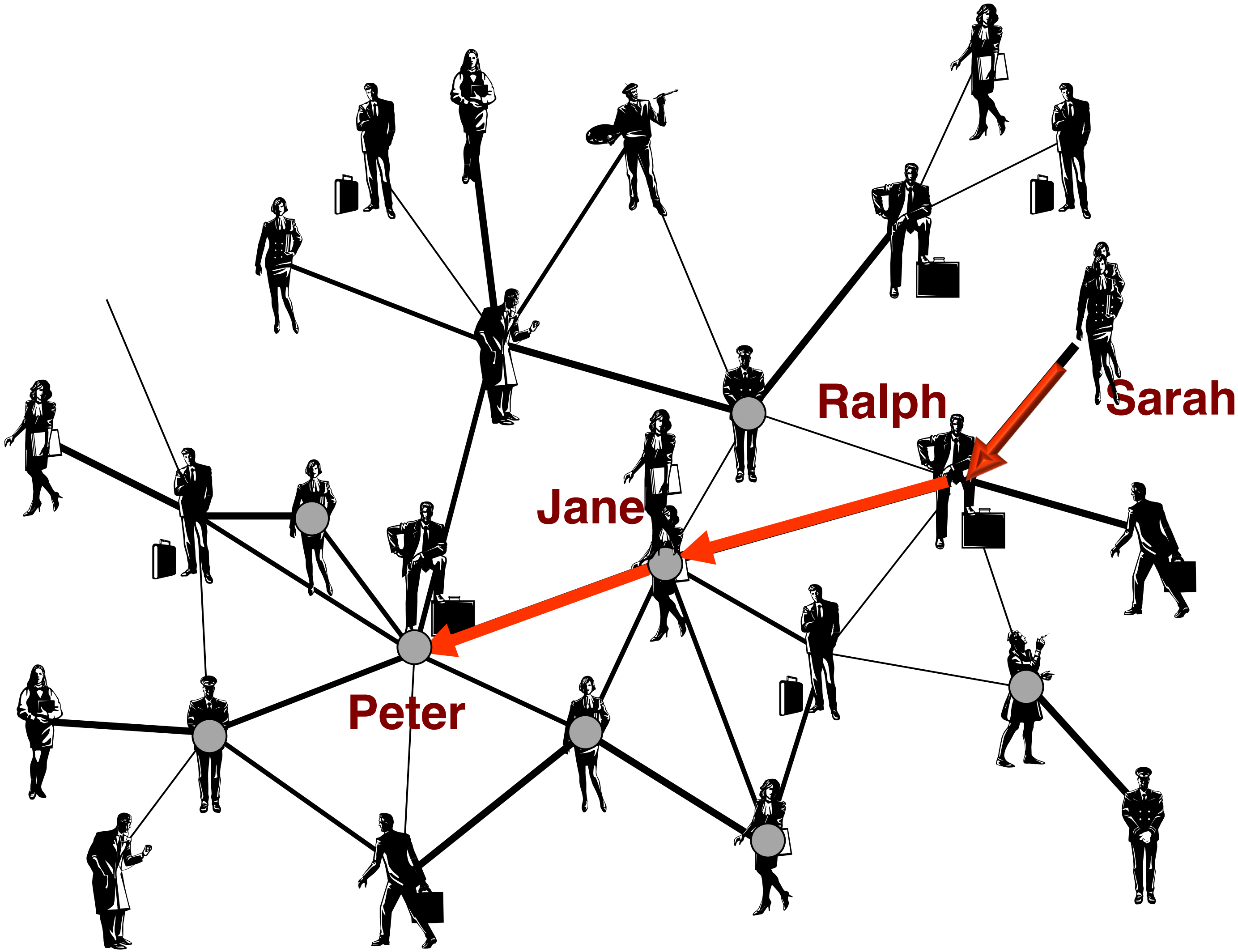
Network	N	L	$<k>$	$\ln N$
Internet	192,244	609,066	6.34	12.17
Power Grid	4,941	6,594	2.67	8.51
Science Collaboration	23,133	186,936	8.08	10.04
Actor Network	212,250	3,054,278	28.78	12.27
Yeast Protein Interactions	2,018	2,930	2.90	7.61

Red conectada

$<k> \ > \ \ln N$

Mundo Pequeño
(Small World)

SEIS GRADOS | small worlds



Frigyes Karinthy, 1929
Stanley Milgram, 1967

1929: Frigyes Kartinthy

Frigyes Karinthy (1887-1938)
Hungarian Writer

1929: *Minden másképpen van* (Everything is Different)
Láncszemek (Chains)

“Mire, Selma Lagerlöf acaba de ganar el Premio Nobel de Literatura, por lo que está obligada a conocer al Rey Gustav de Suecia, después de todo, él fue quien le entregó el Premio, como lo exige la tradición. El Rey Gustav, sin duda, es un apasionado jugador de tenis, que siempre participa en torneos internacionales. Se sabe que jugó con el Sr. Kehrling, a quien, por lo tanto, lo debe conocer con seguridad, y sucede que yo mismo conozco bastante bien al Sr. Kehrling ".

"El trabajador conoce al gerente de la tienda, quién conoce a Ford; Ford tiene una relación amistosa con el director general de Hearst Publications, que el año pasado se hizo muy amigo de Arpad Pasztor, alguien que no solo conozco, sino que es un buen amigo mío. Así que fácilmente podría pedirle que envíe un telegrama a través del director general para que le diga a Ford que debe hablar con el gerente para que le informe al trabajador de la tienda que arme un auto rápidamente para mí, ya que necesito uno."

SEIS GRADOS. 1967: Stanley Milgram

CÓMO PARTICIPAR EN ESTE ESTUDIO

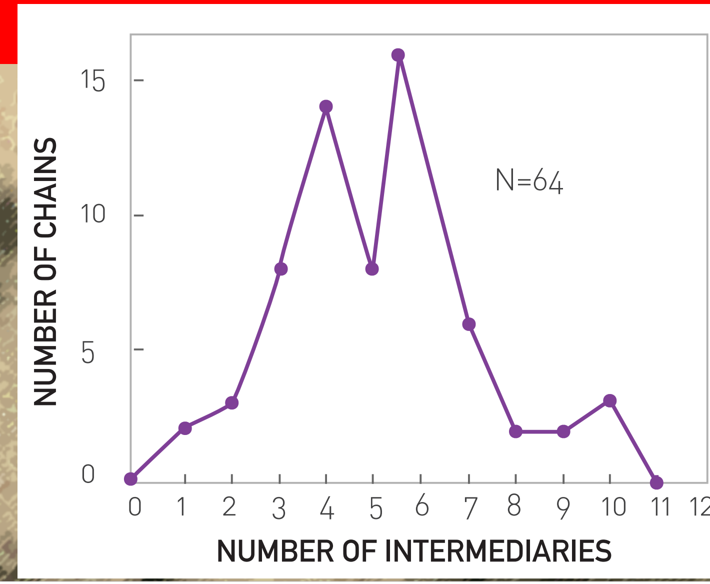
1. AGREGUE SU NOMBRE A LA LISTA EN LA PARTE INFERIOR DE ESTA HOJA, para que la próxima persona que reciba esta carta sepa de quién proviene.

2. SAQUE UNA POSTAL. LLENARLA Y REGRESARLA A HARVARD UNIVERSITY. No se necesita sello. La postal es muy importante. Nos permite realizar un seguimiento del progreso de la carpeta a medida que avanza hacia la persona objetivo.

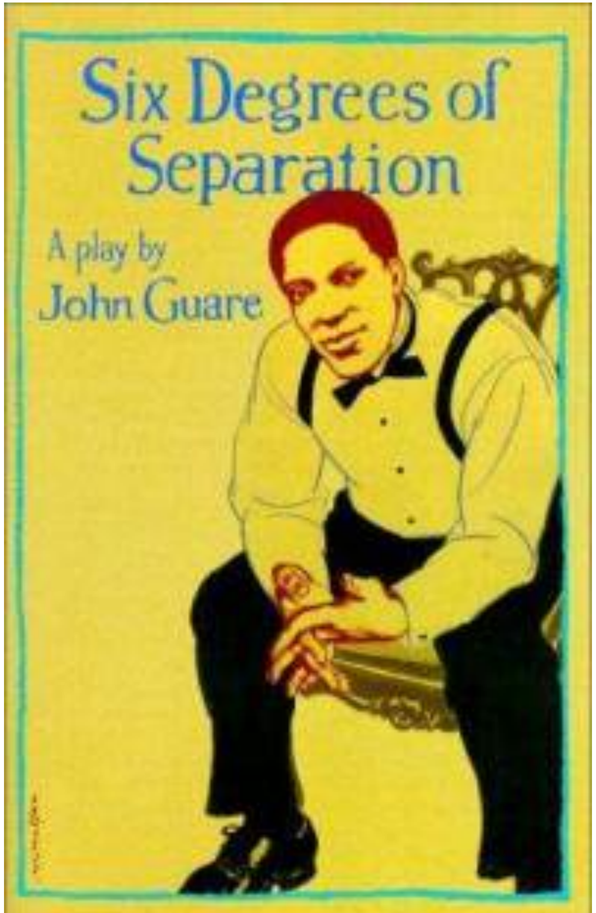
3. SI CONOCES LA PERSONA OBJETIVO A NIVEL PERSONAL, ENVÍELA ESTA CARPETA DIRECTAMENTE A ÉL (ELLA). Haga esto solo si ha conocido previamente a la persona objetivo y se conocen entre sí por su nombre.

4. SI NO CONOCES A LA PERSONA OBJETIVO DE MANERA PERSONAL, NO INTENTES CONTACTARLO DIRECTAMENTE. EN SU LUGAR, ENVÍA ESTA CARPETA (TARJETAS POSTALES Y TODO) A UN CONOCIDO PERSONAL PUEDA CONOCER LA PERSONA OBJETIVO. Puede enviar la carpeta a un amigo, familiar o conocido, pero debe ser alguien que conozca por su nombre de pila.

SEIS GRADOS | 1967: Stanley Milgram



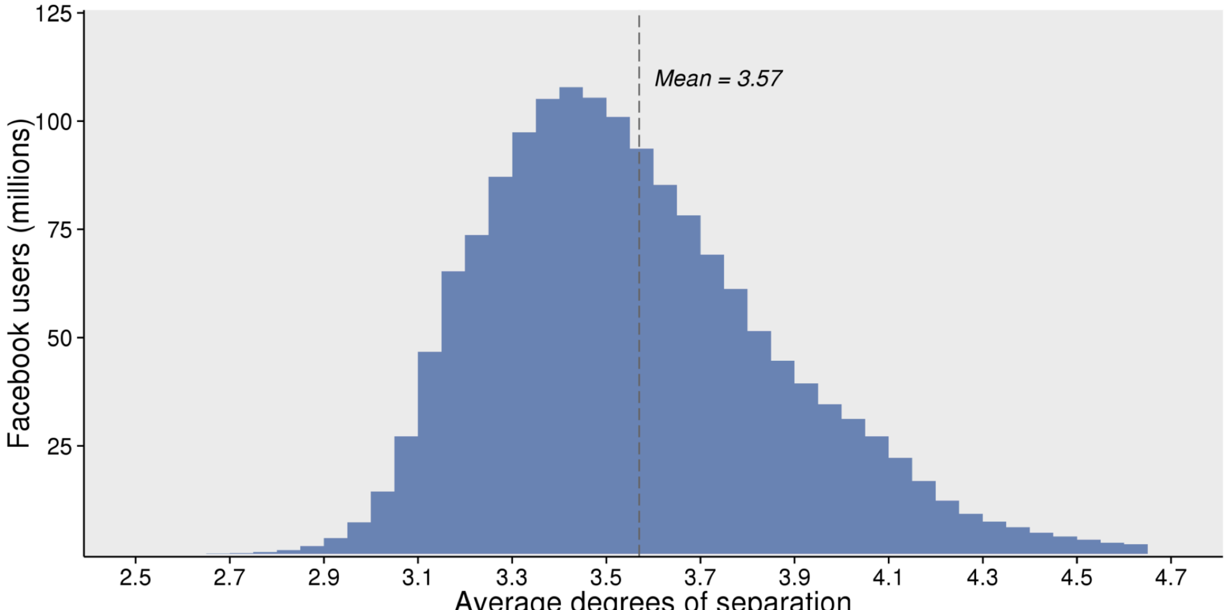
1991: John Guare



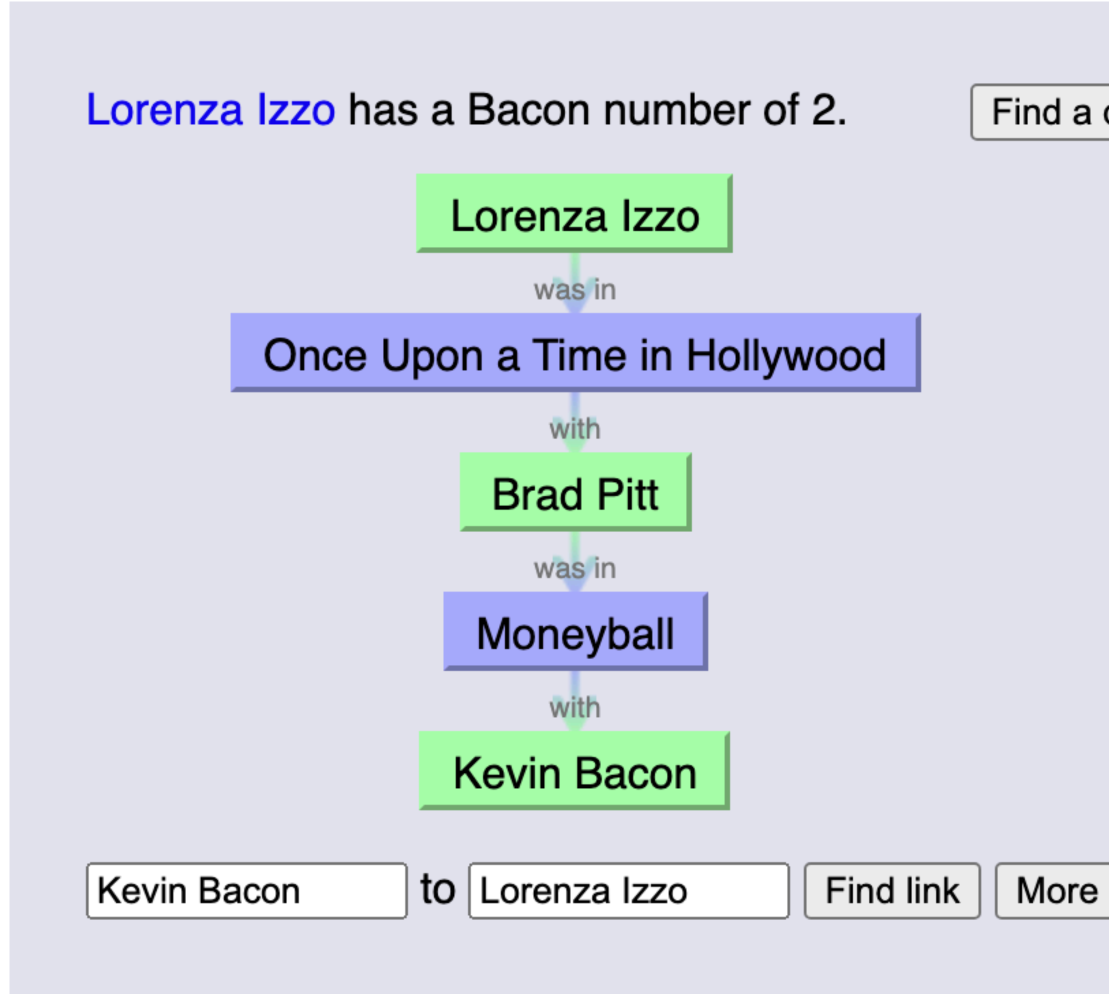
"Todos en este planeta están separados solo por otras seis personas. Seis grados de separación. Entre nosotros y todos los demás en este planeta. El presidente de los Estados Unidos. Un gondolero en Venecia ... No son solo los grandes nombres. Es cualquiera. Un nativo en una selva tropical. Un Tierra del Fuegano. Un esquimal. Estoy atado a todos en este planeta por un rastro de seis personas. Es un pensamiento profundo. Cómo cada persona es una puerta nueva, abriéndose a otros mundos ".

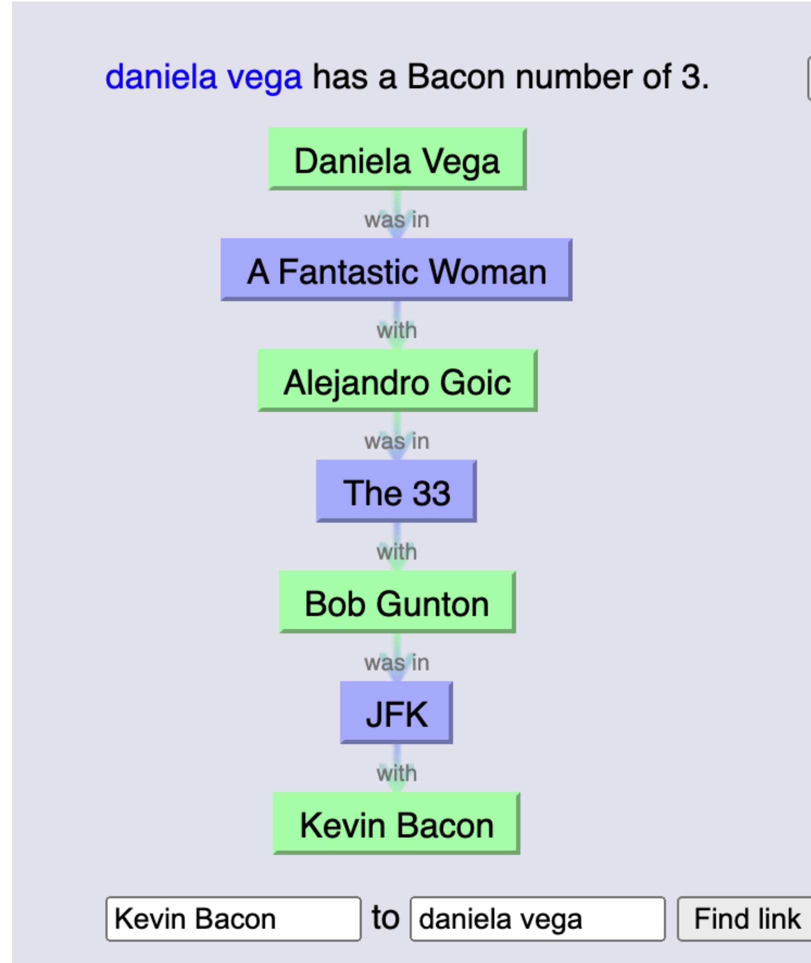
SEIS GRADOS | Experimento 2.0

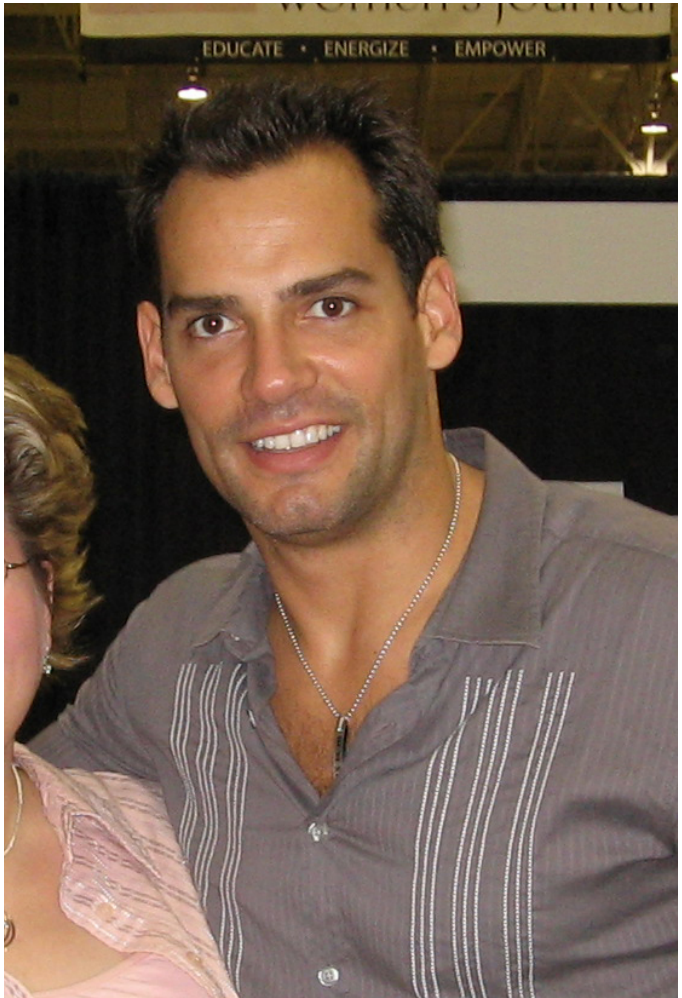
—Informado en el blog de investigación de Facebook en 2016;
—La mayoría de las personas en Facebook tienen promedios de entre 2,9 y 4,2 grados de separación;
—La figura muestra los grados promedio estimados de separación entre todas las personas en Facebook. La persona promedio está conectada con todas las demás personas en un promedio de 3.57 pasos. La mayoría de personas tiene un promedio de entre 3 y 4 pasos;

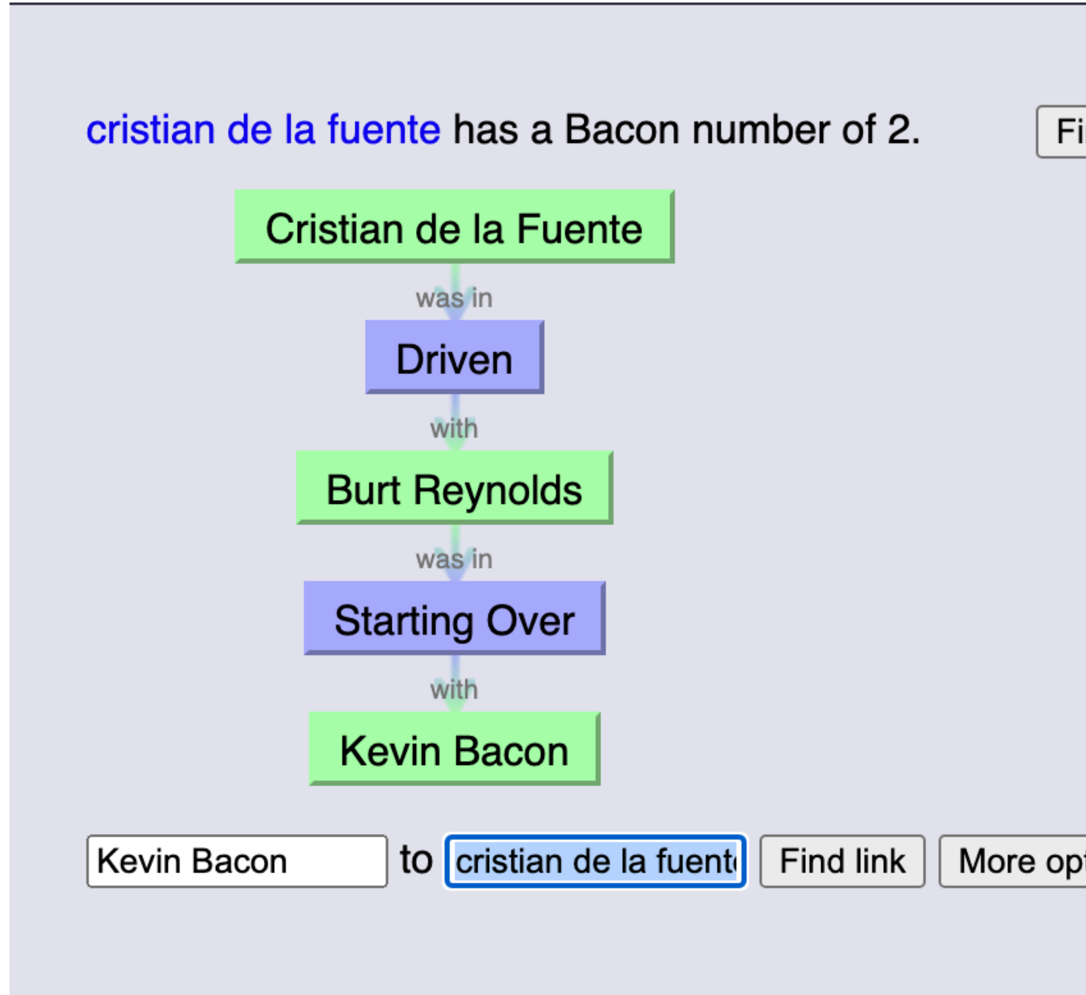


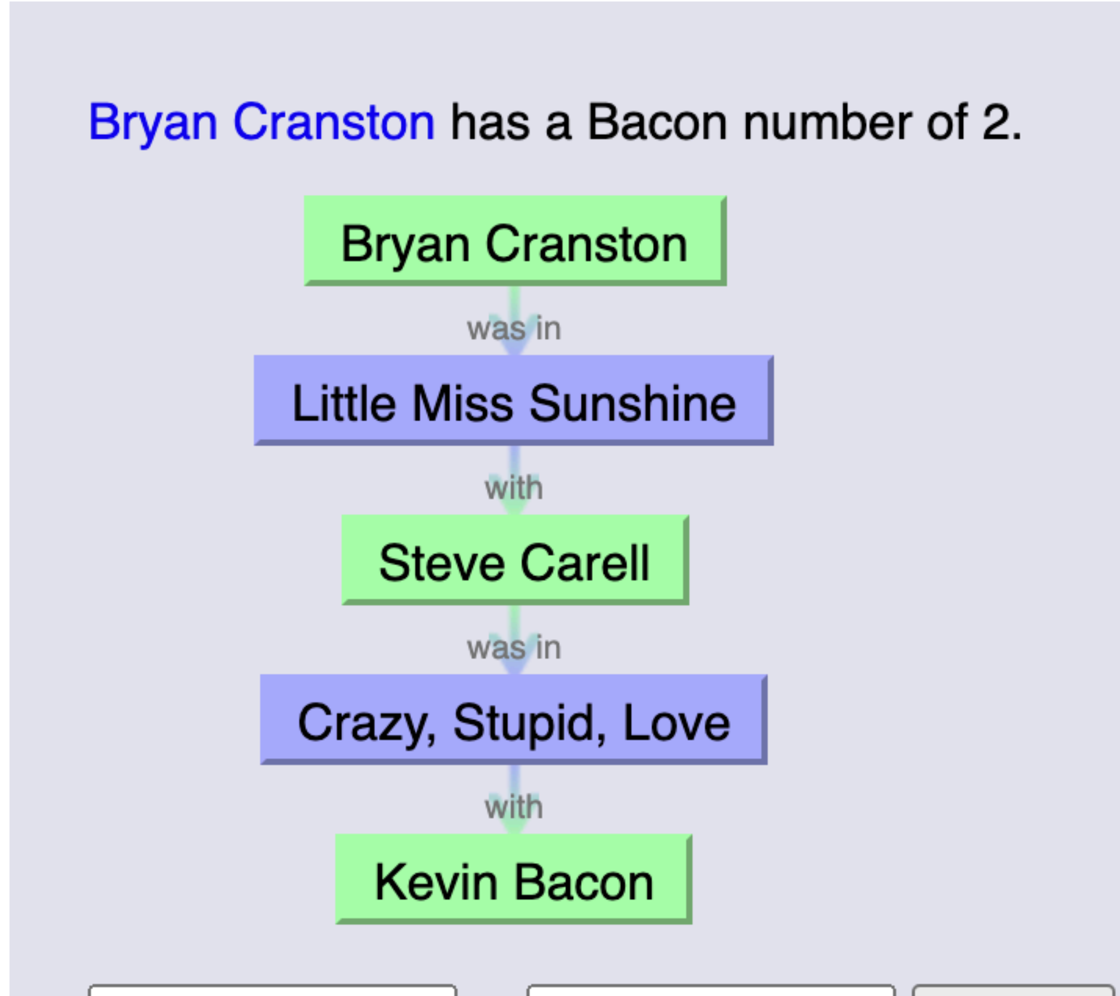
https://research.fb.com/blog/2016/02/three-and-a-half-degrees-of-separation/


Lorenza Izzo has a Bacon number of 2.
Find a di
Lorenza Izzo
was in
Once Upon a Time in Hollywood
with
Brad Pitt
was in
Moneyball
with
Kevin Bacon
Kevin Bacon
to
Lorenza Izzo
Find link
More o


daniela vega has a Bacon number of 3.
Daniela Vega
was in
A Fantastic Woman
with
Alejandro Goic
was in
The 33
with
Bob Gunton
was in
JFK
with
Kevin Bacon
Kevin Bacon
to
daniela vega
Find link

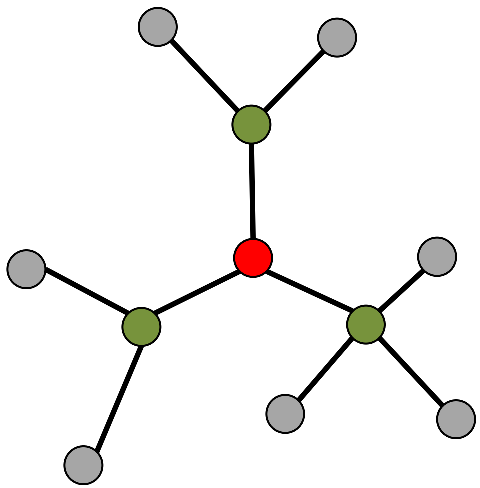

EDUCATE • ENERGIZE • EMPOWER


cristian de la fuente has a Bacon number of 2.
Cristian de la Fuente
was in
Driven
with
Burt Reynolds
was in
Starting Over
with
Kevin Bacon
Kevin Bacon
to
cristian de la fuent
Find link


Bryan Cranston has a Bacon number of 2.
Bryan Cranston
was in
Little Miss Sunshine
with
Steve Carell
was in
Crazy, Stupid, Love
with
Kevin Bacon

DISTANCIA EN GRAFOS ALEATORIOS

Los grafos aleatorios tienden a tener una topología en forma de árbol con grados de nodo casi constantes.



$<k>$ nodes at distance one (d=1).

$<k>^2$ nodes at distance two (d=2).

$<k>^3$ nodes at distance three (d =3).

...

$<k>^d$ nodes at distance d.

$$N = 1 + \langle k \rangle + \langle k \rangle^2 + \ldots + \langle k \rangle^{d_{\max}} = \frac{\langle k \rangle^{d_{\max}+1} - 1}{\langle k \rangle - 1} \approx \langle k \rangle^{d_{\max}} \Rightarrow \qquad d_{\max} = \frac{\log N}{\log \langle k \rangle}$$

DISTANCIA EN GRAFOS ALEATORIOS

$$d_{\max} = \frac{\log N}{\log\langle k\rangle}$$

En la mayoría de las redes, esto ofrece una mejor aproximación a la distancia promedio entre dos nodos elegidos al azar, ⟨d⟩, que d_{max} .

$$< d >= \frac{\log N}{\log\langle k\rangle}$$

Llamaremos a este fenómeno de mundo pequeño la propiedad de que la longitud promedio de la trayectoria o el diámetro depende logarítmicamente del tamaño del sistema. Por lo tanto, "pequeño" significa que ⟨d⟩ es proporcional al logaritmo de N, en lugar de a N.

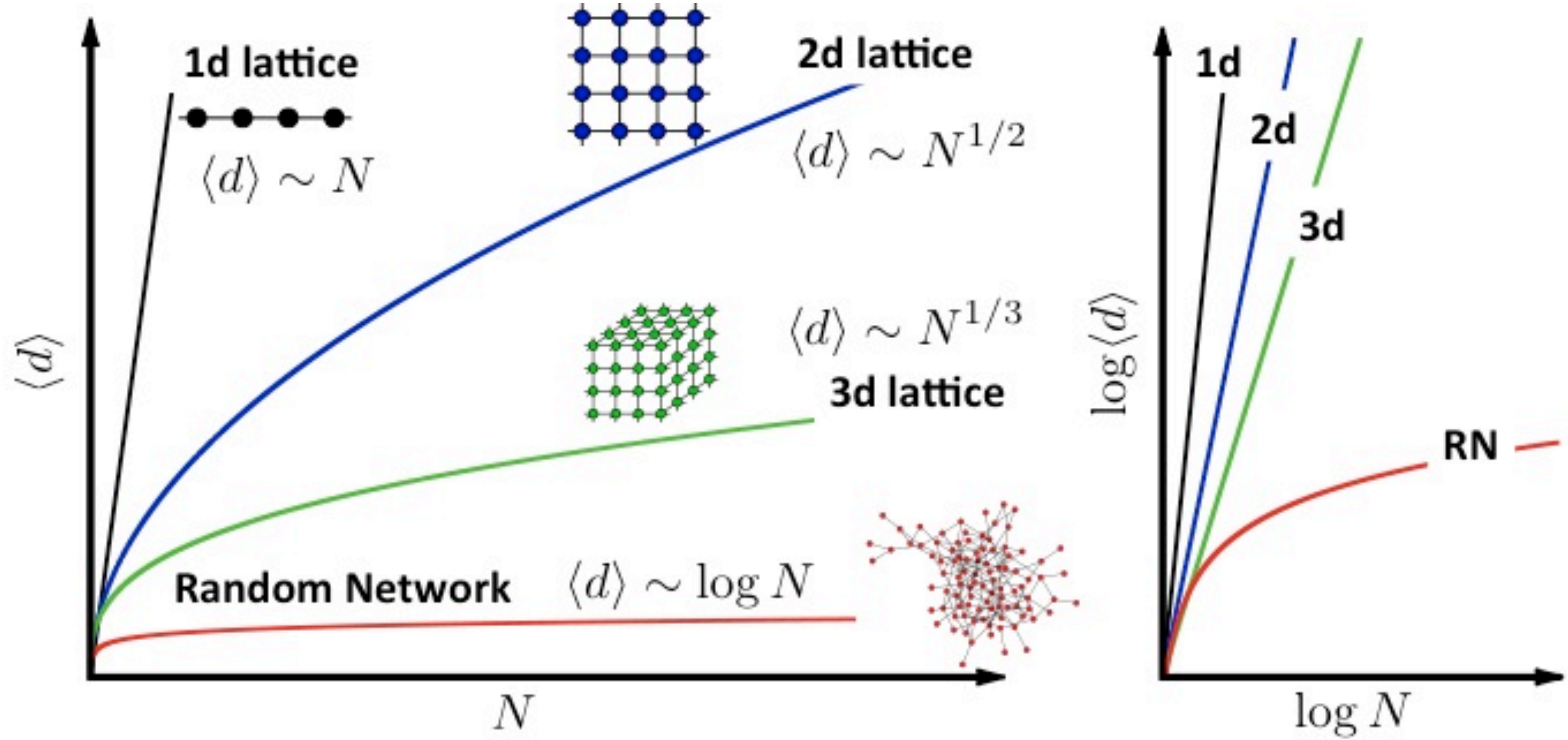
El término 1/log⟨k⟩ implica cuanto más densa sea la red, menor será la distancia entre los nodos.

DISTANCIAS EN GRÁFOS ALEATORIOS comparados con datos reales

Network	N	L	$\langle k \rangle$	$\langle d \rangle$	d_{max}	$\ln N / \ln \langle k \rangle$
Internet	192,244	609,066	6.34	6.98	26	6.58
WWW	325,729	1,497,134	4.60	11.27	93	8.31
Power Grid	4,941	6,594	2.67	18.99	46	8.66
Mobile-Phone Calls	36,595	91,826	2.51	11.72	39	11.42
Email	57,194	103,731	1.81	5.88	18	18.4
Science Collaboration	23,133	93,437	8.08	5.35	15	4.81
Actor Network	702,388	29,397,908	83.71	3.91	14	3.04
Citation Network	449,673	4,707,958	10.43	11.21	42	5.55
E. Coli Metabolism	1,039	5,802	5.58	2.98	8	4.04
Protein Interactions	2,018	2,930	2.90	5.61	14	7.14

Dadas las enormes diferencias en alcance, tamaño y grado promedio, la aproximación es excelente.

¿Por qué sorprenden los pequeños mundos? Sorprendente en comparación con qué?



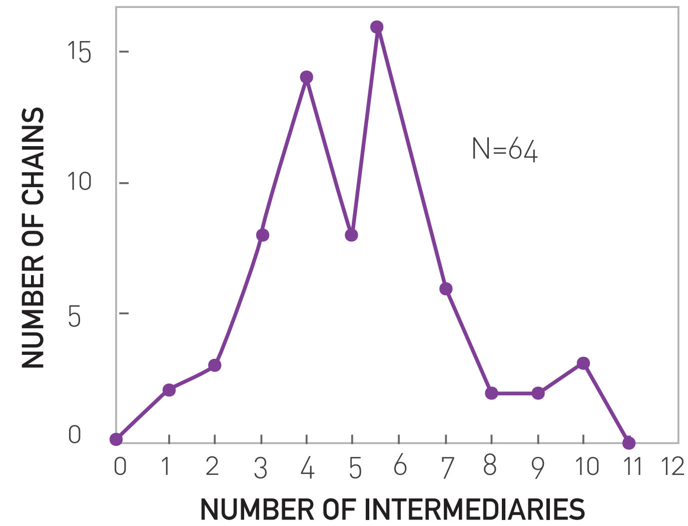
Tres, cuatro o seis grados?

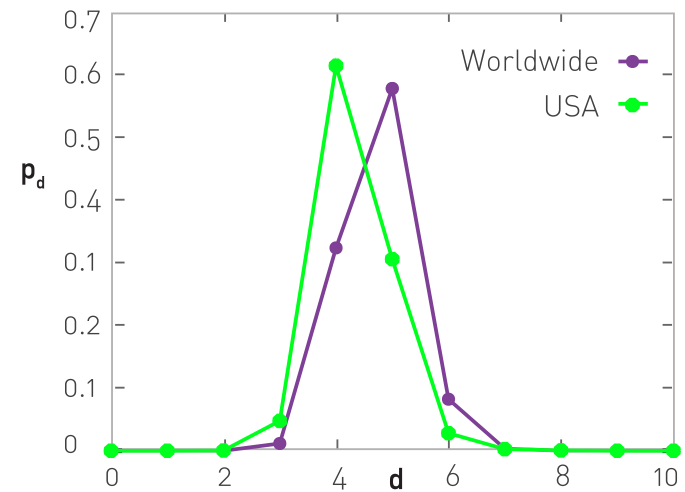
Para las redes sociales del mundo:

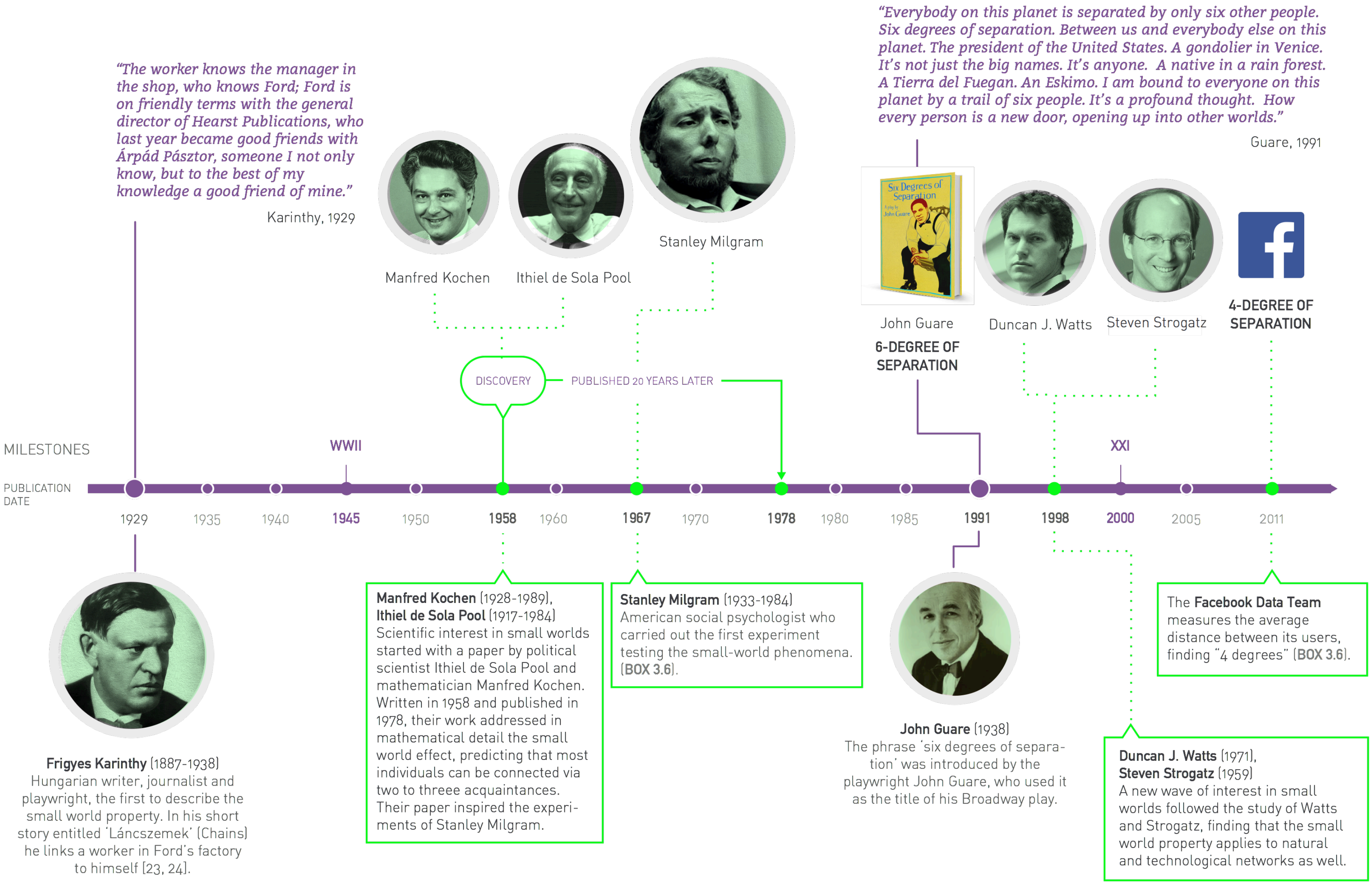
$\langle k \rangle \simeq 10^3$

$N \simeq 7 \times 10^9$ para la población mundial.

$$< d > = \frac{\ln(N)}{\ln\langle k \rangle} = 3.28$$



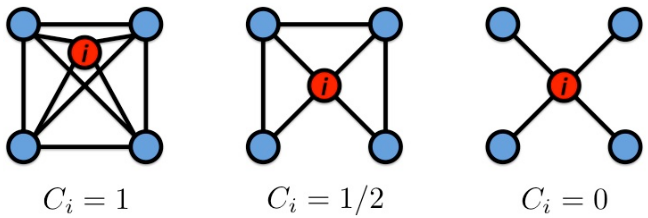



"The worker knows the manager in the shop, who knows Ford; Ford is on friendly terms with the general director of Hearst Publications, who last year became good friends with Árpád Pásztor, someone I not only know, but to the best of my knowledge a good friend of mine."
Karinthy, 1929
Manfred Kochen
Ithiel de Sola Pool
Stanley Milgram
"Everybody on this planet is separated by only six other people. Six degrees of separation. Between us and everybody else on this planet. The president of the United States. A gondolier in Venice. It's not just the big names. It's anyone. A native in a rain forest. A Tierra del Fuegan. An Eskimo. I am bound to everyone on this planet by a trail of six people. It's a profound thought. How every person is a new door, opening up into other worlds."
Guare, 1991
Six Degrees of Separation
John Guare
6-DEGREE OF SEPARATION
Duncan J. Watts
Steven Strogatz
4-DEGREE OF SEPARATION
DISCOVERY
PUBLISHED 20 YEARS LATER
MILESTONES
WWII
XXI
PUBLICATION DATE
1929
1935
1940
1945
1950
1958
1960
1967
1970
1978
1980
1985
1991
1998
2000
2005
2011
Frigyes Karinthy (1887-1938)
Hungarian writer, journalist and playwright, the first to describe the small world property. In his short story entitled 'Láncszemek' (Chains) he links a worker in Ford's factory to himself [23, 24].
Manfred Kochen (1928-1989),
Ithiel de Sola Pool (1917-1984)
Scientific interest in small worlds started with a paper by political scientist Ithiel de Sola Pool and mathematician Manfred Kochen. Written in 1958 and published in 1978, their work addressed in mathematical detail the small world effect, predicting that most individuals can be connected via two to threee acquaintances. Their paper inspired the experiments of Stanley Milgram.
Stanley Milgram (1933-1984)
American social psychologist who carried out the first experiment testing the small-world phenomena. (BOX 3.6).
John Guare (1938)
The phrase 'six degrees of separation' was introduced by the playwright John Guare, who used it as the title of his Broadway play.
Duncan J. Watts (1971),
Steven Strogatz (1959)
A new wave of interest in small worlds followed the study of Watts and Strogatz, finding that the small world property applies to natural and technological networks as well.
The **Facebook Data Team** measures the average distance between its users, finding "4 degrees" (BOX 3.6).

Coeficiente de Agrupamiento (Clustering)

COEFICIENTE DE AGRUPAMIENTO

$$C_i \equiv \frac{2 < L_i >}{k_i(k_i - 1)}$$



Dado que los enlaces son independientes y tienen la misma probabilidad p,

$$< L_i > \cong p \frac{k_i(k_i - 1)}{2} \Rightarrow C_i = \frac{2\langle L_i \rangle}{k_i(k_i - 1)} = p = \frac{\langle k \rangle}{N}.$$

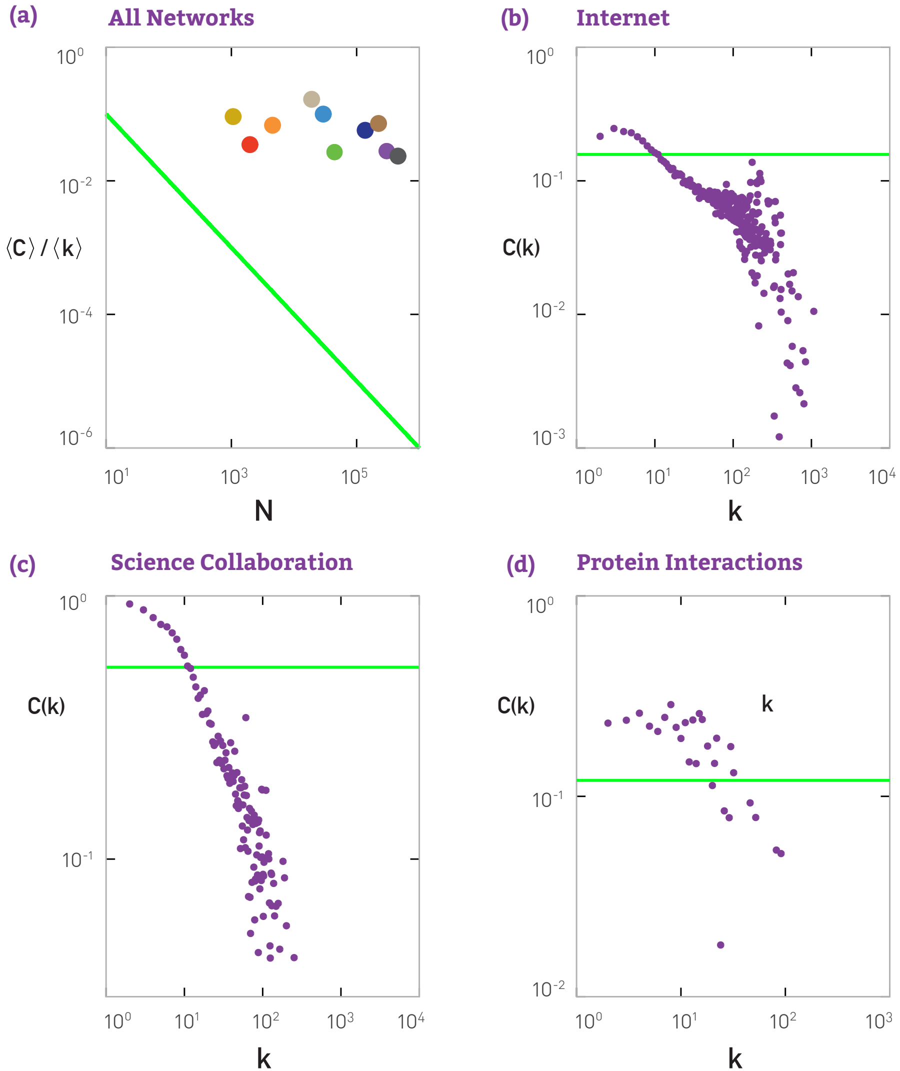
k_i=número de vecinos del nodo i
L_i=Número de enlaces que pasan por el nodo I

•El coeficiente de agrupamiento de los grafos aleatorios es pequeño.

•Para grados fijos, C disminuye con el tamaño del sistema N.

•C es independiente del grado k de nodos individuales.

COEFICIENTE DE AGRUPAMIENTO



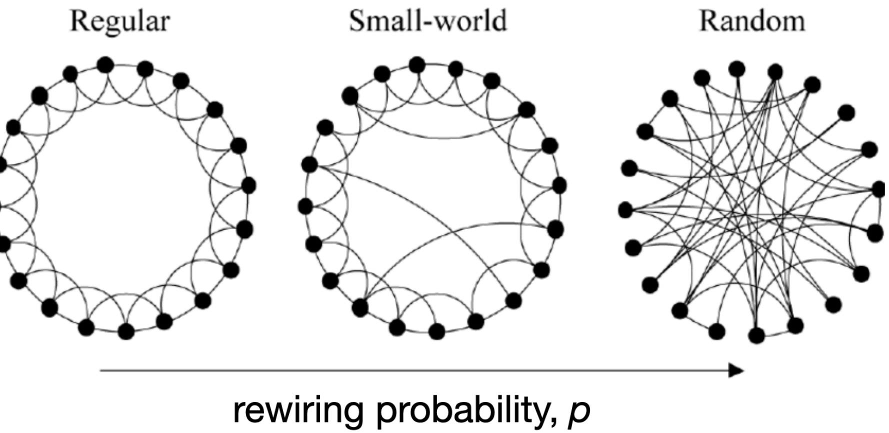
$$C_i = \frac{2\langle L_i \rangle}{k_i(k_i - 1)} = p = \frac{\langle k \rangle}{N}.$$

C disminuye con el tamaño del sistema N.

C es independiente del grado k de los nodos individuales.

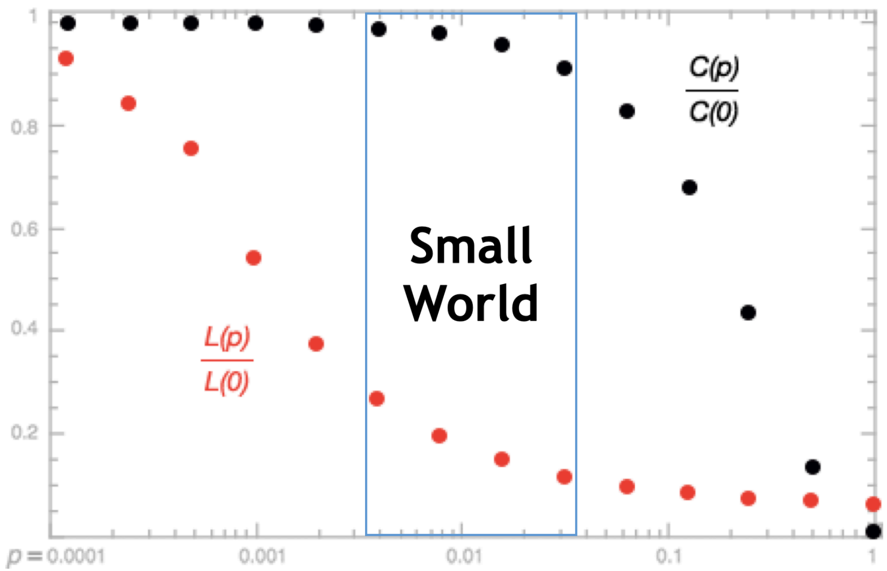
Small World Networks

Duncan J Watts   Steven Strogatz



Watts & Strogatz algorithm

A partir de una red de anillo regular, destruya cada enlace con probabilidad p, creando un nuevo enlace a un nodo aleatorio.



Las redes reales no son aleatorias.

¿Las redes reales son como los grafos aleatorios?

A medida que se dispone de datos cuantitativos sobre redes reales, podemos compara su topología con las predicciones de la teoría de grafos aleatorios.

Tenga en cuenta que una vez que tenemos N y <k> para una red aleatoria, de ella podemos derivar cada propiedad medible. De hecho, tenemos:

<u>Longitud media del camino:</u>

$$< l_{rand} > \approx \frac{\log N}{\log\langle k\rangle}$$

<u>Coeficiente de agrupamiento:</u>

$$C_i = \frac{2\langle L_i\rangle}{k_i(k_i-1)} = p = \frac{\langle k\rangle}{N}.$$

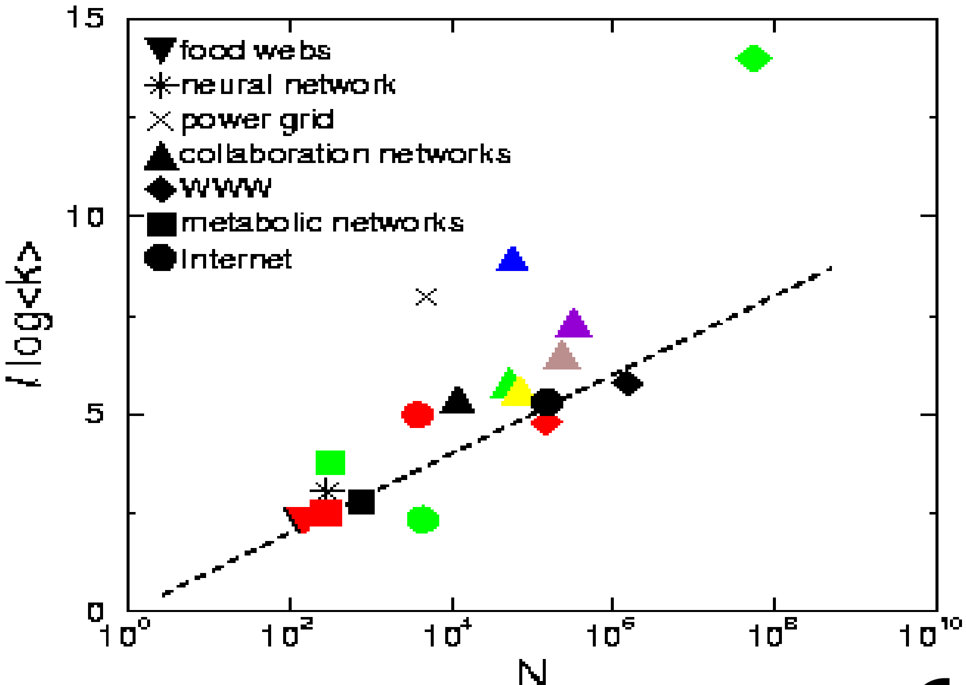
<u>Distribución de Grados:</u>

$$P(k) = e^{-<k>}\frac{<k>^{k}}{k!}$$

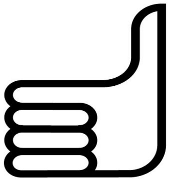
LONGITUD DE CAMINOS EN REDES REALES

Predicción:

$$< d >= \frac{\log N}{\log\langle k\rangle}$$



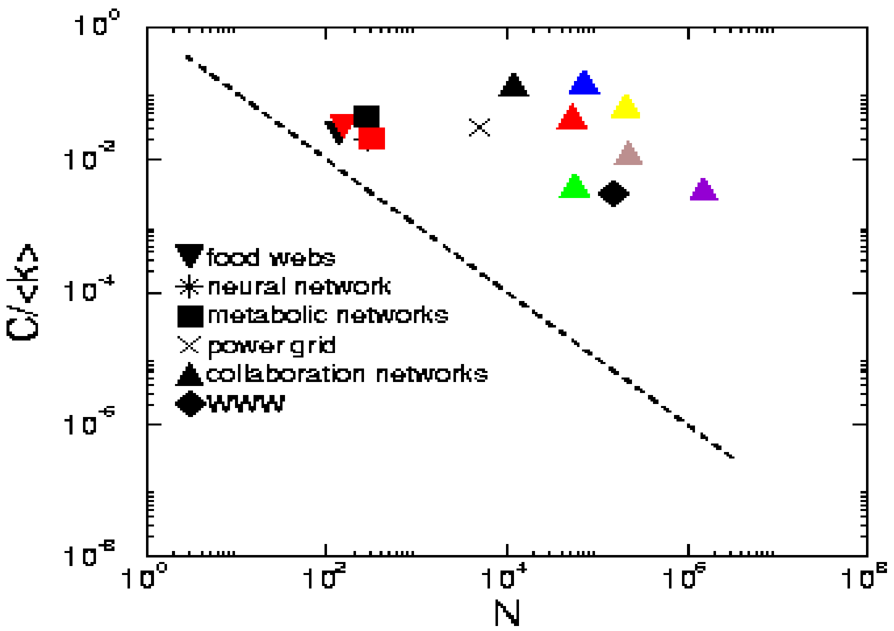
Las redes reales tienen distancias cortas como los grafos aleatorios.



COEFICIENTE DE CLUSTERING

Predicción:

$$C_i = \frac{2\langle L_i \rangle}{k_i(k_i - 1)} = p = \frac{\langle k \rangle}{N}.$$



C_{rand} Subestima en órdenes de magnitudes el coeficiente de agrupamiento de las redes reales.

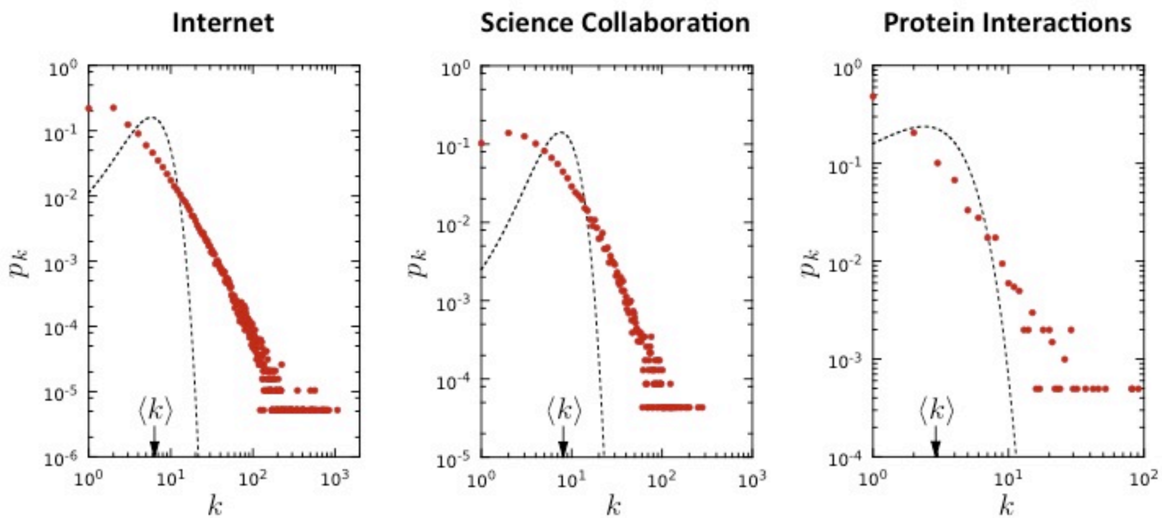
LA DISTRIBUCIÓN DE GRADO

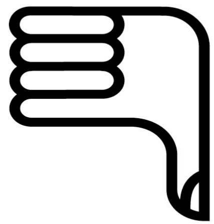
Predicción:

$$P(k) = e^{-<k>} \frac{<k>^k}{k!}$$

Data:

$$P(k) \approx k^{-\gamma}$$





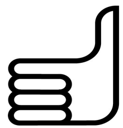
¿Las redes reales son como los grafos aleatorios?

A medida que se dispone de datos cuantitativos sobre redes reales, podemos compara su topología con las predicciones de la teoría de grafos aleatorios.

Tenga en cuenta que una vez que tenemos N y <k> para una red aleatoria, de ella podemos derivar cada propiedad medible. De hecho, tenemos:

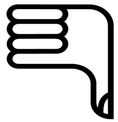
<u>Longitud media del camino:</u>

$$< l_{rand} > \approx \frac{\log N}{\log\langle k\rangle}$$



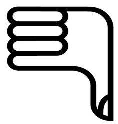
<u>Coeficiente de agrupamiento:</u>

$$C_i = \frac{2\langle L_i\rangle}{k_i(k_i - 1)} = p = \frac{\langle k\rangle}{N}.$$



<u>Distribución de Grados:</u>

$$P(k) = e^{-<k>}\frac{<k>^k}{k!}$$



¿ES EL MODELO GRÁFICO ALEATORIO RELEVANTE PARA LOS SISTEMAS REALES?

(B) Lo más importante: debemos preguntarnos, ¿son las redes reales aleatorias?

La respuesta simplemente es no

No hay ninguna red en la naturaleza que sepamos que sería descrita por el modelo de red aleatorio.

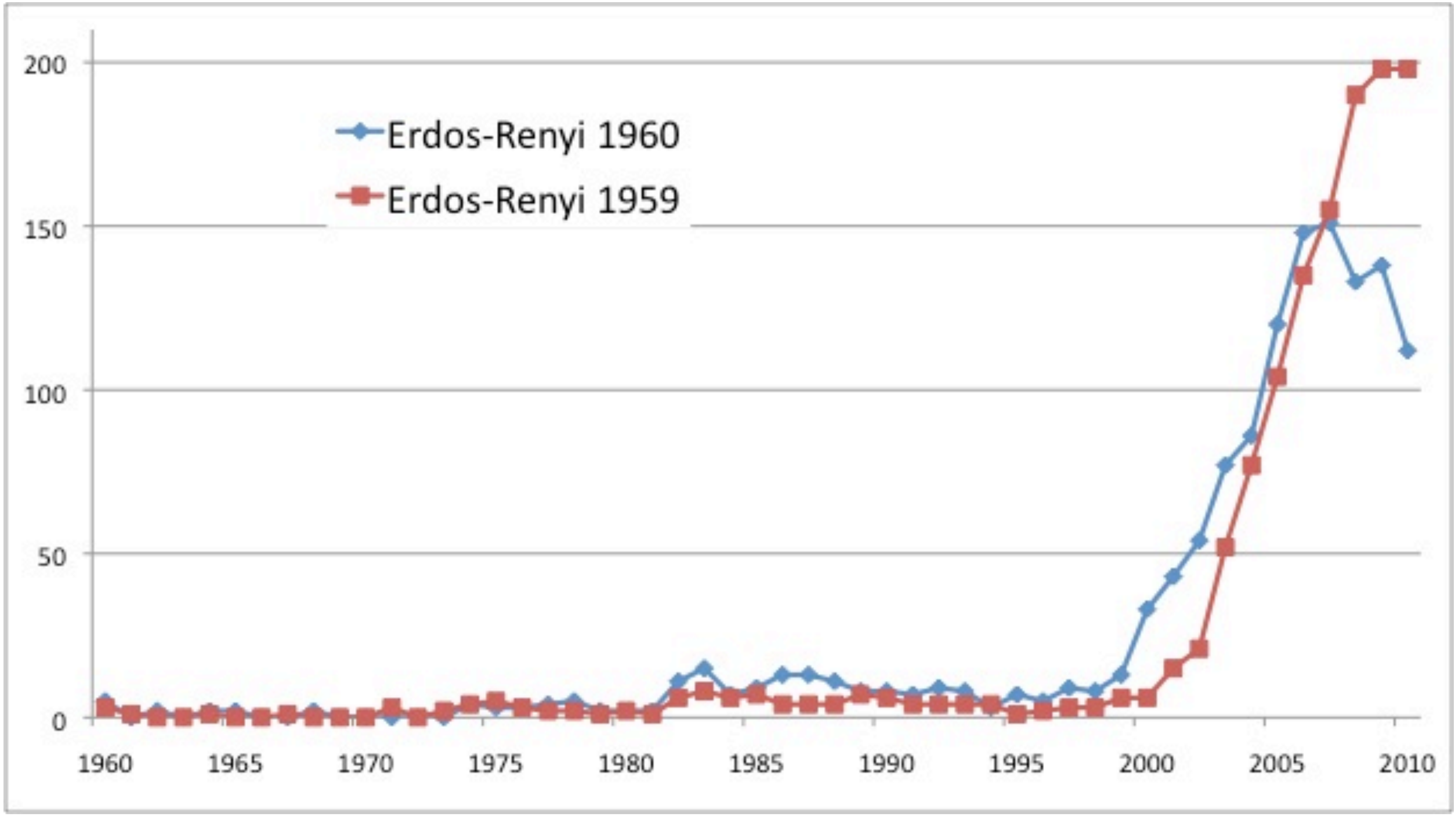
SI ES INCORRECTO E IRRELEVANTE, ¿POR QUÉ DEDICAMOS A UNA CLASE COMPLETA?

- Es el modelo de referencia para el resto de la clase.
- Nos ayudará a calcular muchas cantidades, que luego se pueden comparar con los datos reales, entendiendo hasta qué punto es una propiedad particular el resultado de algún proceso aleatorio.
- Patrones en redes reales, que son compartidos por una gran cantidad de redes reales, pero que se desvían de las predicciones del modelo de red aleatorio.
- Para identificarlos, debemos entender cómo se vería una propiedad en particular si está totalmente controlada por procesos aleatorios.

Si bien es EQUIVOCADO e IRRELEVANTE, ¡resultará extremadamente ÚTIL!

Resumen

Modelo Erdös-Rényi (1960)



NOTA HISTÓRICA

Anatol Rapoport
1911- 2007

1951, Rapoport y Solomonoff:

- Primer estudio sistemático de un grafo aleatorio.
- Demuestra la transición de fase.

Sistemas naturales: redes neuronales; las redes sociales de contactos físicos (epidemias); genética.

1959: *G(N,p)*

Edgar N. Gilbert
(**b**.1923)

¿Por qué lo llamamos el modelo aleatorio de Erdos-Renyi?

NOTA HISTÓRICA

DATOS DE RED: REDES DE COLABORACIÓN DE CIENCIAS

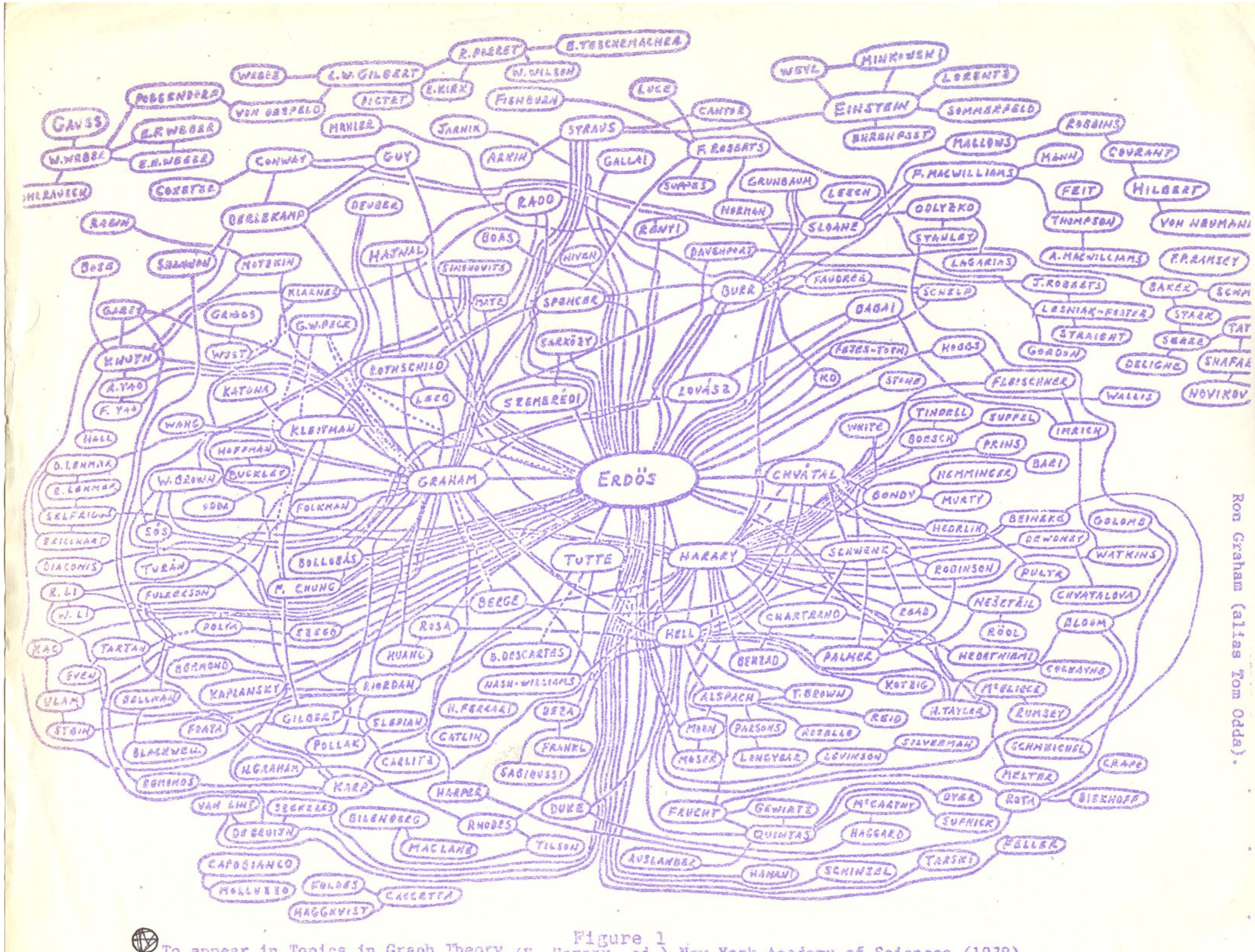


Figure 1
To appear in Topics in Graph Theory (F. Harary, ed.) New York Academy of Sciences (1979).

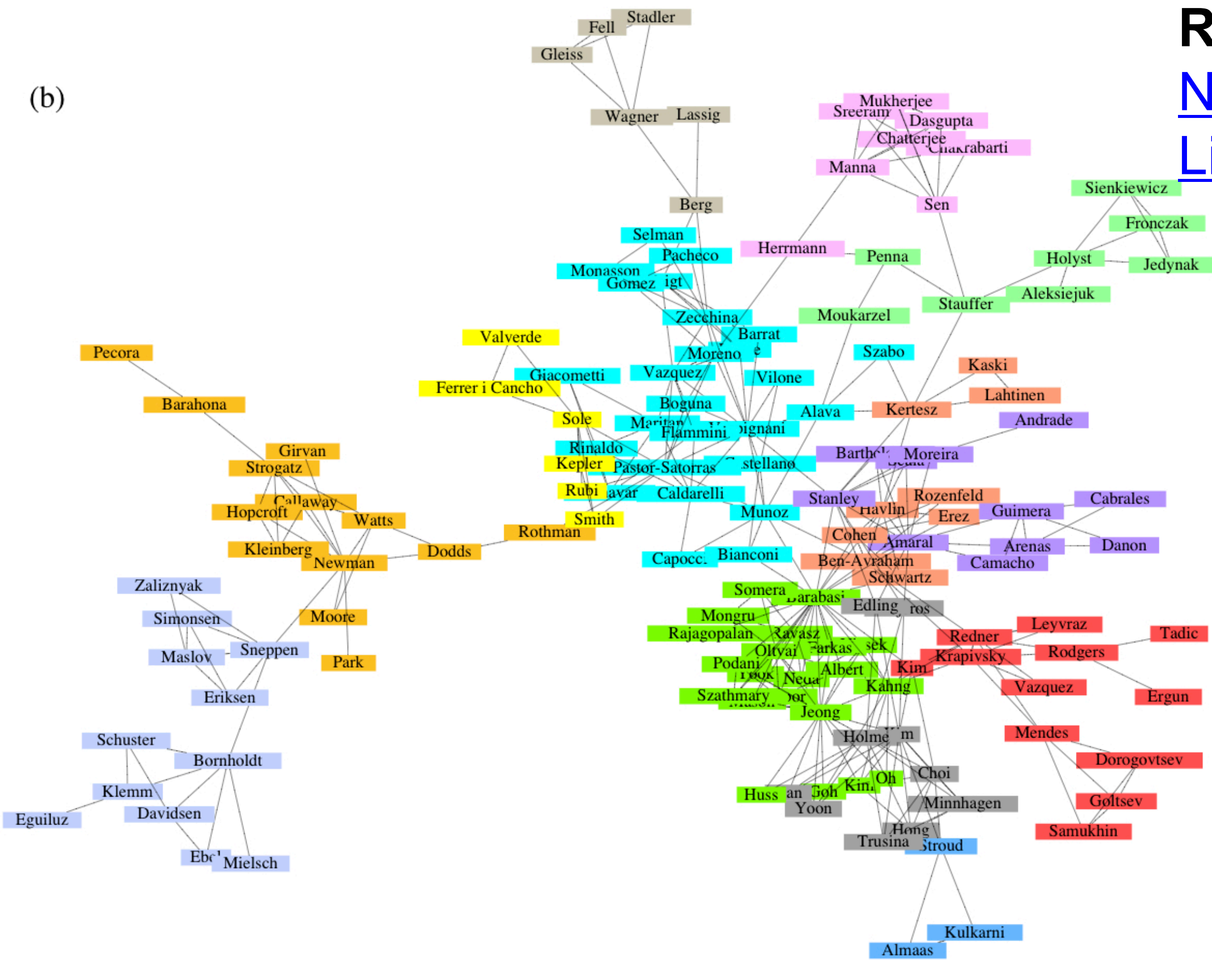
Erdos:
1,400 papers
507 coauthors

Einstein: EN=2
Paul Samuelson EN=5

….

ALB: EN= 3
CCV: EN=4

DATOS DE RED: REDES DE COLABORACIÓN DE CIENCIAS

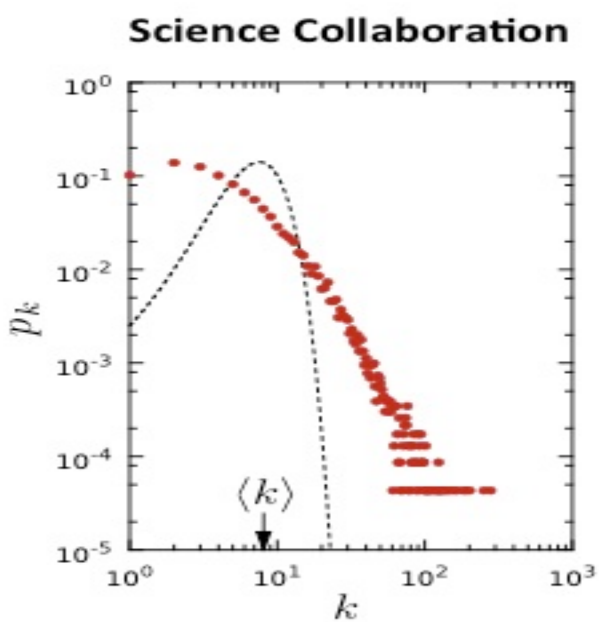


Red de colaboración
Nodos: Científicos
Links: Publicaciones conjuntas

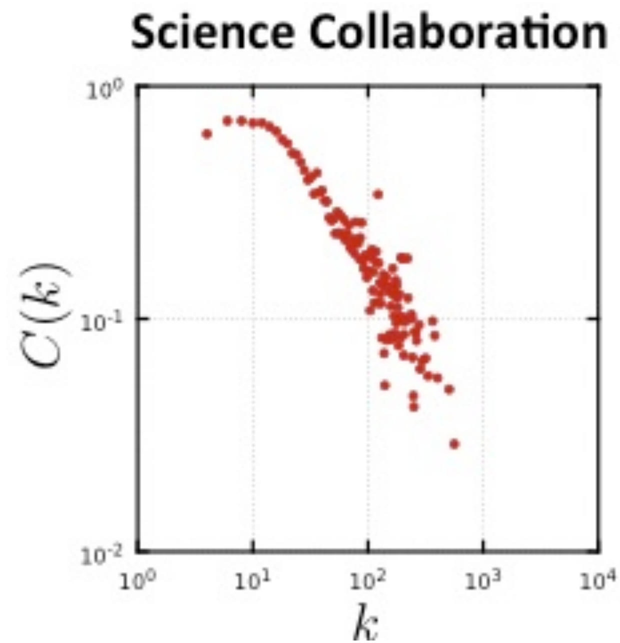
Physical Review:
1893 – 2009.

N=449,673
L=4,707,958

Ver también Stanford Large Network database
http://snap.stanford.edu/data/#canets.



Scale-free



Hierarchical