



**Redes y Sistemas Complejos**  
**Cuarto Curso del Grado en Ingeniería Informática**

**Tema 5: Modelos de Redes**  
**5.1. Redes Aleatorias**

***Oscar Cordon García***

*Dpto. Ciencias de la Computación e Inteligencia Artificial*  
*ocordon@decsai.ugr.es*

## JUSTIFICACIÓN DEL USO DE MODELOS DE REDES COMPLEJAS

Uno de los objetivos fundamentales de la Ciencia de las Redes es construir modelos que reproduzcan fielmente las propiedades de las redes complejas que se observan en los sistemas complejos reales

### ¿Por qué y para qué tener modelos de redes complejas?

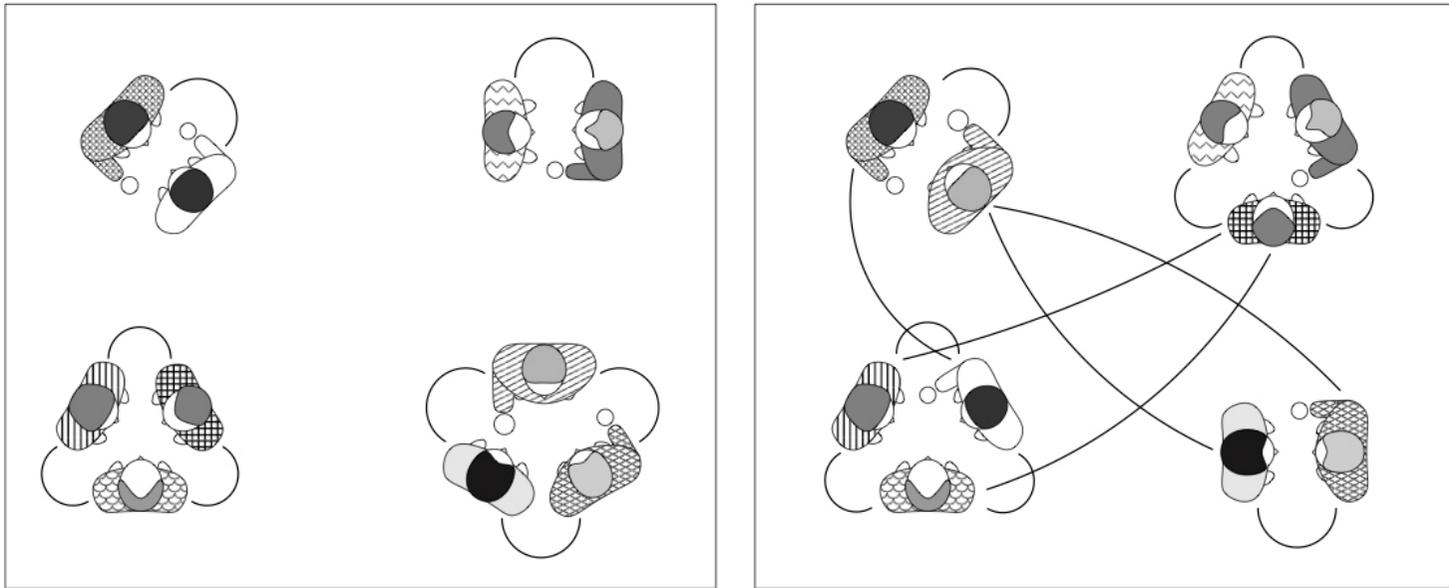
- Para tener representaciones simples de las redes
- Para poder derivar sus propiedades de forma matemática
- Para poder predecir propiedades y resultados

Y además para tener una referencia, básica y extremadamente simplificada, contra la que comparar:

- ¿Cuáles son las características diferenciales de mi red compleja real con respecto al modelo hipotético?
- ¿Qué conocimiento puede extraerse de estas diferencias?

# REDES ALEATORIAS

## EJEMPLO INTRODUCTORIO



Fiesta con 100 invitados tomando vino y charlando en grupos con encuentros aleatorios

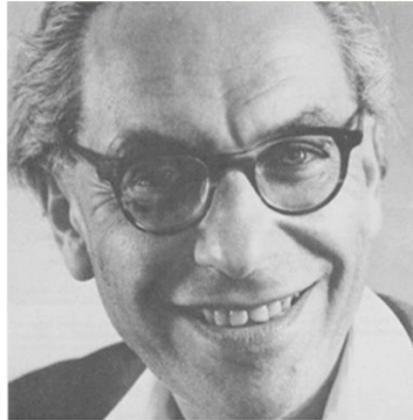
Comentario a uno de ellos sobre la calidad superior de un vino de los que se están sirviendo

**En teoría**, si cada encuentro dura 10 minutos, se tardarían 16 horas en que todos se enteraran

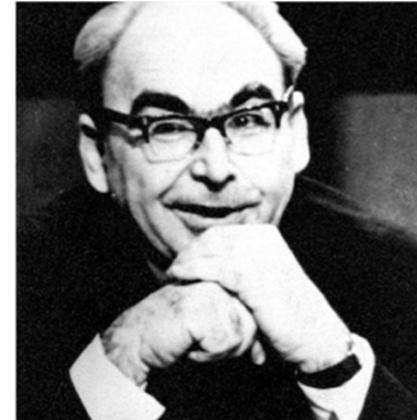
En realidad, el tiempo es mucho menor debido a la aparición de una **red aleatoria**

# MODELO DE RED ALEATORIA

**Pál Erdős**  
(1913-1996)

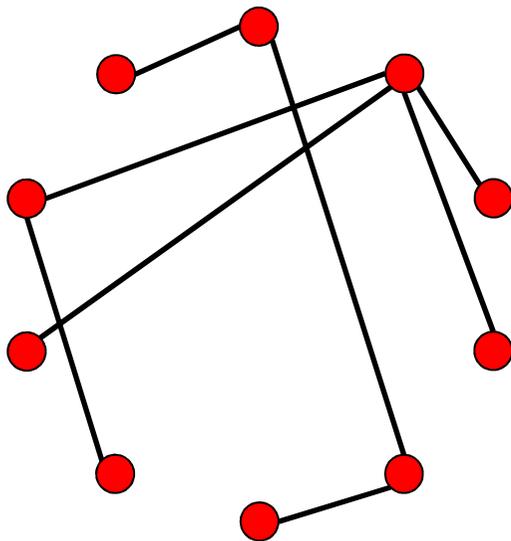


**Alfréd Rényi**  
(1921-1970)



$N=10$   $p=1/6$

$\langle k \rangle \sim 1.5$



## Modelo Erdős-Rényi (1960):

Un **grafo aleatorio**  $G(N,p)$  es un grafo no dirigido con  $N$  nodos donde cada par de nodos está conectado aleatoriamente con una probabilidad prefijada  $p$

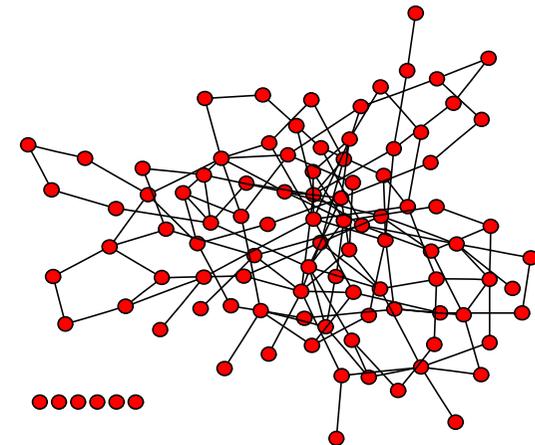
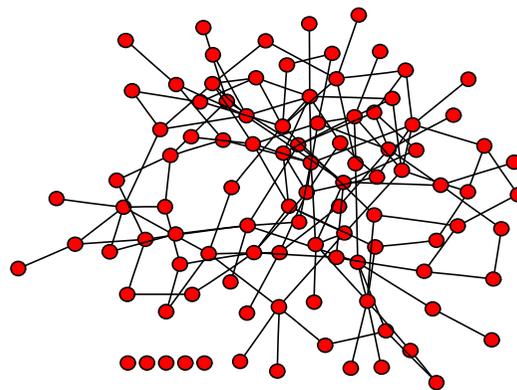
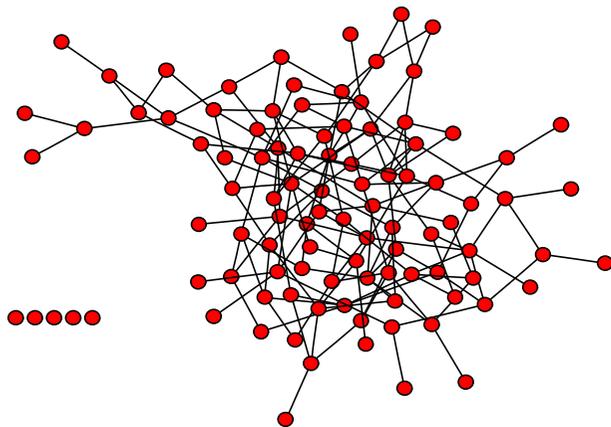
A primera vista, muchas redes reales parecen aleatorias

Este primer modelo de red compleja se basa en modelar esa aparente aleatoriedad mediante redes realmente aleatorias

# MODELO DE RED ALEATORIA:

# Ejemplo

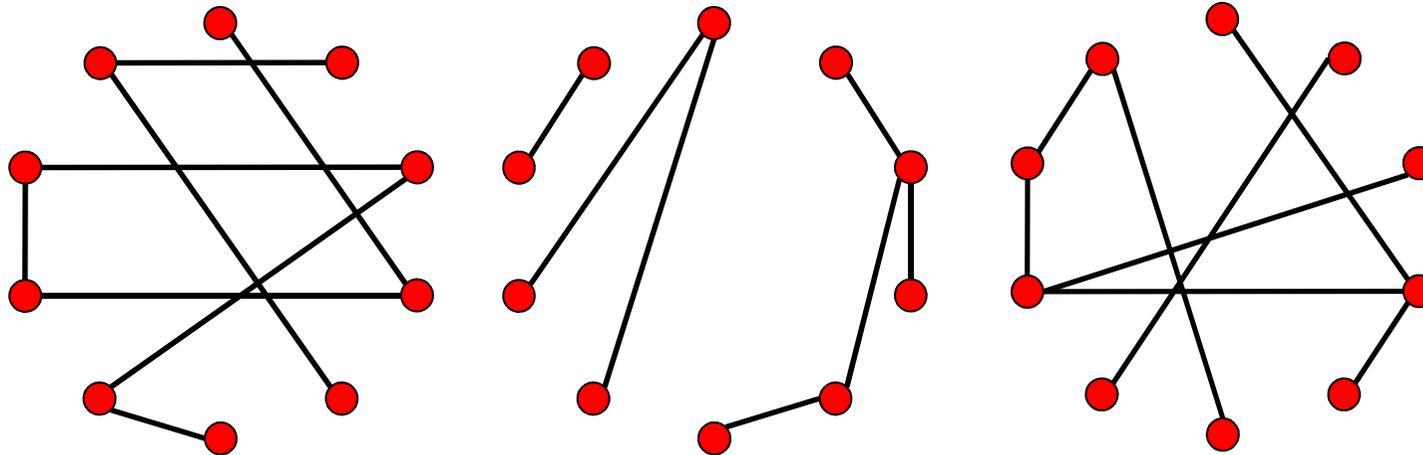
$N=100$   
 $p=0.03$



## MODELO DE RED ALEATORIA:

## Variabilidad de los grafos aleatorios

$N$  y  $p$  no definen unívocamente la red— puede haber muchas instanciaciones, diferentes en el número de enlaces  $L$  y en los enlaces concretos. **¿Cuántas?**



$N=10$   
 $p=1/6$

La probabilidad de generar un grafo aleatorio *concreto*  $G(N,p)$  con  $L$  enlaces es:

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

Es decir, cada grafo  $G(N,p)$  aparece con probabilidad  $P(G(N,p))$

## MODELO DE RED ALEATORIA:

## Variabilidad del número de enlaces

$P(L)$ : probabilidad de obtener exactamente  $L$  enlaces en una red de  $N$  nodos con probabilidad de conexión  $p$ :

$$P(L) = \underbrace{\binom{\binom{N}{2}}{L}}_{\text{Número de combinaciones distintas resultantes de escoger } L \text{ enlaces de entre todos los enlaces potenciales}} p^L (1-p)^{\frac{N(N-1)}{2} - L}$$

Número máximo de enlaces en una red de  $N$  nodos

Número de combinaciones distintas resultantes de escoger  $L$  enlaces de entre todos los enlaces potenciales

### Distribución binomial:

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

$$\langle x \rangle = Np$$

$$\langle x^2 \rangle = p(1-p)N + p^2 N^2$$

$$\sigma_x = (\langle k^2 \rangle - \langle k \rangle^2)^{1/2} = [p(1-p)N]^{1/2}$$

## MODELO DE RED ALEATORIA:

## Grado medio

$P(L)$ : probabilidad de obtener una red con exactamente  $L$  enlaces ([distribución binomial](#)):

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

- El número esperado de enlaces  $\langle L \rangle$  de un grafo aleatorio  $G(N,p)$  se calcula como el producto de la probabilidad de conexión y el número de enlaces ([media de la distribución](#)):

$$\langle L \rangle = \sum_{L=0}^{\binom{N(N-1)}{2}} LP(L) = p \frac{N(N-1)}{2}$$

- El grado medio es:**

$$\langle k \rangle = 2L / N = p(N-1)$$

- La varianza es:

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

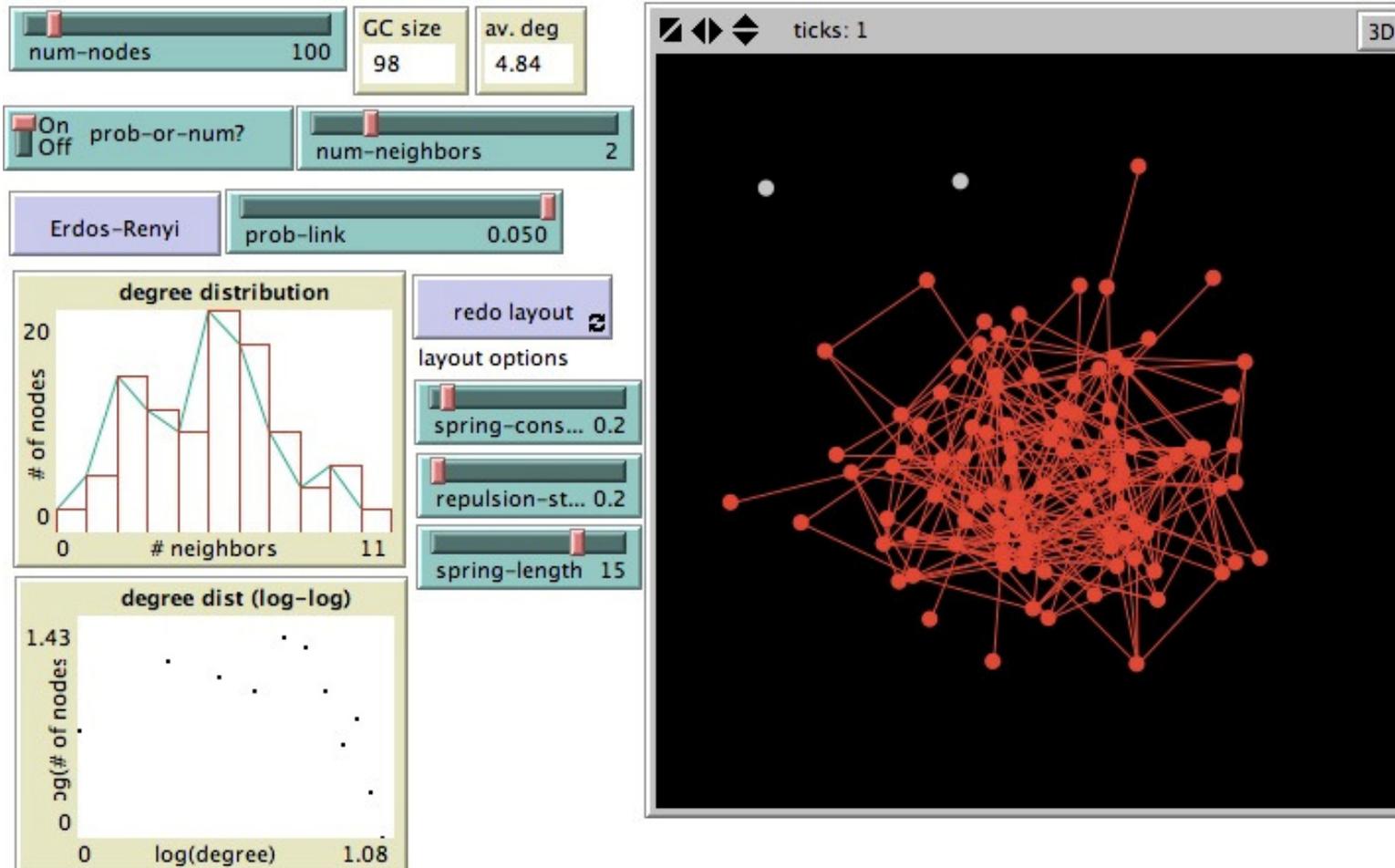
Si se mantiene fijo el valor de la probabilidad de conexión entre nodos,  $p$ , y se va aumentando el tamaño de la red, ¿qué ocurre con el grado medio de la red?

- a) Se mantiene constante
- b) Aumenta
- c) Disminuye

<http://ladamic.com/netlearn/NetLogo501/ErdosRenyiDegDist.html>

## MODELO DE RED ALEATORIA:

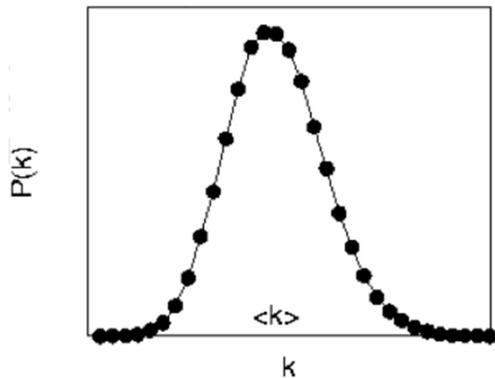
## Pregunta sobre el grado medio (2)



<http://ladamic.com/netlearn/NetLogo501/ErdosRenyiDegDist.html>

## MODELO DE RED ALEATORIA:

## Distribución (binomial) del grado



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

Selección de  $k$  nodos de los  $N-1$  a los que el nodo actual puede conectarse

Probabilidad de que  $k$  enlaces estén presentes

Probabilidad de que los otros  $N-1-k$  enlaces no estén presentes

$$\langle k \rangle = p(N-1)$$

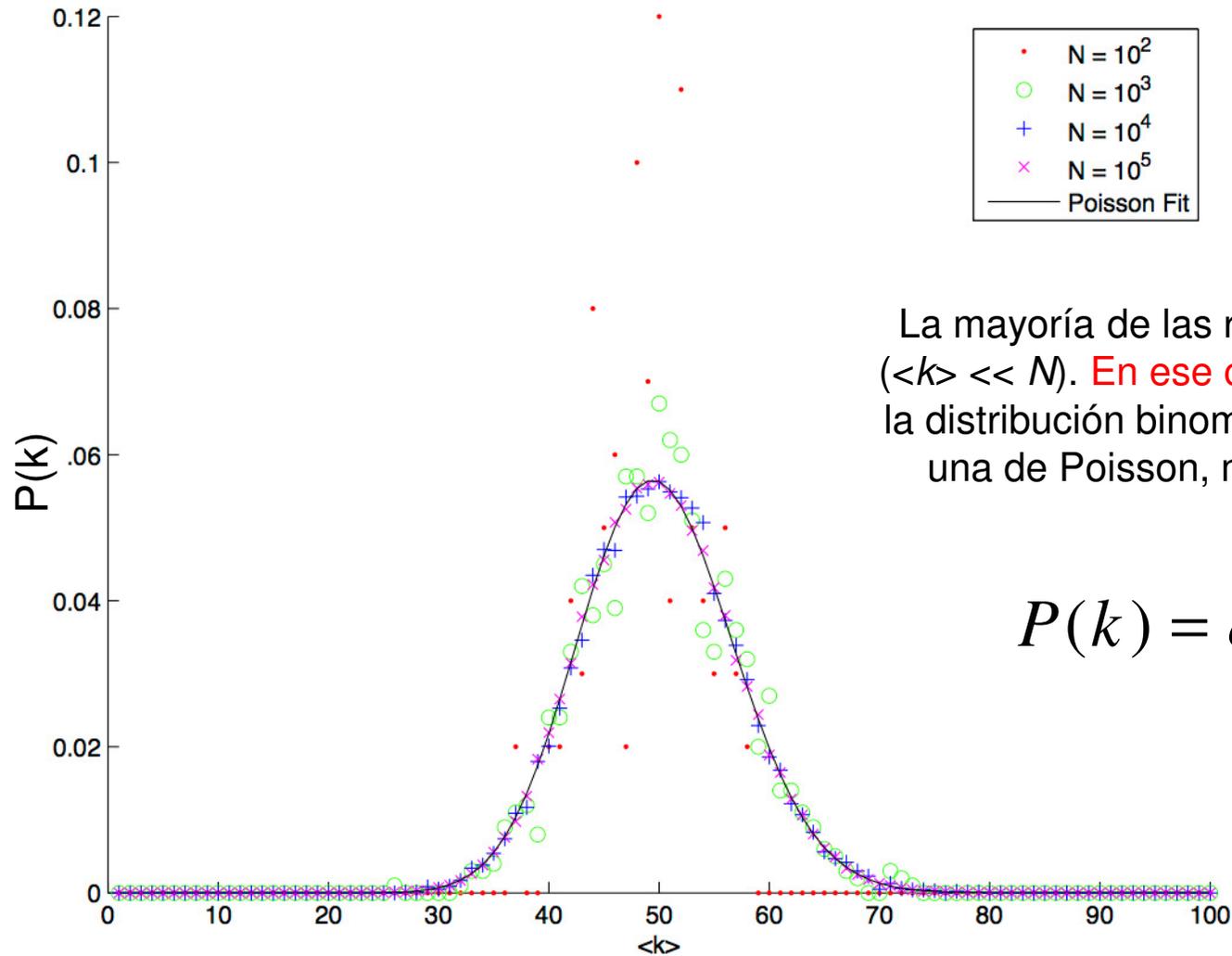
$$\sigma_k^2 = p(1-p)(N-1)$$

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

**Conforme aumenta el tamaño de la red, la distribución se va volviendo más estrecha**, es decir, tenemos una seguridad creciente de que el grado de un nodo está en la vecindad de  $\langle k \rangle$

## MODELO DE RED ALEATORIA:

## Distribución (Poisson) del grado



La mayoría de las redes reales son dispersas ( $\langle k \rangle \ll N$ ). En ese caso y cuando  $N$  es grande, la distribución binomial se puede aproximar por una de Poisson, más cómoda de manejar:

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

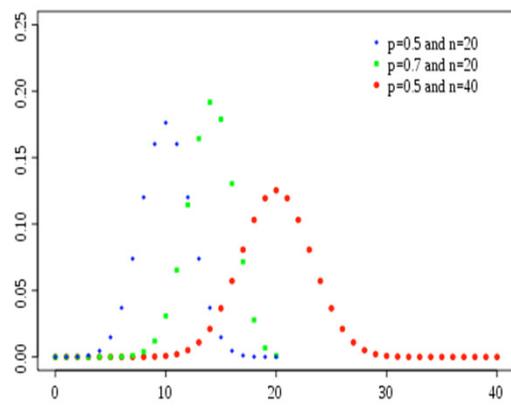
# MODELO DE RED ALEATORIA:

# Comparativa de distribuciones del grado

## Resultado Exacto

-distribución binomial-

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



Función de Distribución de Probabilidad (PDF)

$$\langle k \rangle = (N-1)p$$

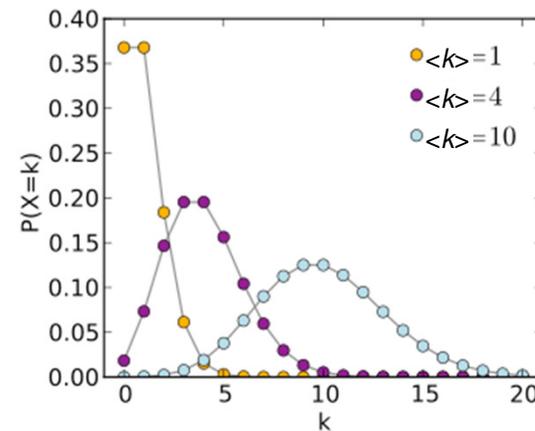
$$\langle k^2 \rangle = p(1-p)(N-1) + p^2(N-1)^2$$

$$\sigma_k = (\langle k^2 \rangle - \langle k \rangle^2)^{1/2} = [p(1-p)(N-1)]^{1/2}$$

## Límite con $\langle k \rangle \ll N$ y $N$ grande

-distribución de Poisson-

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$



$$\langle k \rangle = \langle k \rangle$$

$$\langle k^2 \rangle = \langle k \rangle (1 + \langle k \rangle)$$

$$\sigma_k = (\langle k^2 \rangle - \langle k \rangle^2)^{1/2} = \langle k \rangle^{1/2}$$

**Poisson:** Todo depende de un solo parámetro  $\langle k \rangle$  y no depende explícitamente del número de nodos  $N$ . Predice que la distribución del grado es la misma en redes de distintos tamaños en tanto en cuanto tengan el mismo grado medio

## MODELO DE RED ALEATORIA:

El grado de los nodos es bastante parecido

### ¿Por qué? Formalismo continuo:

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

En una red aleatoria con grado medio  $\langle k \rangle$  la probabilidad de tener un nodo cuyo grado sea mayor que un valor  $k_0$  es:

$$P(k > k_0) = \int_{k_0}^{\infty} e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} dk$$

Por ejemplo, con  $\langle k \rangle = 10$ ,

- La probabilidad de encontrar un nodo cuyo grado sea al menos el doble del grado medio es 0.00158826.
- La probabilidad de encontrar un nodo cuyo grado sea al menos diez veces el grado medio es  $1.79967152 \times 10^{-13}$
- La probabilidad de encontrar un nodo cuyo grado sea menor que un décimo del grado medio es 0.00049

<http://www.elektro-energetika.cz/calculations/po.php?language=espanol>

### ¿Por qué? Formalismo discreto:

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

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

- La probabilidad de hallar un nodo con un grado muy alto o muy bajo es exponencialmente pequeña
- Muchos nodos tienen valores de grado similares
- Cuanto mayor es el tamaño de la red aleatoria, mas parecidos son los grados de los nodos

## MODELO DE RED ALEATORIA:

No hay nodos con valores de grado extremos

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

Asumamos que la **red social mundial** se describe adecuadamente con un modelo de red aleatoria

En sociología, una persona cualquiera conoce directamente a otras mil,  $k \sim 1000$

La probabilidad de encontrar un individuo con grado  $k > 2000$  es  $10^{-27}$

**Dado que  $N \sim 10^9$ , la probabilidad de encontrar un individuo con 2000 conocidos es tan pequeña que esos nodos son virtualmente inexistentes en una sociedad aleatoria**

→ una sociedad aleatoria estaría formada principalmente por individuos promedio con más o menos el mismo número de amigos:  $\langle k \rangle = 1000 \rightarrow \sigma_k = \langle k \rangle^{1/2} = 31.62$

→ No presentaría outliers, individuos que sean bien altamente populares o bien solitarios

## MODELO DE RED ALEATORIA:

## Distribuciones de grado de redes reales

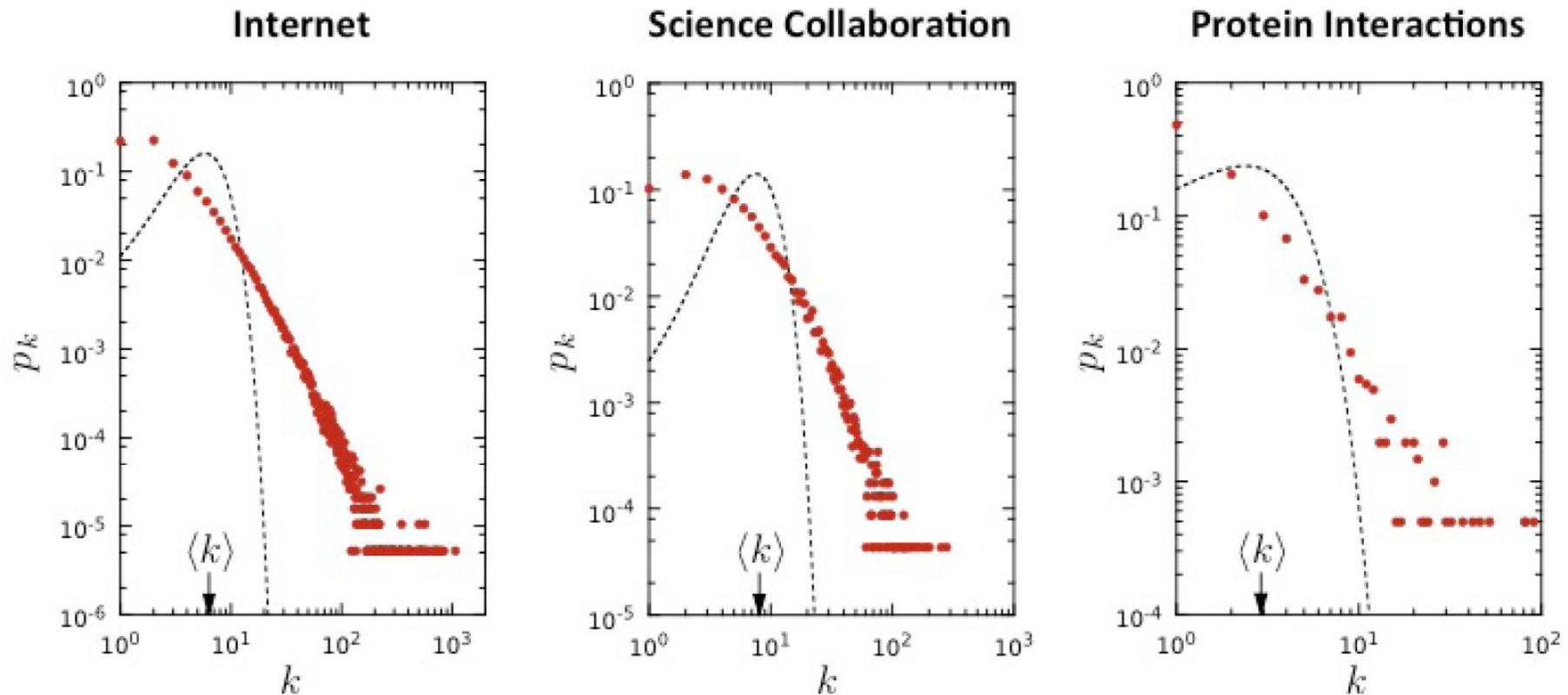


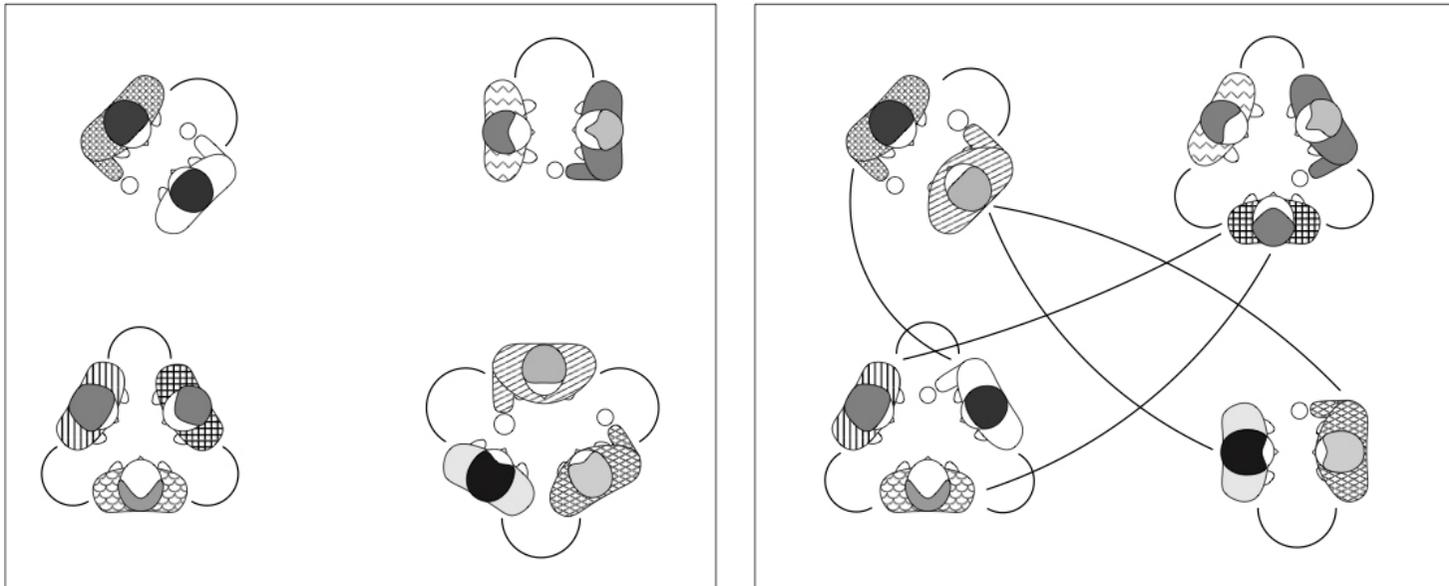
Image 3.5

### Degree distribution of real networks.

The degree distribution of the Internet, science collaboration network, and the protein interaction network of yeast (Table 2.1). The dashed line corresponds to the Poisson prediction, obtained by measuring  $\langle k \rangle$  for the real network and then plotting Eq. (8). The significant deviation between the data and the Poisson fit indicates that the random network model underestimates the size and the frequency of highly connected nodes, or hubs.

## MODELO DE RED ALEATORIA:

## Evolución de una red aleatoria (1)



Fiesta con 100 invitados tomando vino y charlando en grupos con encuentros aleatorios

Comentario a uno de ellos sobre la calidad superior de un vino de los que se están sirviendo

**En teoría**, si cada encuentro dura 10 minutos, se tardarían 16 horas en que todos se enteraran

En realidad, el tiempo es mucho menor debido a la aparición de una **red aleatoria**

## MODELO DE RED ALEATORIA:

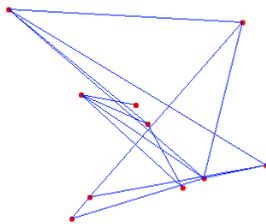
## Evolución de una red aleatoria (2)

Hasta ahora nos hemos centrado en las propiedades estáticas de un grafo aleatorio con un valor de  $p$  fijo

¿Qué pasa cuando variamos el valor del parámetro  $p$ ?

Java Applet that Creates a Random Graph with n Nodes:

Number of Nodes:  Connect Probability:



Number of Edges:

Connect Probability

IR A: <http://cs.gmu.edu/~astavrou/random.html>

Elegir  $N=100$

Ver como  $p$  aumenta en incrementos de 0.001, para los cuales, con  $N=100$ ,  $L=p \cdot N \cdot (N-1)/2 \sim p \cdot 50,000$ , i.e. cada incremento tiene en realidad 50 líneas nuevas

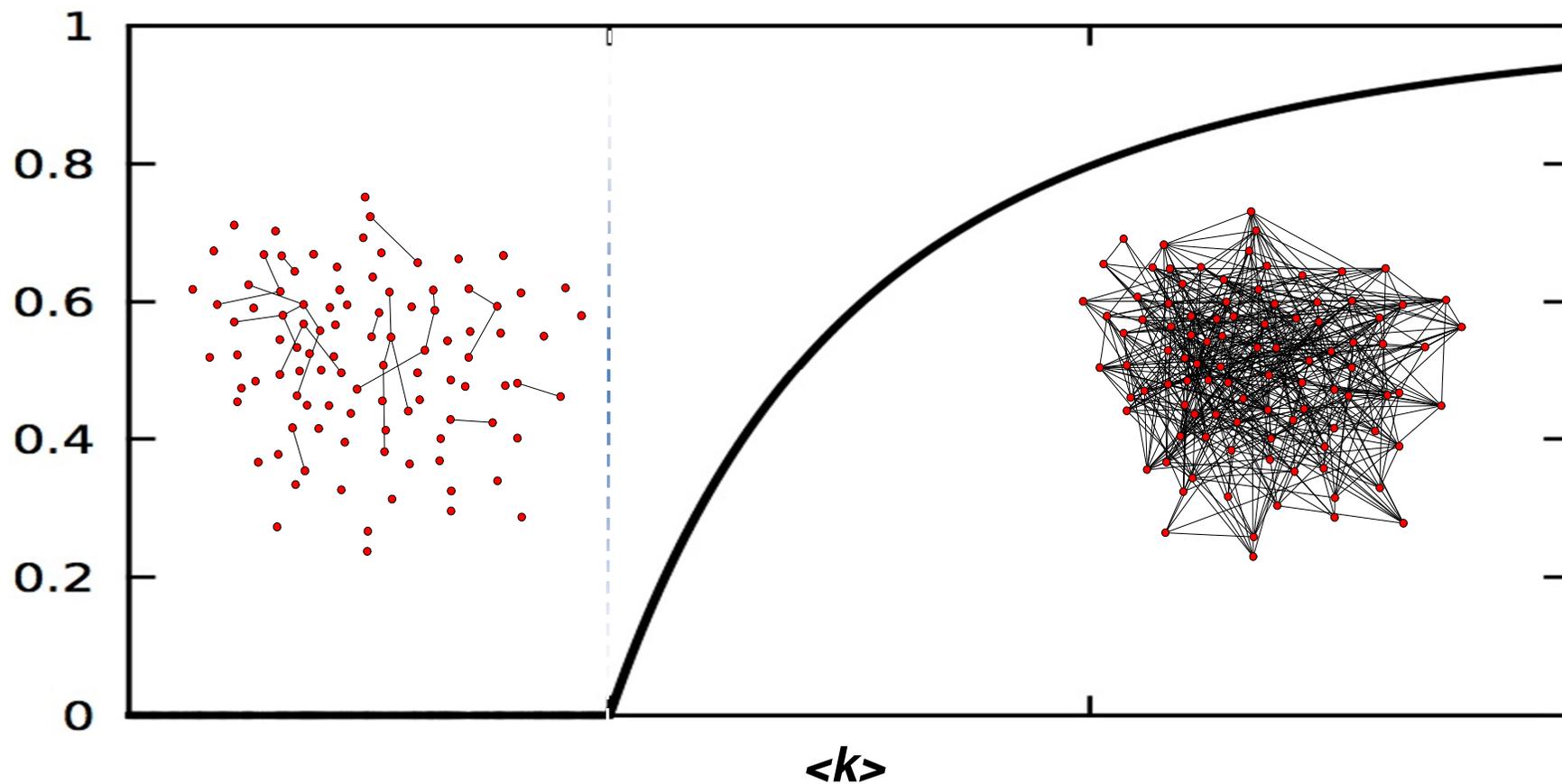
# MODELO DE RED ALEATORIA:

# Evolución de una red aleatoria (3)

Conjunto de nodos desconectados



RED



¿Cómo se da esta transición?

La transición de fase se produce cuando  $\langle k \rangle = 1$  (Erdos y Renyi, 1959):

Sea  $u = 1 - N_G/N$  la fracción de nodos que NO forman parte de la componente gigante (CG)  $N_G$

Para que un nodo  $i$  sea parte de la CG es necesario que se conecte via otro nodo  $j$ . Si  $i$  NO forma parte de la CG, puede deberse a dos razones:

Caso A: El nodo  $i$  no está conectado al nodo  $j$ , Probabilidad:  $1-p$

Caso B: El nodo  $i$  está conectado a  $j$  pero  $j$  no está conectado a la CG: Probabilidad:  $p \cdot u$

La probabilidad total de que  $i$  no sea parte de la CG via el nodo  $j$  es:  $1-p+p \cdot u$

La probabilidad de que  $i$  no esté enlazado a la CG via *cualquier otro nodo* es  $(1-p+p \cdot u)^{N-1}$

Por tanto:

$$u = (1 - p + pu)^{N-1}$$

Para cualquier  $p$  y  $N$  esta ecuación da el tamaño de la componente gigante como  $N_G = N \cdot (1-u)$

$$u = (1 - p + pu)^{N-1}$$

Usando  $p = \langle k \rangle / (N-1)$ , aplicando logaritmos en los dos lados y usando  $\langle k \rangle \ll N$  obtenemos:

$$\ln u = (N - 1) \ln \left[ 1 - \frac{\langle k \rangle}{N - 1} (1 - u) \right] \approx -(N - 1) \frac{\langle k \rangle}{N - 1} (1 - u) = -\langle k \rangle (1 - u)$$

Aplicando ahora una exponencial en los dos lados de la ecuación tenemos que:

$$u = e^{-\langle k \rangle (1 - u)}$$

Alternativamente, si notamos por  $S$  la fracción de nodos en la CG,  $S = N_G / N$ , i.e.  $S = 1 - u$

$$S = 1 - e^{-\langle k \rangle S}$$

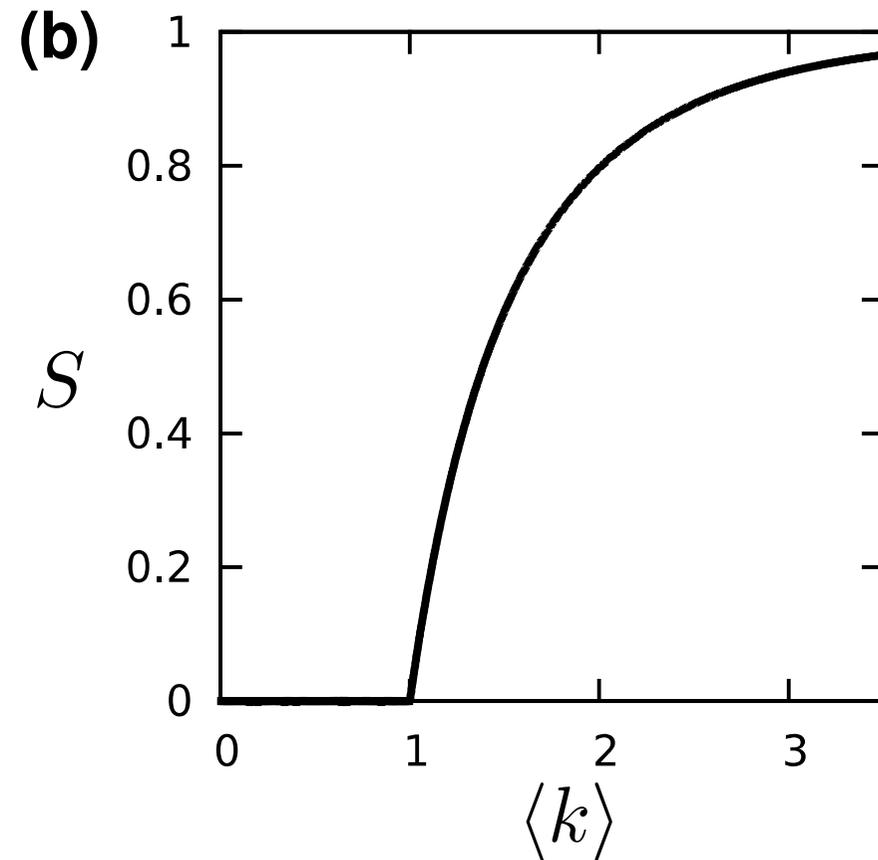
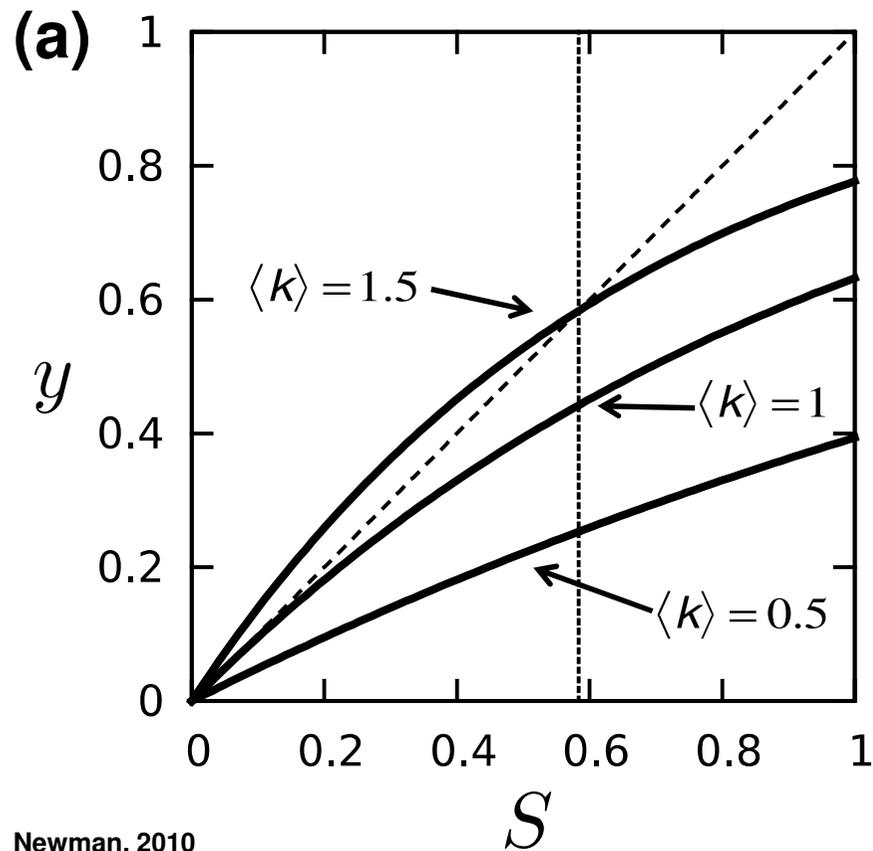
**(Erdos y Renyi, 1959)**

## MODELO DE RED ALEATORIA:

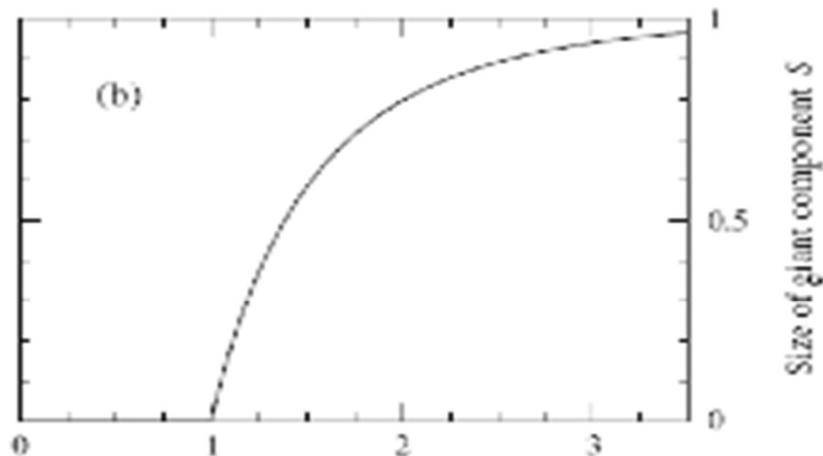
## Evolución de una red aleatoria (6)

$$S = 1 - e^{-\langle k \rangle S} \quad S: \text{fracción de nodos de la CG, } S = N_G/N$$

Punto de transición de fase:  $\frac{d}{dS}(1 - e^{-\langle k \rangle S}) = 1 \quad \langle k \rangle e^{-\langle k \rangle S} = 1$  Fijando  $S=0$ , obtenemos una transición de fase en  $\langle k \rangle = 1$



Resultado analítico



Resultado numérico

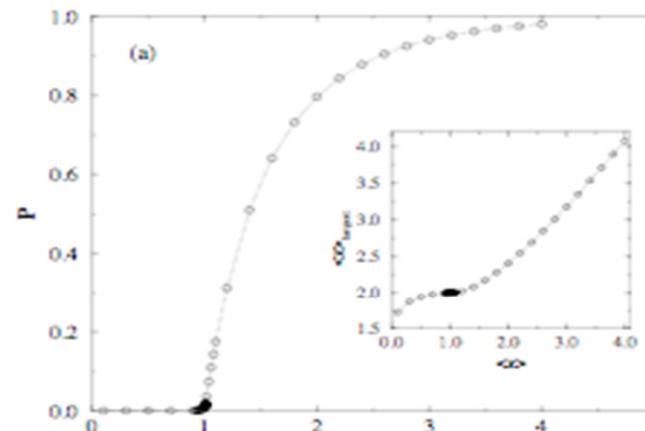


FIG. 1. (a) The strength  $P = \max_i \{n_i/N\}$  of the largest cluster as a function of the mean coordination number  $\langle z \rangle$ . There is a phase transition at  $\langle z \rangle_c = 1.00 \pm 0.01$  above which a macroscopic number of the species will be interconnected. The inset shows how the average coordination number in the largest cluster  $\langle z \rangle_{largest}$  increases with  $\langle z \rangle$  and reaches a value of  $1.9998 \pm 0.0005$  at  $\langle z \rangle_c$ . Both curves were obtained with an ensemble average over 100 systems with  $N = 10^6$ .

Evolution of Random Networks

Kim Christensen,<sup>1,2,\*</sup> Raul Donangelo,<sup>1</sup> Belita Koiller,<sup>1</sup> and Kim Sneppen<sup>3</sup>

<sup>1</sup>Instituto de Física, Universidade Federal do Rio de Janeiro, CP 68528, 21945-970 Rio de Janeiro, Brazil

<sup>2</sup>Department of Mathematics, Imperial College, 180 Queen's Gate, London SW7 2BZ, United Kingdom

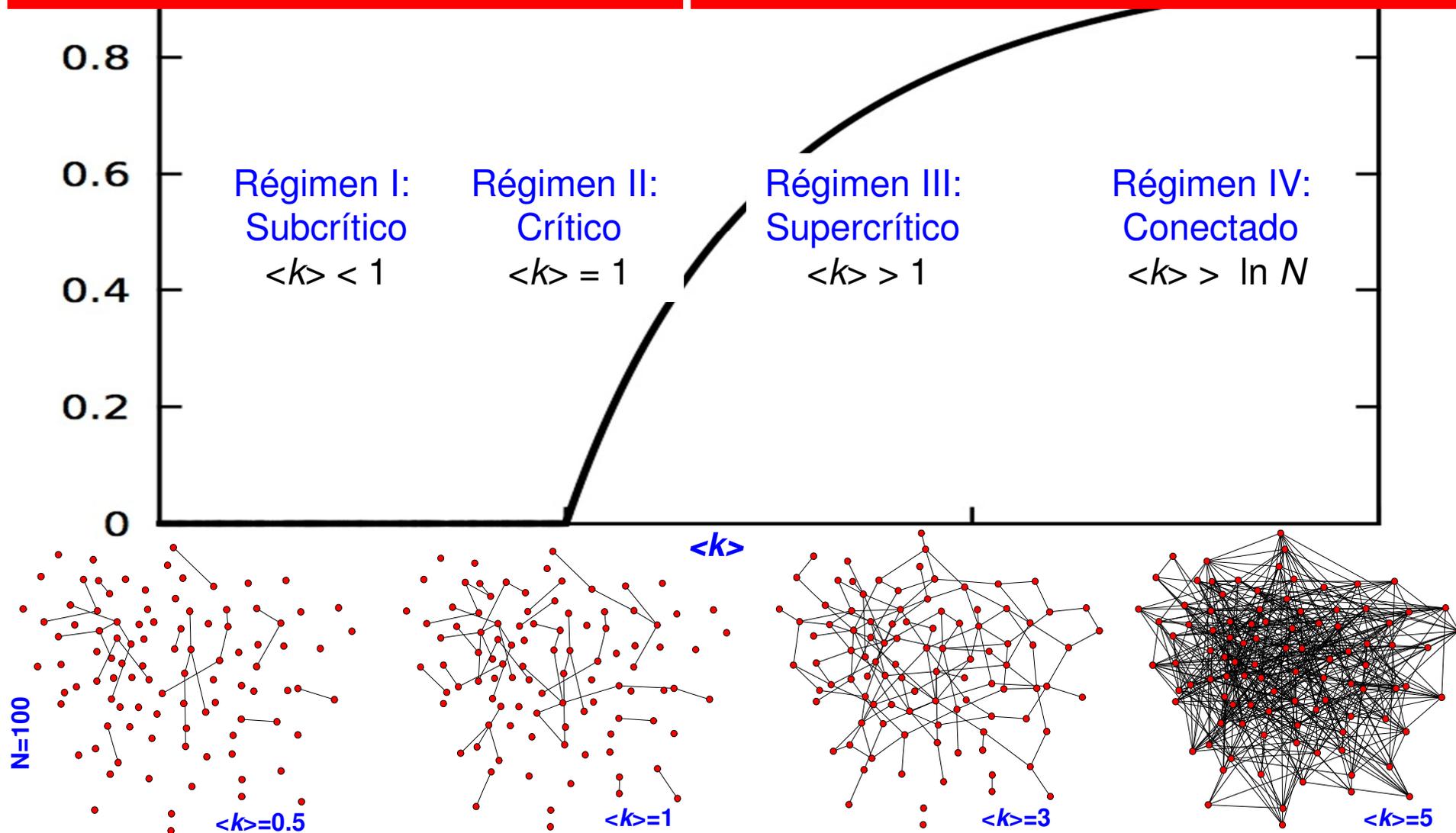
<sup>3</sup>NORDITA, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark

(Received 12 May 1998)

We investigate extremal dynamics on random networks. In the quenched case, after a transient time the dynamics is localized in the largest cluster. The activity in the largest cluster is nonergodic, with hot spots of activity typically centered around nodes with a high coordination number. The nonergodicity of the activity opens for models of evolving networks, which can self-organize into fractal geometries. [S0031-9007(98)07126-9]

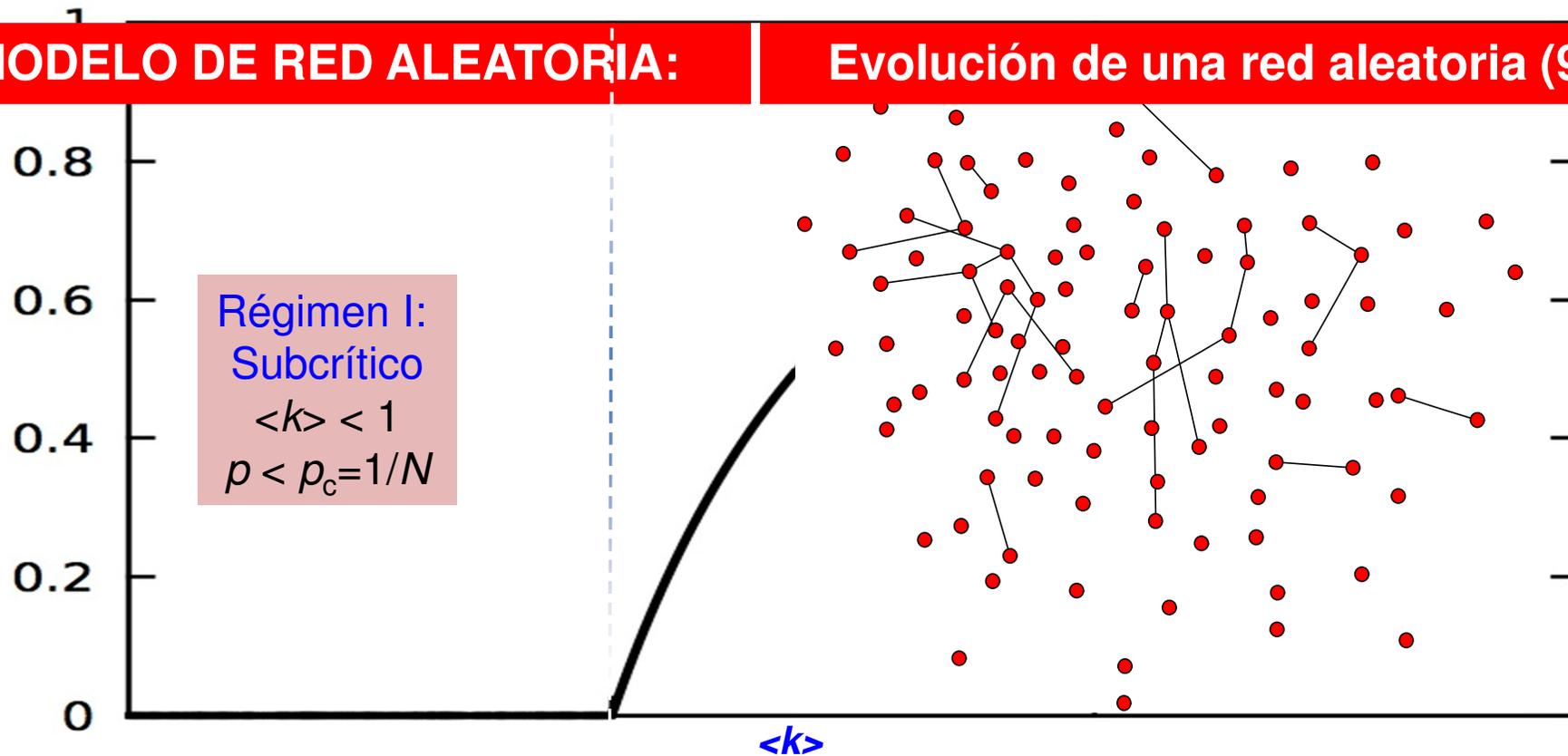
# MODELO DE RED ALEATORIA:

# Evolución de una red aleatoria (8)



## MODELO DE RED ALEATORIA:

## Evolución de una red aleatoria (9)



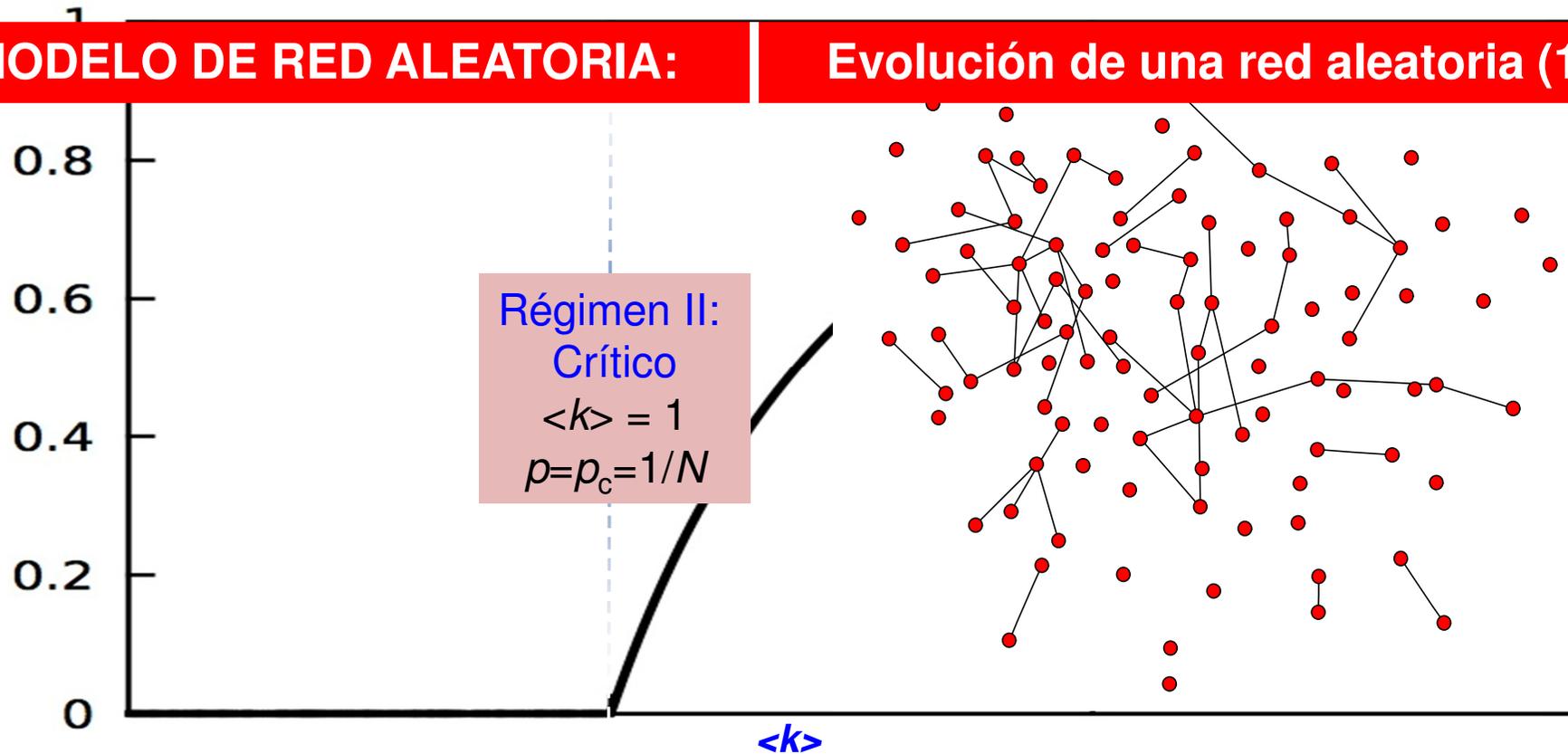
No existe una verdadera componente gigante.  $S = N_G/N \rightarrow 0$

Hay  $N-L$  clusters aislados. La distribución del tamaño de los clusters es exponencial:  $p(s) \sim e^{-\alpha \cdot s}$

El cluster más grande es un árbol de tamaño  $N_G \sim \ln N$

## MODELO DE RED ALEATORIA:

## Evolución de una red aleatoria (10)



Punto en el que **aparece** una única CG:  $N_G \sim N^{2/3}$   
→ aún contiene una fracción muy pequeña de los nodos,  $N_G/N \sim N^{1/3}$   
→ las componentes pequeñas son árboles, la CG tiene ciclos

La distribución del tamaño de los clusters es:  $p(s) \sim s^{-3/2}$

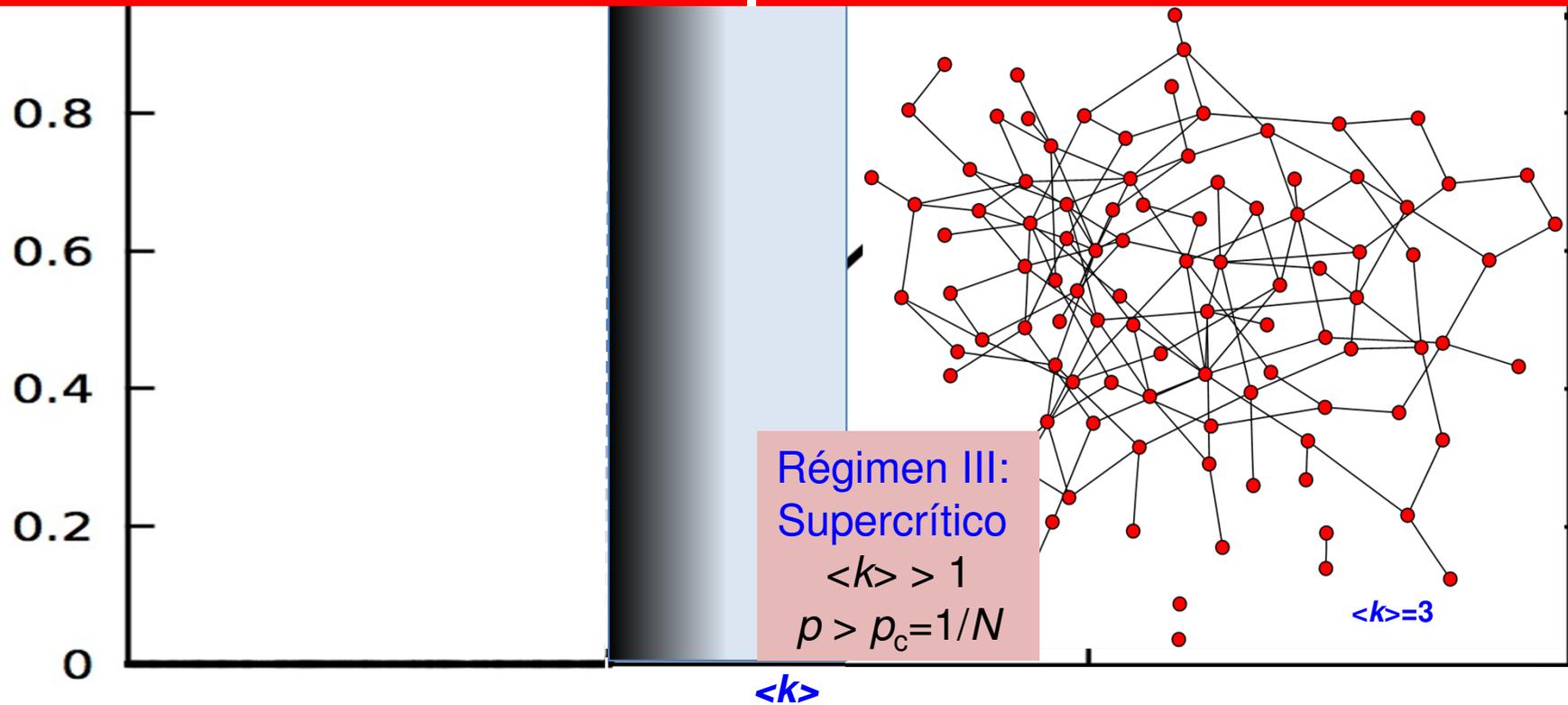
Se produce un salto en el tamaño de los clusters:

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

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

## MODELO DE RED ALEATORIA:

## Evolución de una red aleatoria (11)



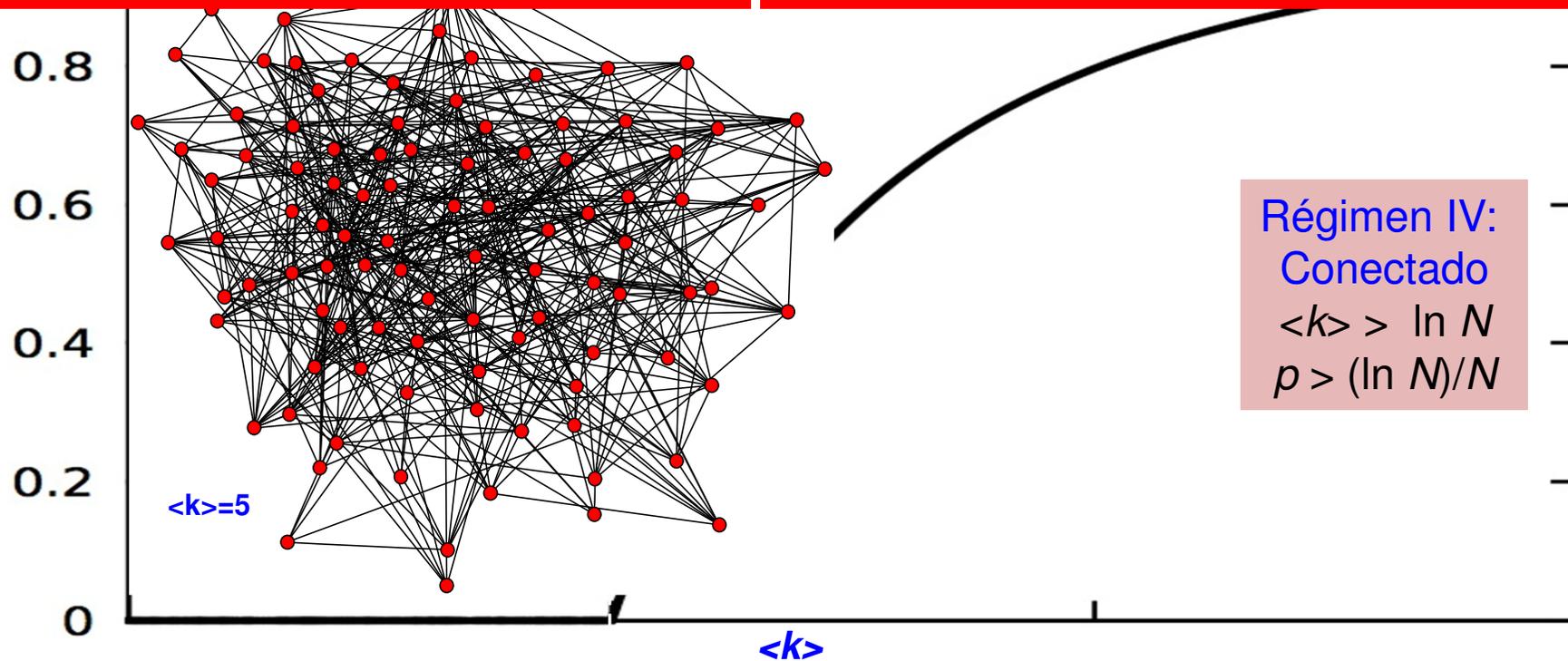
Hay una única CG real, con más nodos según nos alejamos del punto crítico:  $N_G \sim (p - p_c) \cdot N$   
→ La CG tiene muchos ciclos, las componentes pequeñas son árboles

La distribución del tamaño de los clusters es exponencial

Este estado se mantiene hasta que todos los nodos son absorbidos por la CG

## MODELO DE RED ALEATORIA:

## Evolución de una red aleatoria (12)



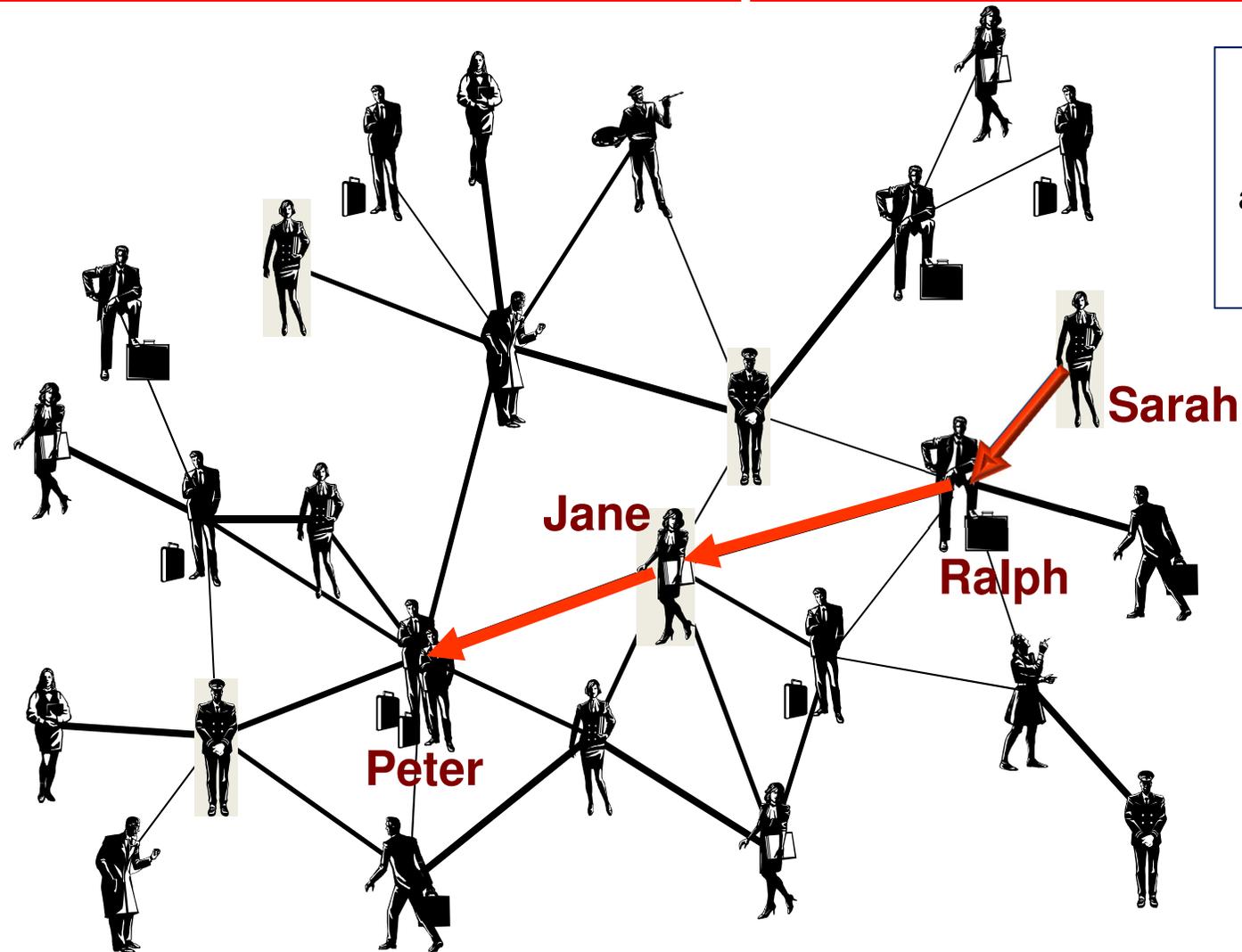
Para valores suficientemente grandes de  $p$  la CG absorbe todos los nodos y componentes. Sólo queda un cluster,  $N_G=N$ , con muchos ciclos

Cuando se entra en el estado conectado, la red es aún relativamente dispersa ( $(\ln N)/N \rightarrow 0$  en redes grandes). Sólo se convierte en un grafo completo con  $\langle k \rangle = N-1$

La distribución de tamaño de los clusters se anula

# MODELO DE RED ALEATORIA:

## Seis grados de separación - Mundos pequeños (1)



**Fenómeno de mundos pequeños:** la distancia entre dos nodos elegidos aleatoriamente de una red es sorprendentemente pequeña

*Frigyes Karinthy, 1929  
Stanley Milgram, 1967*



*Frigyes Karinthy (1887-1938)  
Escritor y periodista húngaro*

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

“Look, Selma Lagerlöf just won the Nobel Prize for Literature, thus she is bound to know King Gustav of Sweden, after all he is the one who handed her the Prize, as required by tradition. King Gustav, to be sure, is a passionate tennis player, who always participates in international tournaments. He is known to have played Mr. Kehrling, whom he must therefore know for sure, and as it happens I myself know Mr. Kehrling quite well.”

“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 Arpad Pasztor, someone I not only know, but to the best of my knowledge a good friend of mine. So I could easily ask him to send a telegram via the general director telling Ford that he should talk to the manager and have the worker in the shop quickly hammer together a car for me, as I happen to need one.”

**1967: Stanley Milgram** (1933-1984, Psicólogo social estadounidense)

#### HOW TO TAKE PART IN THIS STUDY

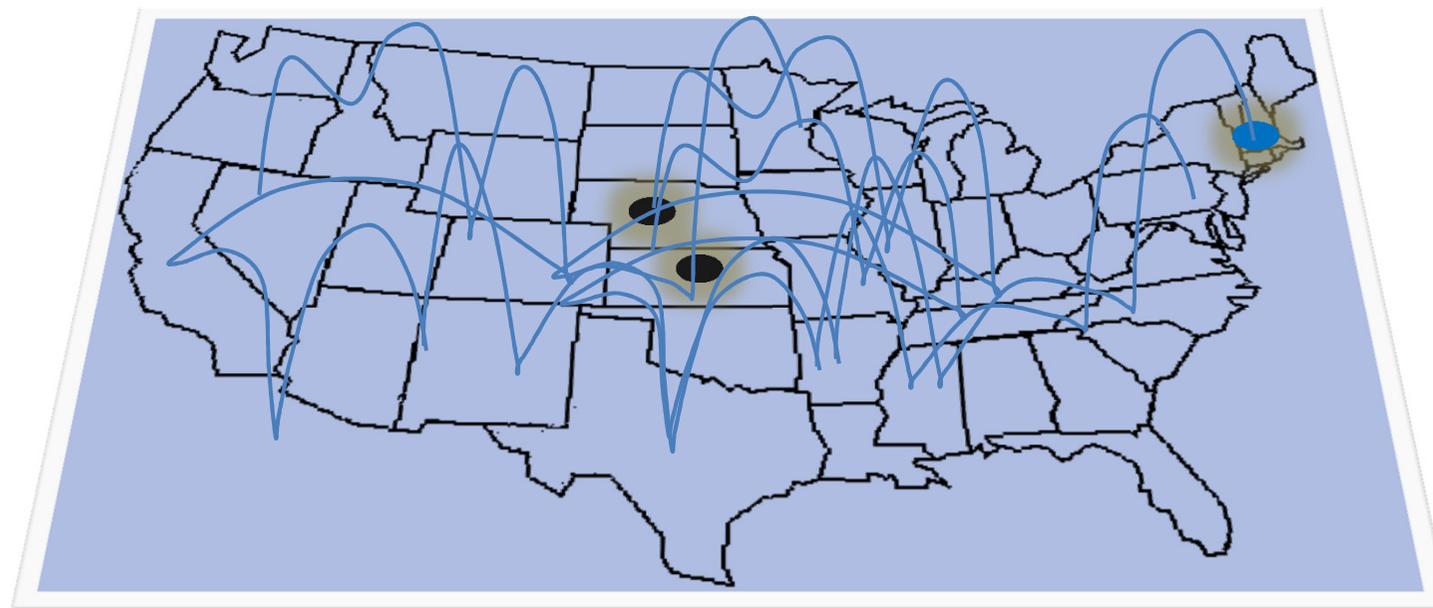
1. ADD YOUR NAME TO THE ROSTER AT THE BOTTOM OF THIS SHEET, so that the next person who receives this letter will know who it came from.
2. DETACH ONE POSTCARD. FILL IT AND RETURN IT TO HARVARD UNIVERSITY. No stamp is needed. The postcard is very important. It allows us to keep track of the progress of the folder as it moves toward the target person.
3. IF YOU KNOW THE TARGET PERSON ON A PERSONAL BASIS, MAIL THIS FOLDER DIRECTLY TO HIM (HER). Do this only if you have previously met the target person and know each other on a first name basis.
4. IF YOU DO NOT KNOW THE TARGET PERSON ON A PERSONAL BASIS, DO NOT TRY TO CONTACT HIM DIRECTLY. INSTEAD, MAIL THIS FOLDER (POST CARDS AND ALL) TO A PERSONAL ACQUAINTANCE WHO IS MORE LIKELY THAN YOU TO KNOW THE TARGET PERSON. You may send the folder to a friend, relative or acquaintance, but it must be someone you know on a first name basis.

## MODELO DE RED ALEATORIA:

## Seis grados de separación - Mundos pequeños (4)

Las primeras cartas en cadena... Orígenes: Omaha, Nebraska y Wichita, Kansas, EEUU. Destino: Boston y Sharon, Massachusetts, EEUU

Se enviaron 296 cartas. La primera llegó en pocos días, pasando sólo por 2 enlaces. Al final llegaron 64 con un máximo de 12 intermediarios



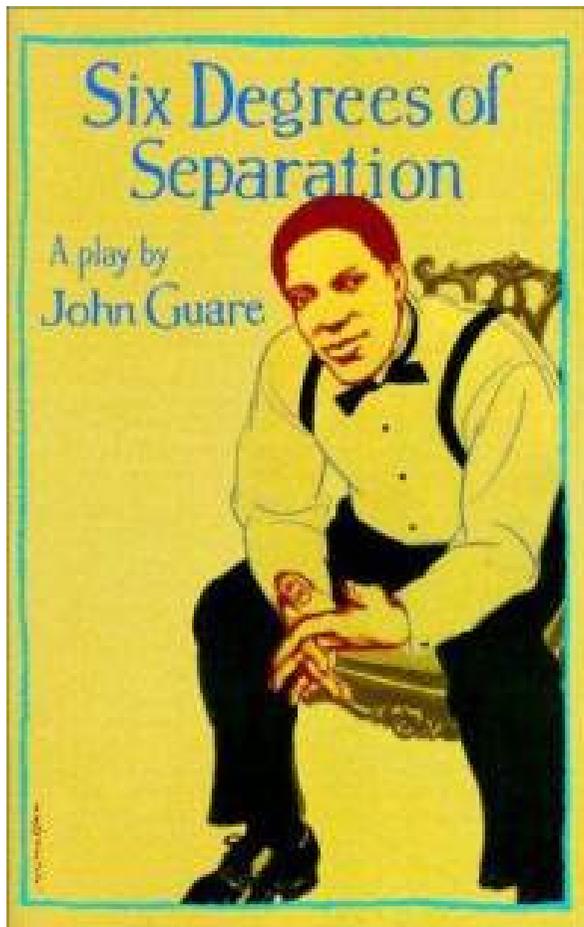
## SEIS GRADOS DE SEPARACIÓN:

La mediana de intermediarios fue de entre 5.5 y 6

Travers y Milgram, Sociometry 32, 425 (1969)

## MODELO DE RED ALEATORIA:

## Seis grados de separación - Mundos pequeños (5)



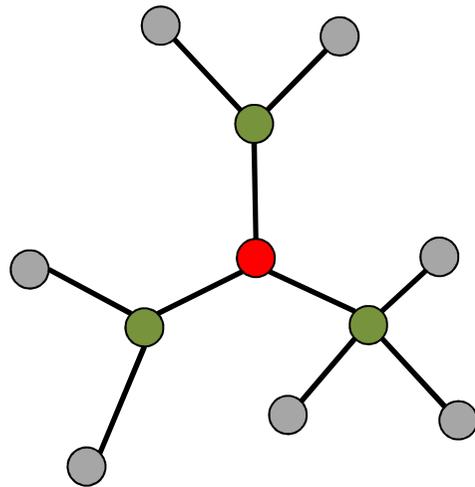
**Origen del nombre: 1991: Obra de Broadway homónima de John Guare** (dramaturgo estadounidense):

"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 Fuego. 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."

## MODELO DE RED ALEATORIA:

## Distancias en redes aleatorias (1)

Los grafos aleatorios tienden a tener una topología en forma de árbol con nodos de grado casi constante:



- no. de vecinos a distancia 1:  $N_1 \cong \langle k \rangle$
- no. de vecinos a distancia 2:  $N_2 \cong \langle k \rangle^2$
- no. de vecinos a distancia  $d$ :  $N_d \cong \langle k \rangle^d$
  
- Estimación de la distancia máxima  
(**fórmula de los mundos pequeños**):

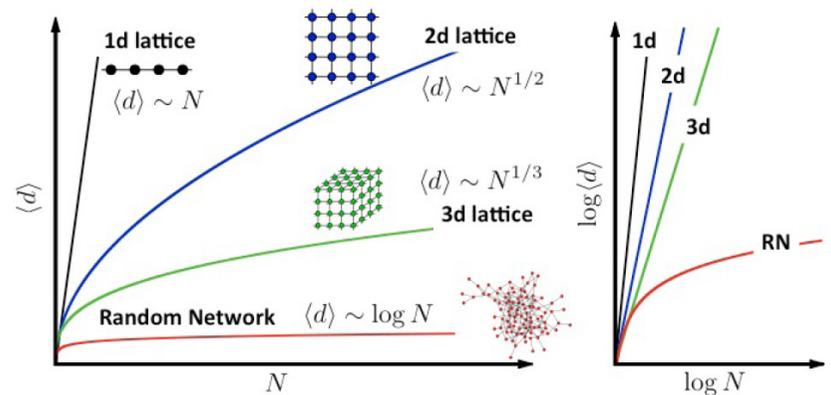
$$N = 1 + \langle k \rangle + \langle k \rangle^2 + \dots + \langle k \rangle^d = \frac{\langle k \rangle^{d+1} - 1}{\langle k \rangle - 1} \approx \langle k \rangle^d \quad \Rightarrow \quad d_{\max} = \frac{\log N}{\log \langle k \rangle}$$

Comentarios:

- En la mayoría de los casos, la fórmula aproxima mejor  $\langle d \rangle$  que  $d_{\max}$ . Esto se debe a que en la práctica  $d_{\max}$  está dominada por unos pocos caminos de longitud extrema mientras que  $\langle d \rangle$  está promediada entre todos los pares de nodos

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

- En general,  $\log N \ll N$ . Por tanto, el hecho de que  $\langle d \rangle$  dependa de  $\log N$  implica que **las distancias en una red aleatoria son varios órdenes de magnitud menores que el tamaño de la red**



- El término  $1/\log \langle k \rangle$  implica que, cuanto más densa es la red, menor es la distancia entre sus nodos

## MODELO DE RED ALEATORIA:

## Distancias en redes aleatorias (3)

En la práctica, la fórmula aproxima mejor  $\langle d \rangle$  que  $d_{\max}$

Validación sobre redes reales:

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

<i>Network Name</i>	<i>N</i>	<i>L</i>	$\langle k \rangle$	$\langle d \rangle$	$d_{\max}$	$\frac{\log N}{\log \langle k \rangle}$
Internet	192,244	609,066	6.34	6.98	26	6.59
WWW	325,729	1,497,134	4.60	11.27	93	8.32
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	186,936	8.08	5.35	15	4.81
Actor Network	212,250	3,054,278	28.78	-	-	-
Citation Network	449,673	4,707,958	10.47	11.21	42	5.55
E Coli Metabolism	1,039	5,802	5.84	2.98	8	4.04
Yeast Protein Interactions	2,018	2,930	2.90	5.61	14	7.14

Dadas las grandes diferencias en temática, tamaño y grado medio, el ajuste es bastante bueno

## MODELO DE RED ALEATORIA:

## Coeficiente de Clustering (1)

Para calcular  $C_i$  en una red aleatoria hay que estimar el número esperado de enlaces  $L_i$  entre los  $k_i$  vecinos del nodo

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

Como los enlaces son independientes y tienen la misma probabilidad  $p$ :

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

El coeficiente de clustering de un nodo de una red aleatoria es pequeño e independiente del grado del nodo

Esto sólo es válido para redes aleatorias, con una distribución del grado arbitraria

Dado un grado medio  $\langle k \rangle$  fijo,  $C$  decrece con el tamaño de la red  $N$

$$C_i \equiv \frac{1}{N} \frac{[\langle k^2 \rangle - \langle k \rangle]^2}{\langle k \rangle^3}$$

$$\begin{aligned} \langle k^2 \rangle &= \langle k \rangle (1 + \langle k \rangle) \\ \langle k^2 \rangle - \langle k \rangle &= \langle k \rangle^2 \end{aligned}$$

$$C_i \equiv \frac{1}{N} \langle k \rangle = p$$

- **Distribución de los grados:**  
*Binomial, Poisson (con colas exponenciales)*
- **Distancia media entre nodos:**  
*Logarítmicamente pequeña*
- **Coefficiente de Clustering:**  
Se anula para redes de gran tamaño

**¿Son aleatorias las  
redes reales?**

## MODELO DE RED ALEATORIA:

## ¿Son aleatorias las redes reales?

Durante cuatro décadas, el modelo de red aleatoria ha dominado la Ciencia de Redes

A finales de los 90, cuando se comenzó a disponer de datos cuantitativos de redes complejas reales, se pudo comparar su topología con las predicciones de la teoría de grafos aleatorios

En cuanto se disponga de  $N$  y  $\langle k \rangle$  para una red aleatoria, es posible derivar todas las propiedades medibles:

Distancia media:  $\langle d_{aleatoria} \rangle \approx \frac{\log N}{\log \langle k \rangle}$

Coefficiente de Clustering:  $C_{aleatoria} = p = \frac{\langle k \rangle}{N}$

Distribución del grado:  $P_{aleatoria}(k) = \binom{N-1}{k} p^k (1-p)^{(N-1)-k}$   $P_{aleatoria}(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$

## MODELO DE RED ALEATORIA:

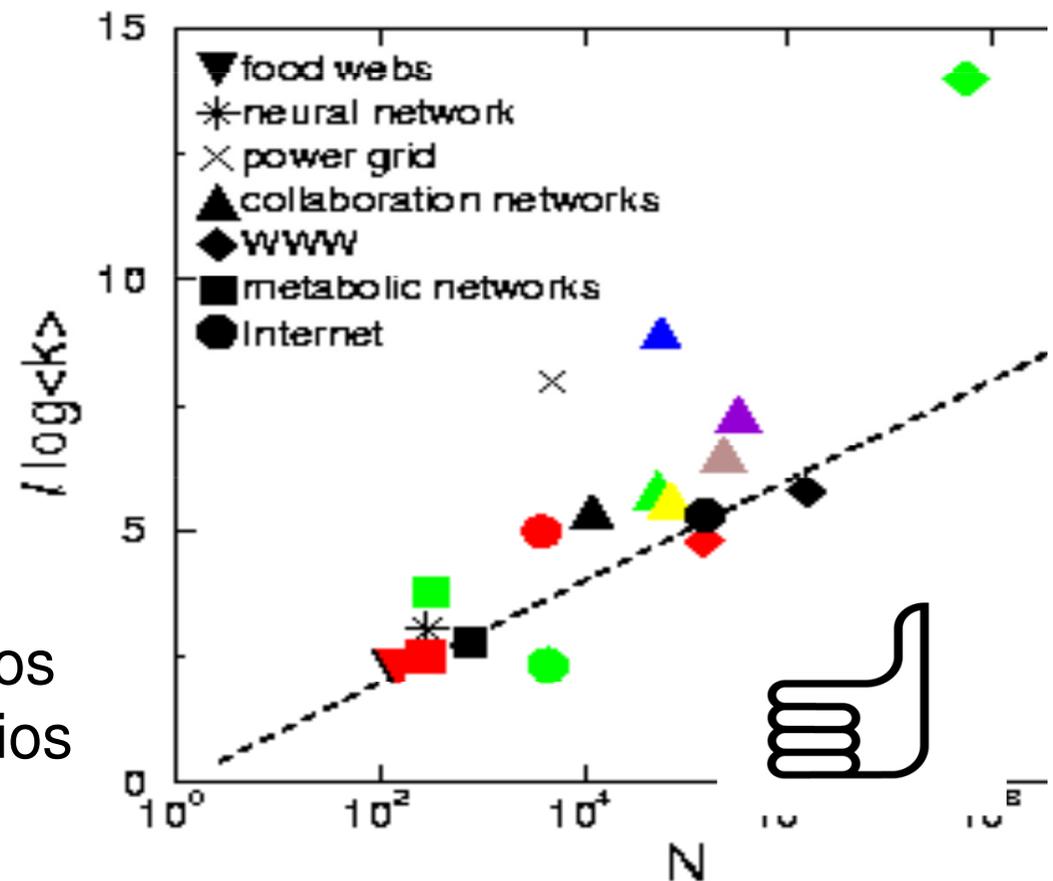
## Longitud de los Caminos

### Predicción del modelo:

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

Las redes reales tienen caminos cortos como los grafos aleatorios

### Datos:



## MODELO DE RED ALEATORIA:

## Coeficiente de Clustering (1)

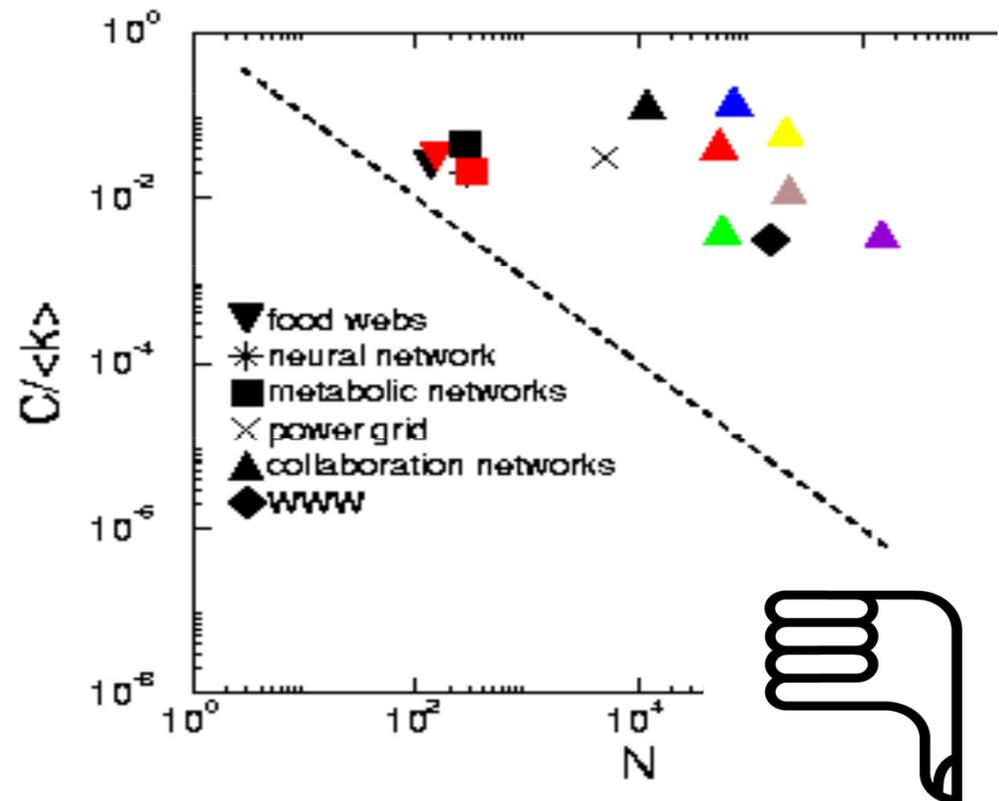
### Predicción:

$$C_{aleatoria} = \frac{\langle k \rangle}{N}$$

$C_{aleatoria}$  infra-estima el coeficiente de clustering de las redes reales en varios órdenes de magnitud

En las redes reales,  $C$  decrece con los grados de los nodos y es altamente independiente del tamaño de la red

### Datos:



## MODELO DE RED ALEATORIA:

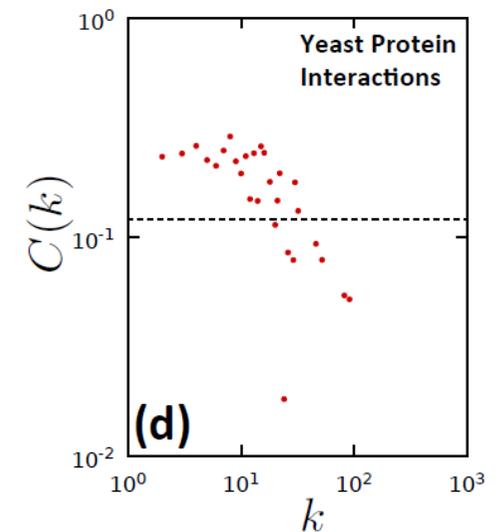
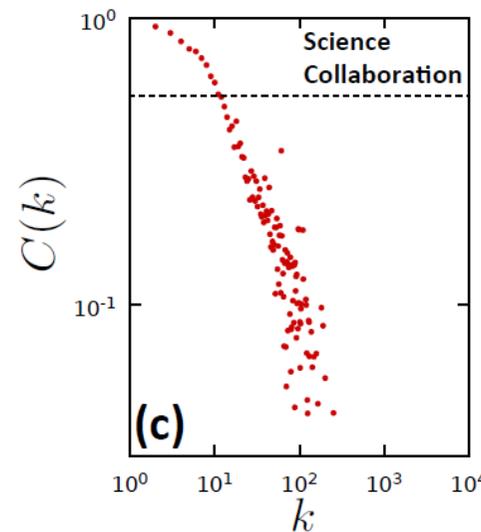
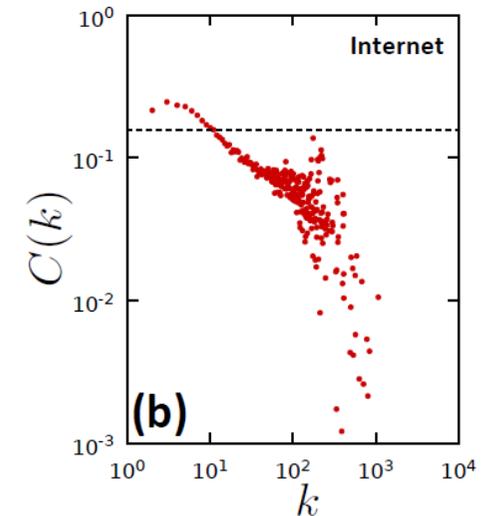
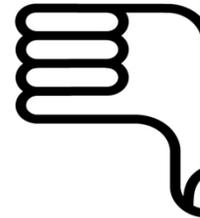
**Predicción:**

$$C_{\text{aleatoria}} = \frac{\langle k \rangle}{N}$$

En las redes reales,  $C$  decrece con los grados de los nodos y es altamente independiente del tamaño de la red

## Coeficiente de Clustering (2)

**Datos:**



# MODELO DE RED ALEATORIA:

# Distribución del grado

## Predicción:

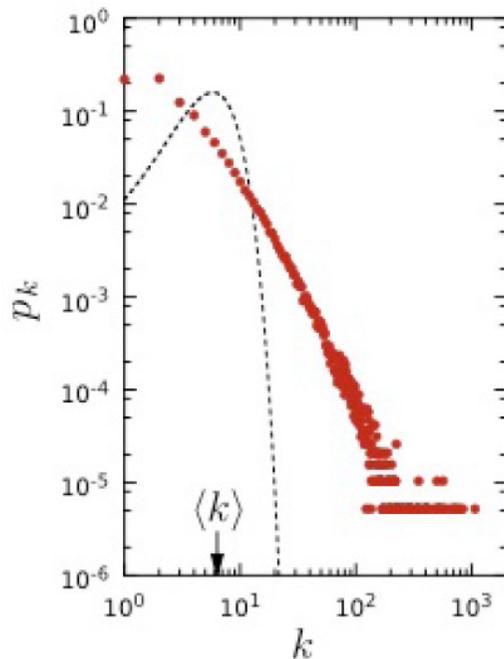
$$P_{aleatoria}(k) = \binom{N-1}{k} p^k (1-p)^{(N-1)-k}$$

## Datos:

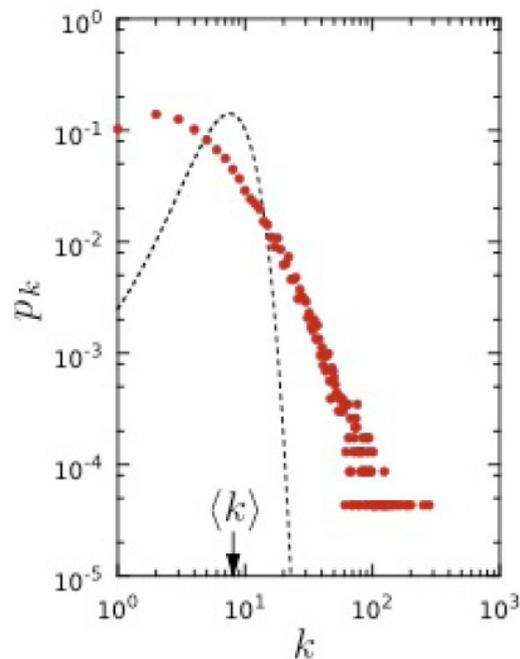
$$P(k) \approx k^{-\gamma}$$

$P_{rand}$  infra-estima el tamaño y la frecuencia de los nodos fuertemente conectados (hubs)

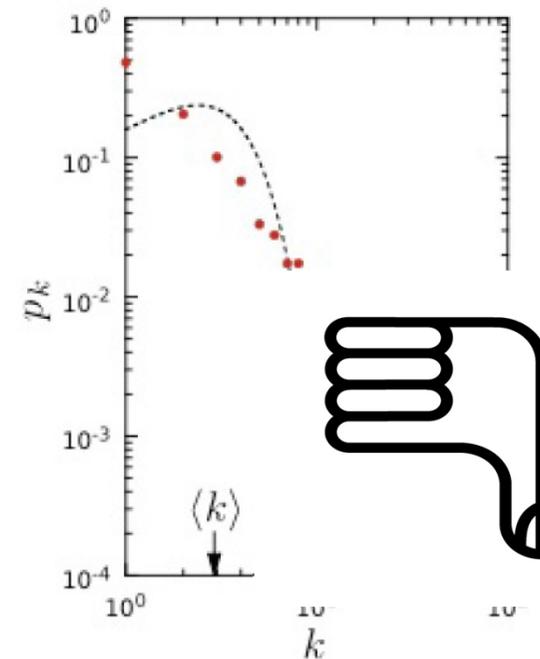
Internet



Science Collaboration



Protein Interactions



## MODELO DE RED ALEATORIA:

## Las redes reales son supercríticas (1)

Hay dos propiedades de la evolución de redes aleatorias que son especialmente importantes para las redes reales:

1. Cuando el grado medio supera el punto  $\langle k \rangle = 1$  emerge una componente gigante con una fracción pequeña de los nodos
2. Cuando  $\langle k \rangle > \log N$  la componente gigante incorpora todas las componentes aisladas y la red queda con una única componente conexa

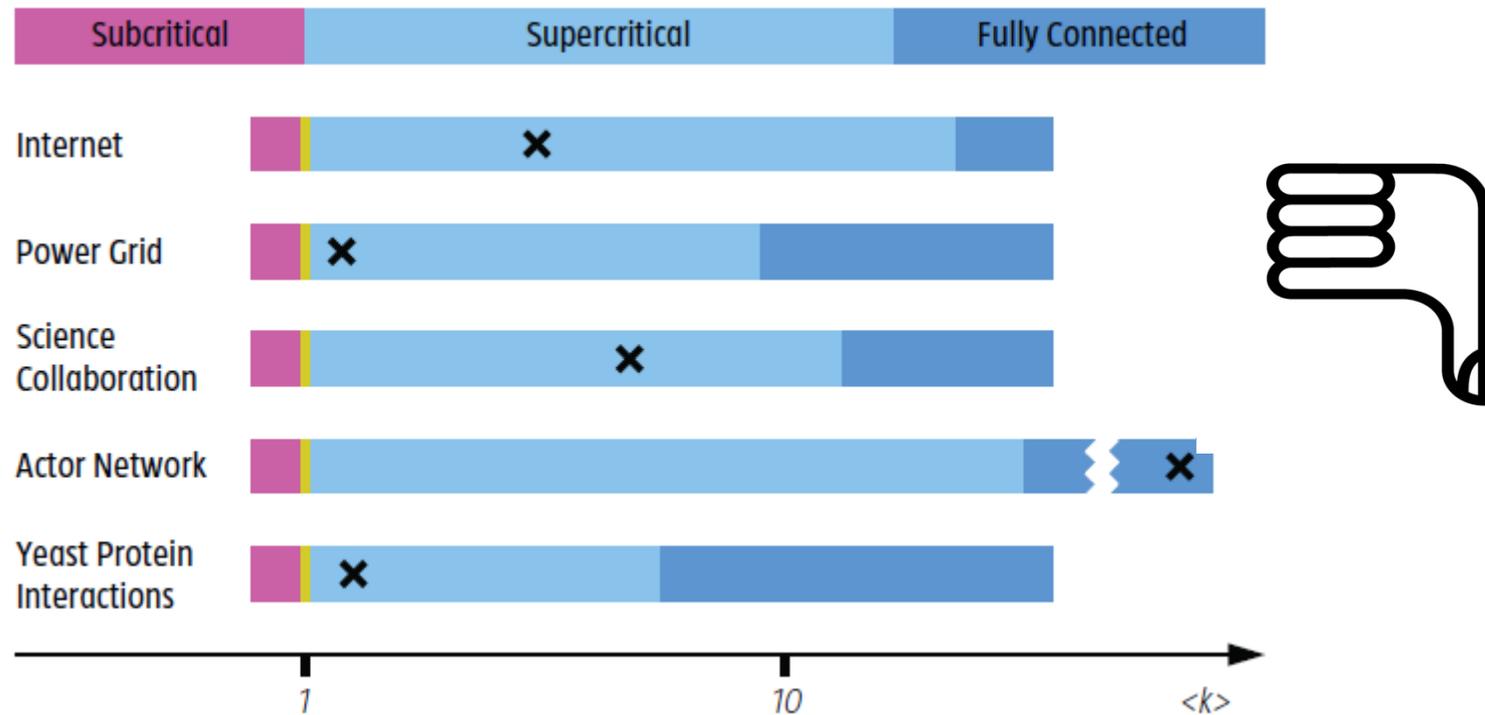
¿Se dan estas propiedades en las redes reales? ¿Aparece la componente gigante con  $\langle k \rangle = 1$ ?  
¿Contiene todos los nodos cuando  $\langle k \rangle > \log N$  o quedan nodos/componentes desconectados?

En muchas redes reales, el grado medio es muy superior a 1 y menor que  $\log N \rightarrow$  **régimen supercrítico**

Network	$N$	$L$	$\langle k \rangle$	$\ln N$
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

## MODELO DE RED ALEATORIA:

## Las redes reales son supercríticas (2)



De acuerdo a esto, las redes reales tienen una componente gigante que coexiste con muchas componentes y nodos aislados (**eso es extraño, sobre todo en la red física de Internet**)

Sin embargo, eso sólo sería así si las redes reales fueran aleatorias. En cambio, **las redes reales pueden permanecer conectadas a pesar de que no cumplen  $\langle k \rangle > \log N$**

## MODELO DE RED ALEATORIA:

## Corrección del modelo de red aleatoria

(A) Problemas con el modelo de red aleatoria:

- En redes reales, el coeficiente de clustering de los nodos depende de su grado
- La distribución de grados difiere de las de las redes reales
- La componente gigante de la mayoría de las redes reales NO emerge de una transición de fase

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

La respuesta es claramente: NO. Por tanto, el modelo es **IRRELEVANTE**

**A → ERRÓNEO**

**B → IRRELEVANTE**

**No existe ninguna red conocida en la naturaleza que pueda ser descrita por el modelo de red aleatoria**

## MODELO DE RED ALEATORIA:

## Utilidad del modelo de red aleatoria

**Y si es erróneo e irrelevante, ¿por qué se estudia?**

A pesar de todo, es un modelo de referencia. Posibilita calcular muchas medidas que pueden confrontarse con los datos reales, permitiendo entender hasta qué punto una propiedad particular es consecuencia de un proceso aleatorio

**Principios organizativos:** patrones que se producen en redes reales, observados en un gran número de ellas, y que se desvían de las predicciones del modelo de red aleatoria

Para identificarlos, necesitamos comprender cómo sería una propiedad particular si estuviese guiada por un proceso puramente aleatorio

Si la propiedad observada está ausente en las redes aleatorias, merece la pena estudiarla porque puede representar algún tipo de orden

**¡Aunque ERRÓNEO e IRRELEVANTE, el modelo resulta ser extremadamente ÚTIL!**

## MODELO DE RED ALEATORIA:

## Impacto en Ciencia de Redes



# RESUMEN DE PROPIEDADES DE LAS REDES ALEATORIAS

## At a glance: Random networks

- *Definition:*  $N$  nodes, where each node pair is connected with probability  $p$ .
- *Average degree:*  $\langle k \rangle = p(N-1)$
- *Average number of links:*  $\langle L \rangle = \frac{p(N-1)}{2}$
- *Degree distribution:*  $p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k}$ .

For sparse networks ( $k \ll N$ ),  $P_k$  has the Poisson form

$$p_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}.$$

- *Giant component ( $N_G$ ):*

$\langle k \rangle < 1$ : no giant component ( $N_G \sim \ln N$ )

$1 < \langle k \rangle < \ln N$ : one giant component and disconnected clusters

$$\left( N_G \sim N^{\frac{2}{3}} \right)$$

$\langle k \rangle > \ln N$ : all nodes join the giant component  $N_G \sim (p - p_i)N$

- *Average distance:*  $\langle d \rangle \propto \frac{\log N}{\log \langle k \rangle}$ ,
- *Clustering coefficient:*  $C = \frac{\langle k \rangle}{N}$ .

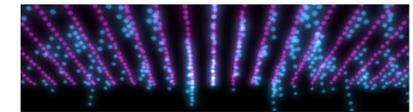
**THE FIRST LAW: SMALL WORLD PROPERTY**  
**In complex networks there are short distances between any pair of nodes.**

# Referencias y Agradecimientos

Para diseñar los materiales de este tema, he hecho uso de material desarrollado por expertos en el área disponible en Internet:

- “Network Science Interactive Book Project” del Laszlo Barabasi Lab:

<http://barabasilab.com/networksciencebook>



- Curso on-line “Social Network Analysis” de Lada Adamic, Coursera, Universidad de Michigan: <https://www.coursera.org/course/sna>

