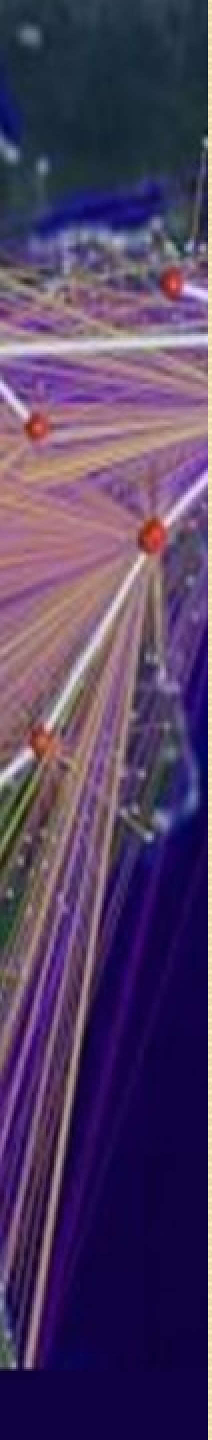
LECTURE 5: BASED ON CHAPTER 8 OF AUXILIARY BOOK

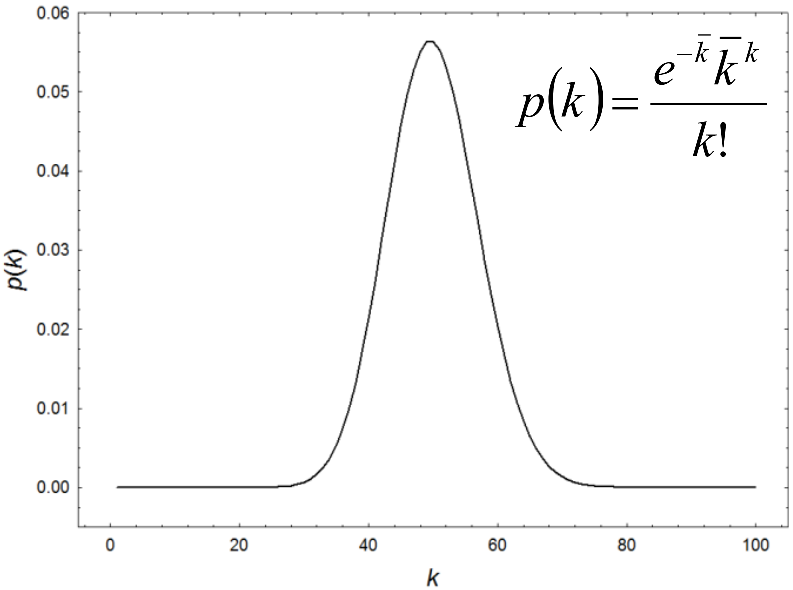
Degree Irregularity. Distributions vs. Algebraic Approaches

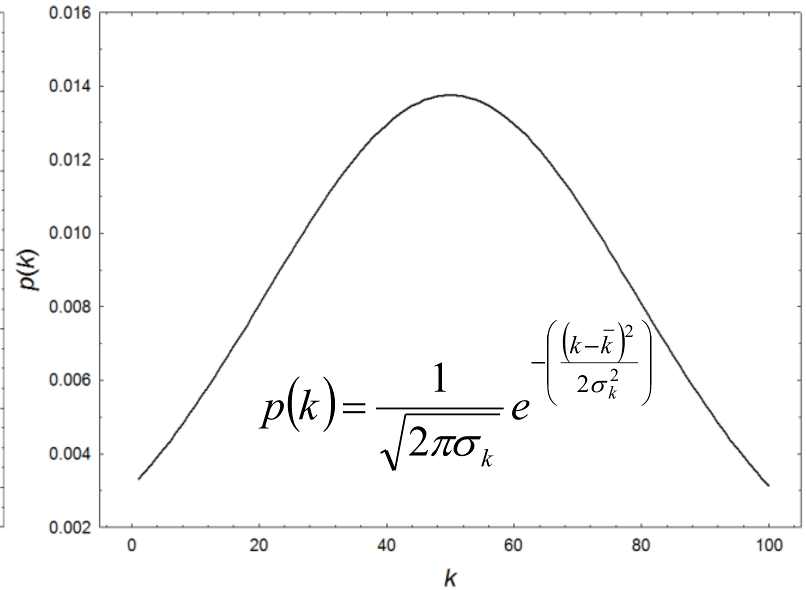
Professor Ernesto Estrada
Department of Mathematics & Statistics
University of Strathclyde
Glasgow, UK

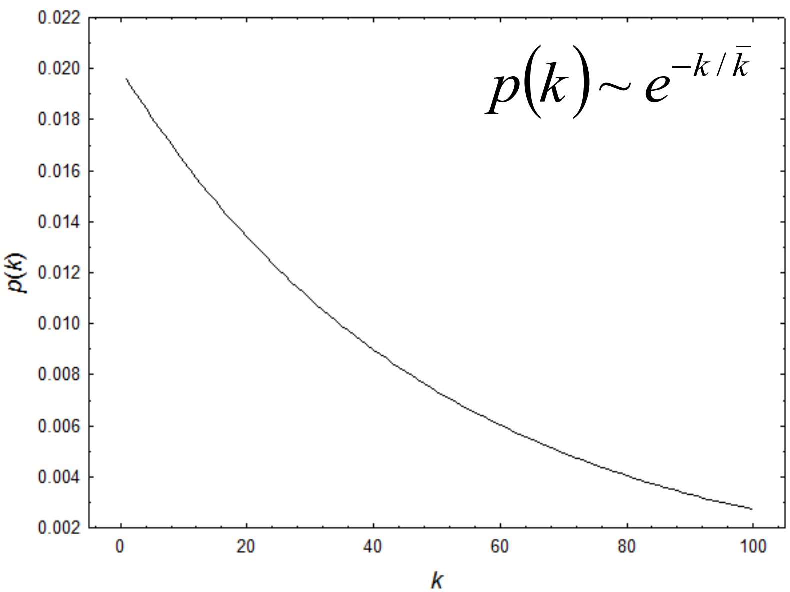
www.estradalab.org

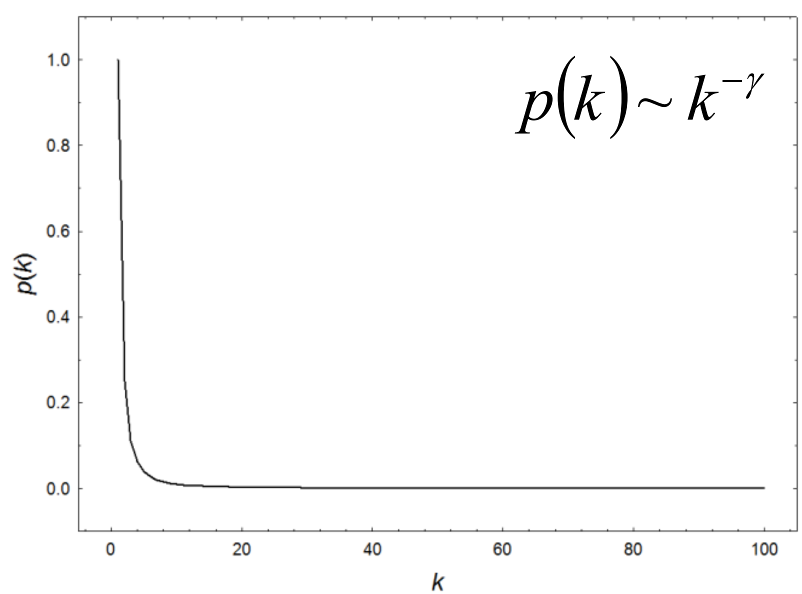


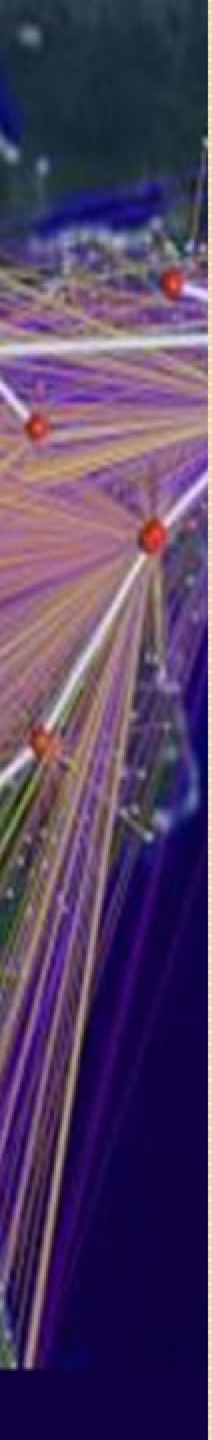
Statistical Distributions





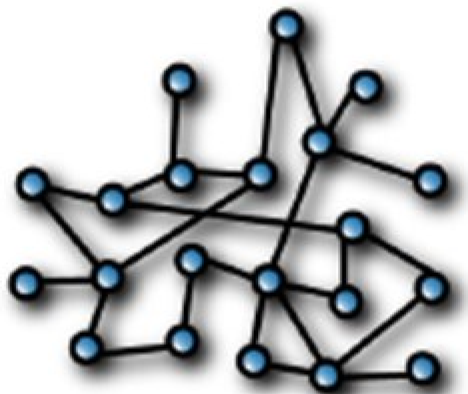


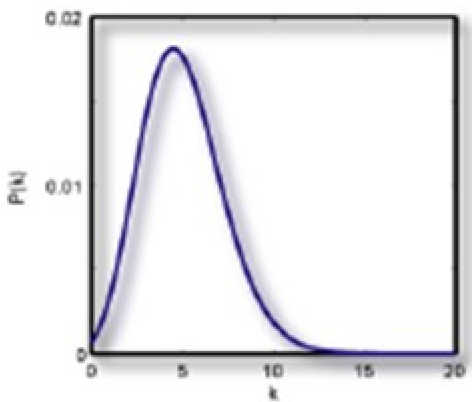




Degree Distributions

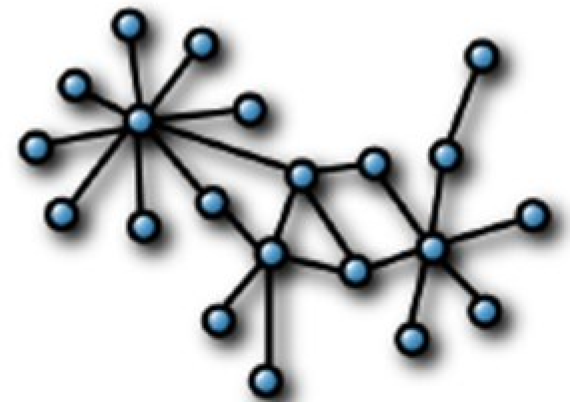
Poisson

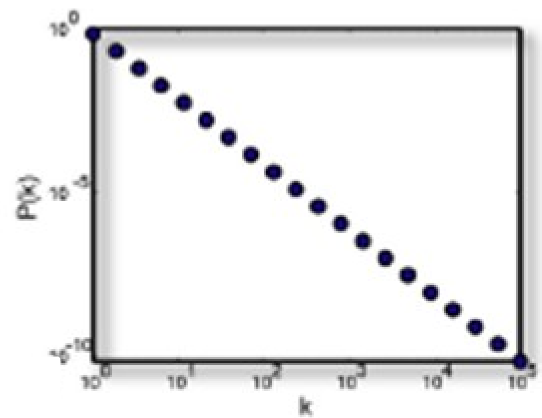




$$p(k)=\frac{e^{-\bar{k}}\,\bar{k}^{k}}{k!}$$

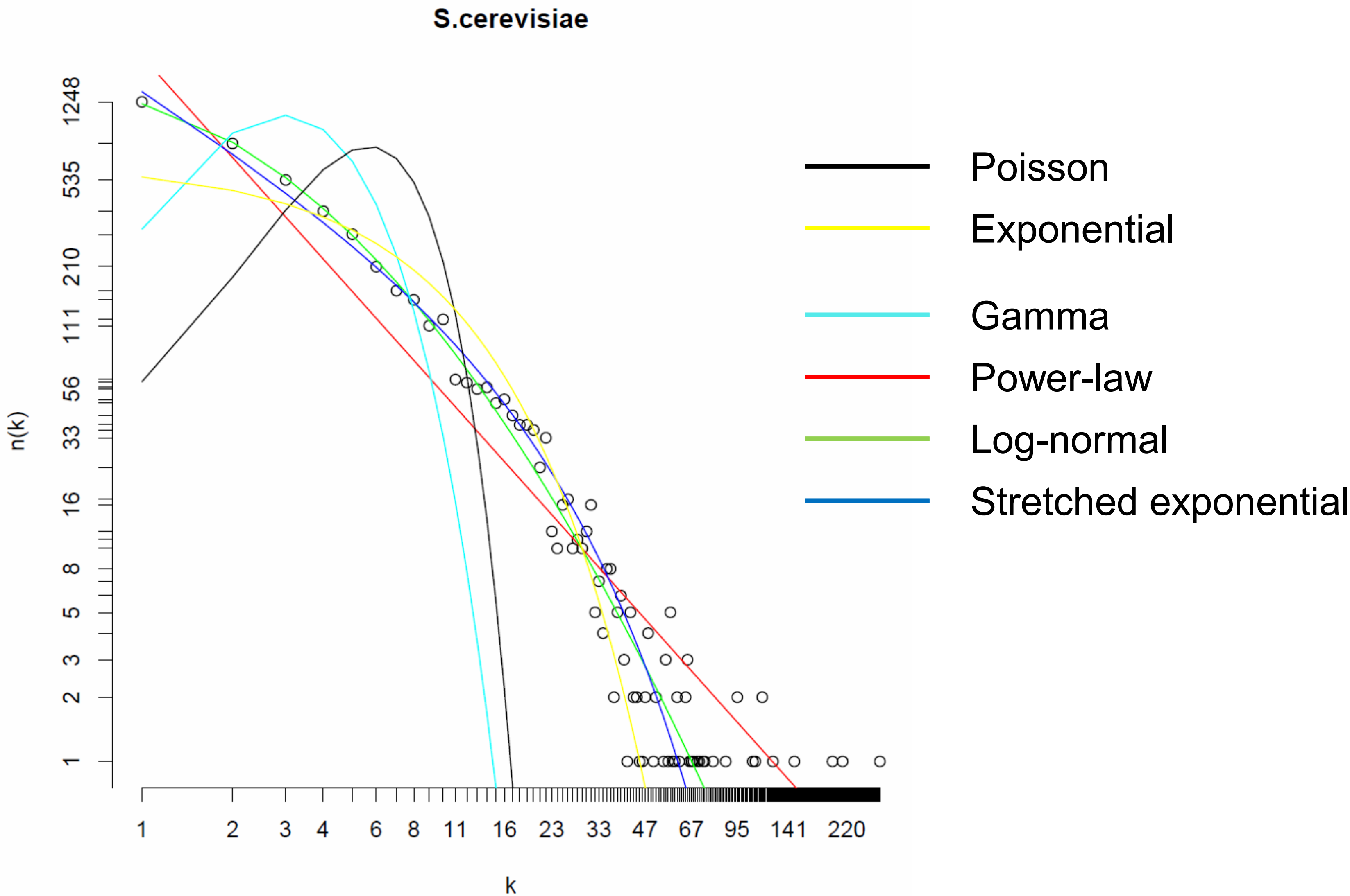
Power-law



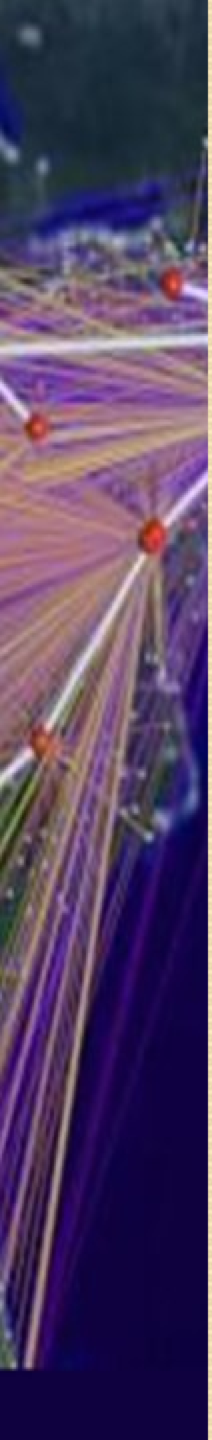


$$p(k)\sim k^{-\gamma}$$

What is the best fit?



Stumpf & Ingram, *Europhys. Lett*. **71** (2005) 152.



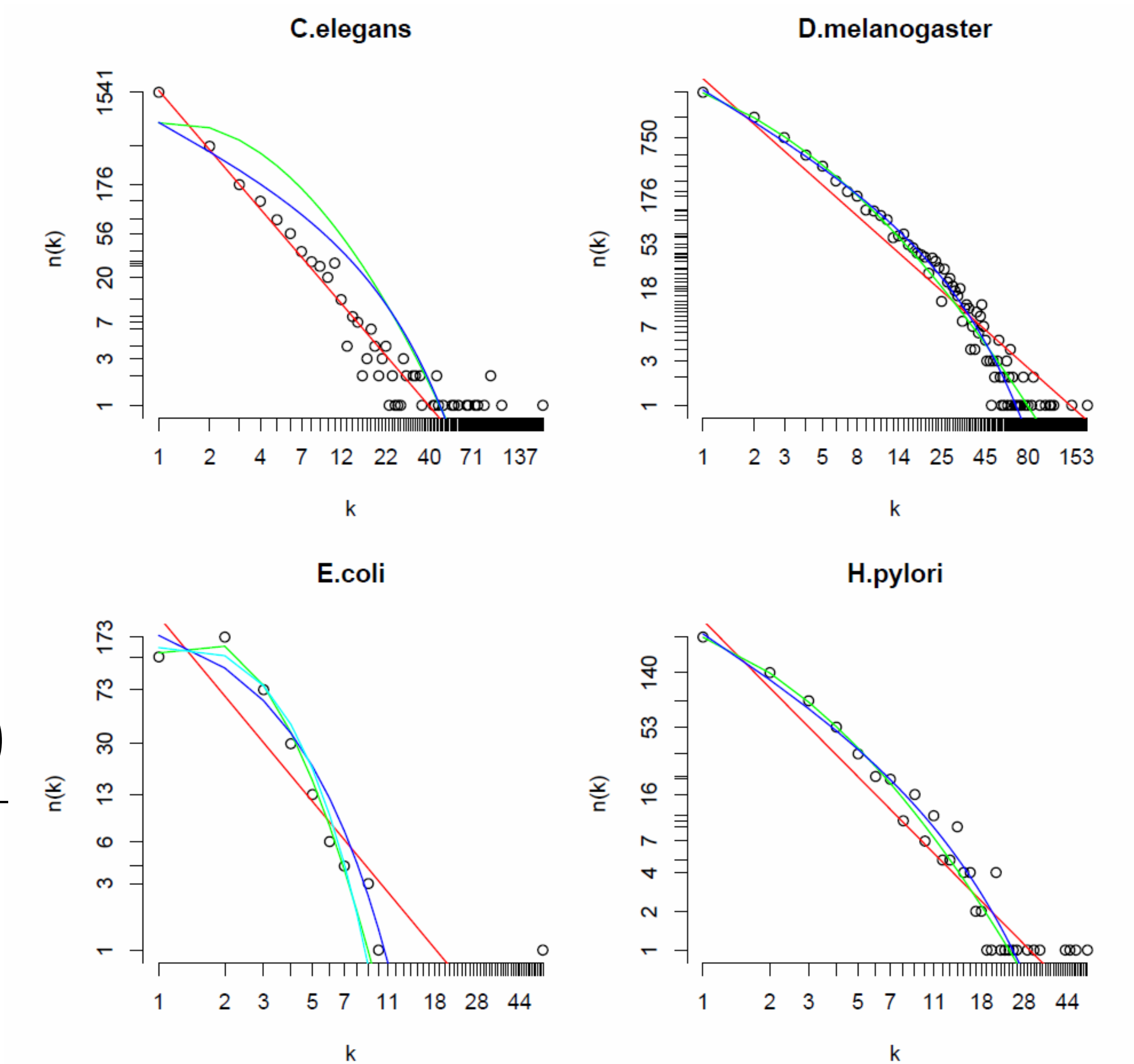
The Zoo of Distributions

- Bernoulli ....
- Beta ....
- Four-parameter Beta...
- Beta-Binomial.............
- Beta-Geometric............................
- Beta-Negative Bininomial ....
- Binomial.......................................
- Bradford.......................................
- Burr..
- Cauchy ....
- Chi...
- Chi-Square(d) ....
- Cumulative Ascending ....
- Cumulative Descending ....
- Dagum ....
- Delaporte ....
- Discrete...
- Discrete Uniform ....
- Error Function ....
- m-Erlang.......................................
- Generalized Error ....
- Exponential ....
- Extreme Value Max ....
- Extreme Value Min............
- F...
- Fatigue Life........................
- Gamma ....
- Geometric ....
- Generalized Logistic ....
- Histogram ....
- Hyperbolic Secant ....
- Hypergeometric ....
- Inverse Gaussian ....
- Inverse Hypergeometric .....
- Johnson Bounded..............
- Johnson Unbounded ....
- Kumaraswamy..........................
- Four-parameter Kumaraswamy..
- Laplace.....................................
- Levy ....
- Logarithmic (Series)..................
- LogGamma ....
- Logistic.....................................
- LogLaplace ....
- LogLogistic ....
- Lognormal ....
- LognormalB..............................
- LognormalE ....
- Modified PERT..........................
- Negative Binomial ....
- Normal.....................................
- Ogive..
- Pareto ....
- Pareto2 ....
- Pearson Type 5.........
- Pearson Type 6.........
- PERT........................
- Poisson....................
- Pólya ....
- Rayleigh ....
- Reciprocal ....
- Relative ....
- Slash ....
- Split Triangle ....
- Step Uniform...........
- Student ....
- Triangle....................
- Uniform...................
- Weibull.................

How to Compare?

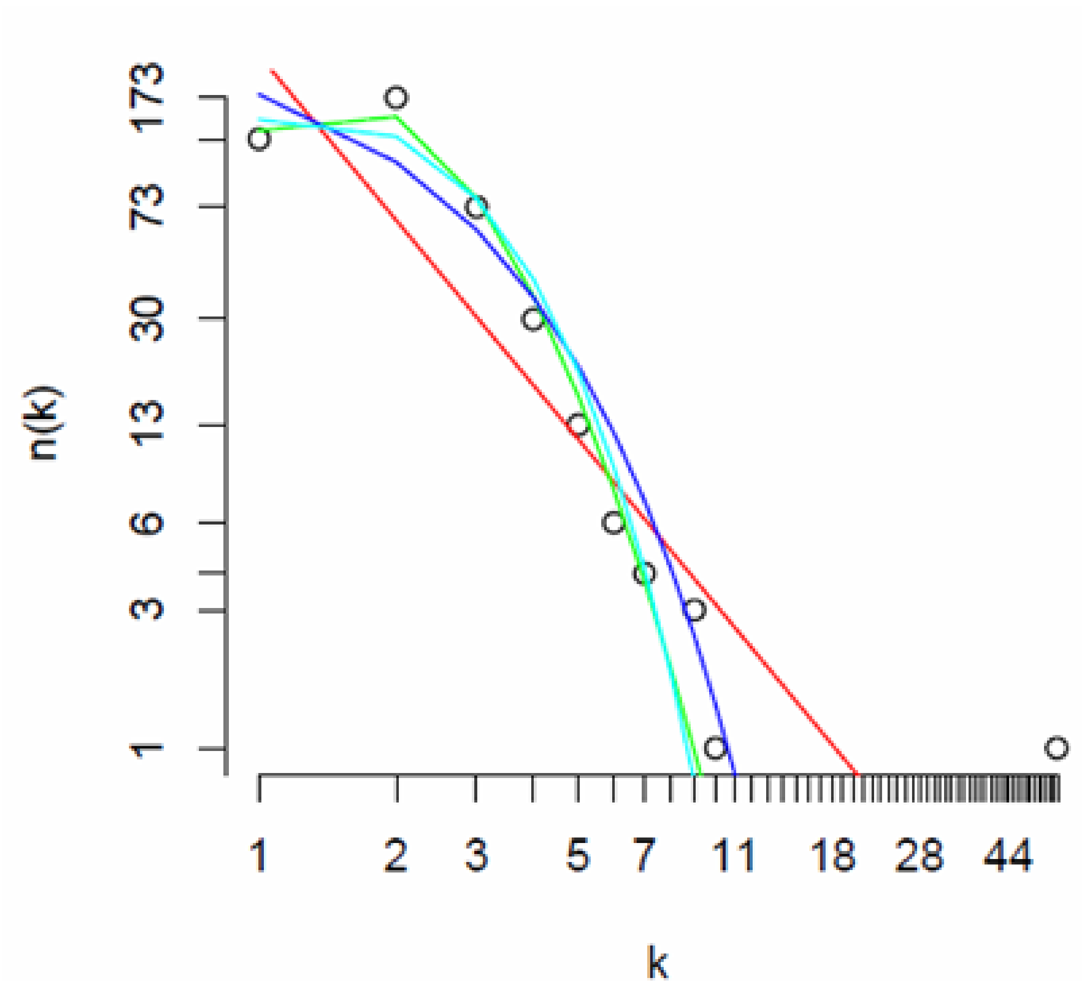
$$p(k) \sim e^{-\alpha k/\bar{k}} k^{-\gamma}$$

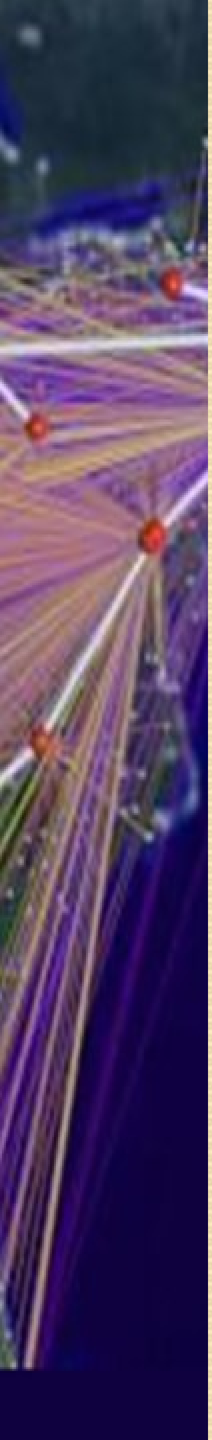
$$p(k) \sim \frac{e^{-\ln((k-\theta)/m)^2/(2\sigma^2)}}{(k-\theta)\sigma\sqrt{2\pi}}$$



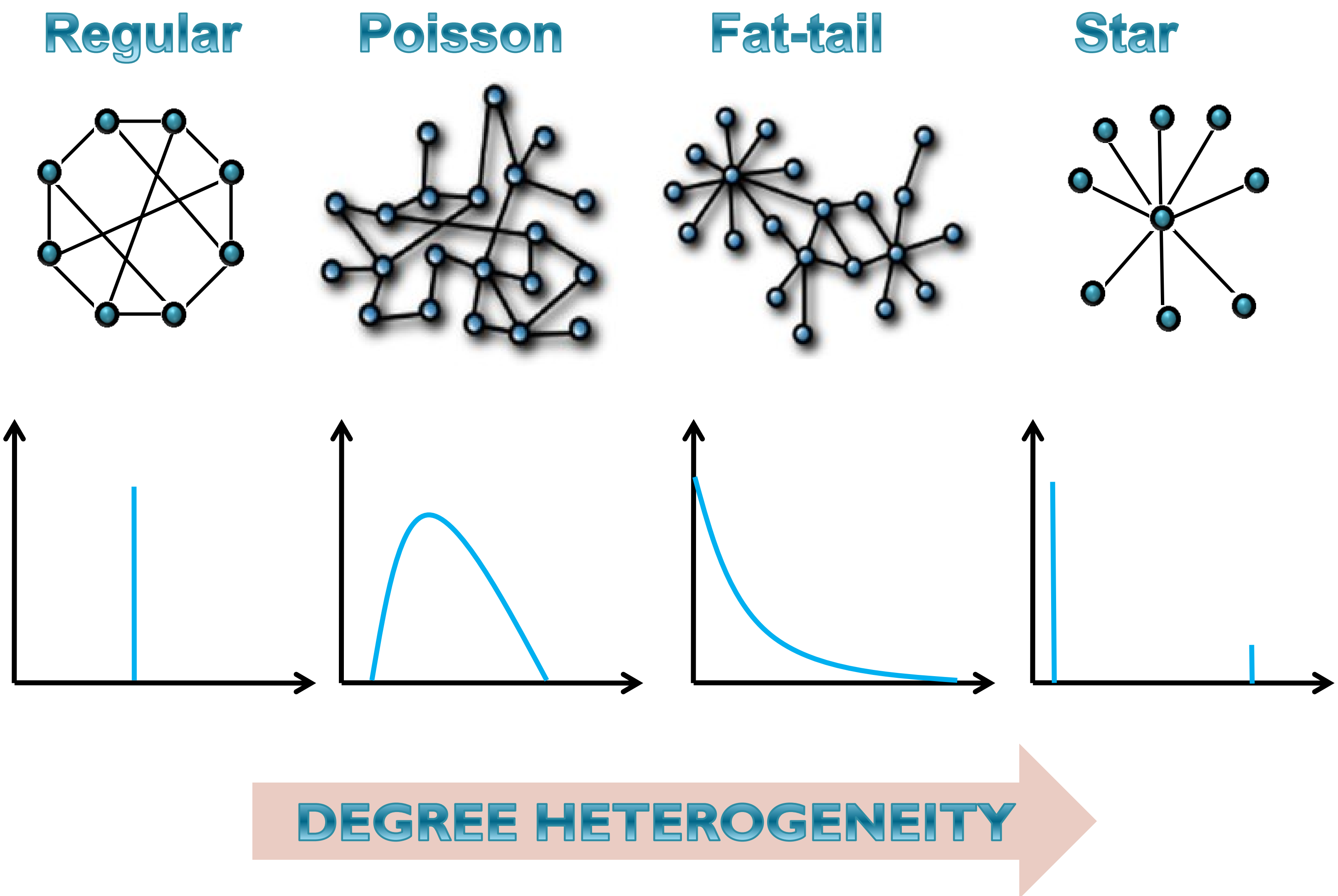
Stumpf & Ingram, *Europhys. Lett*. **71** (2005) 152.

How to Manage Scarce Data?

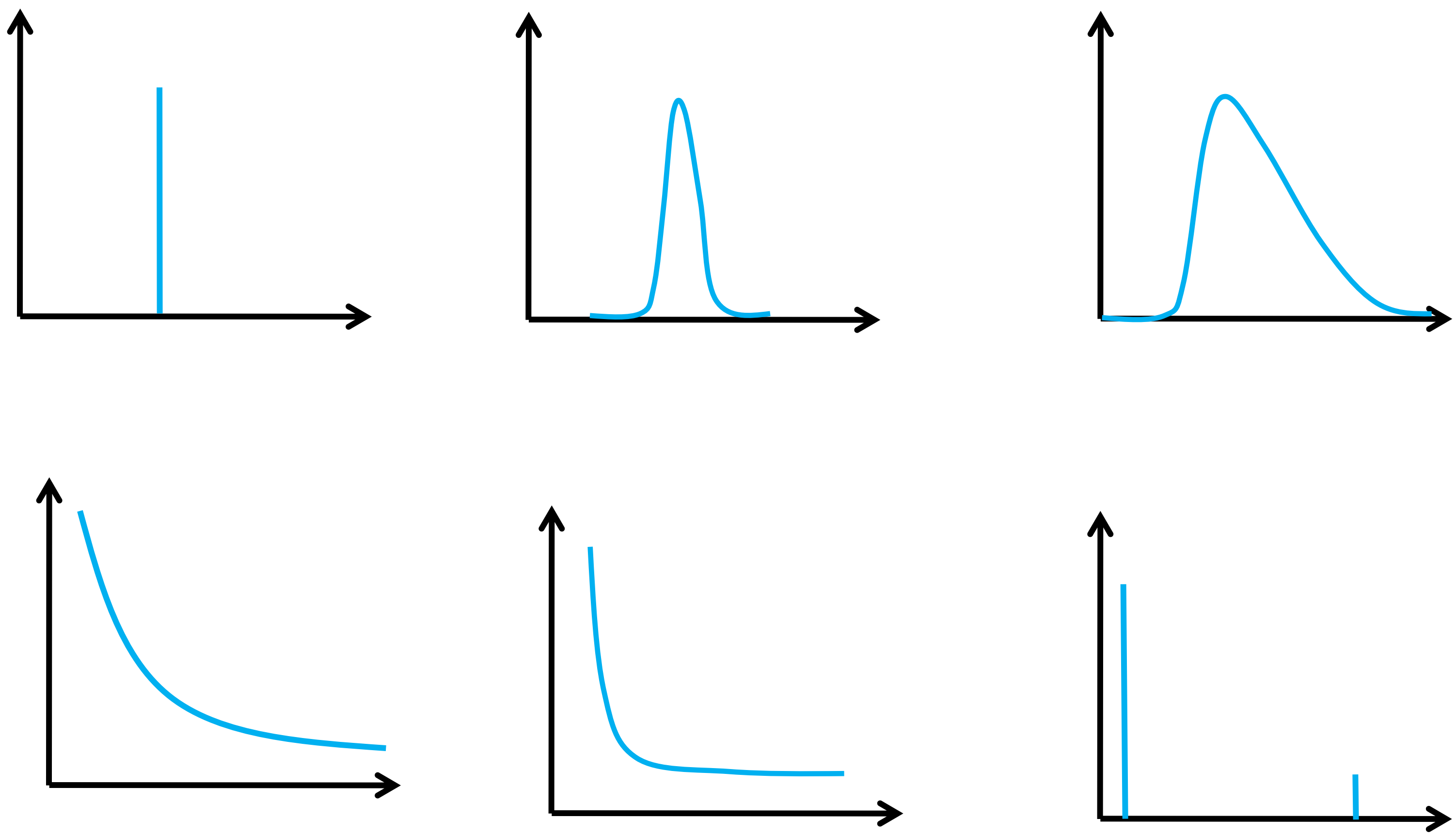


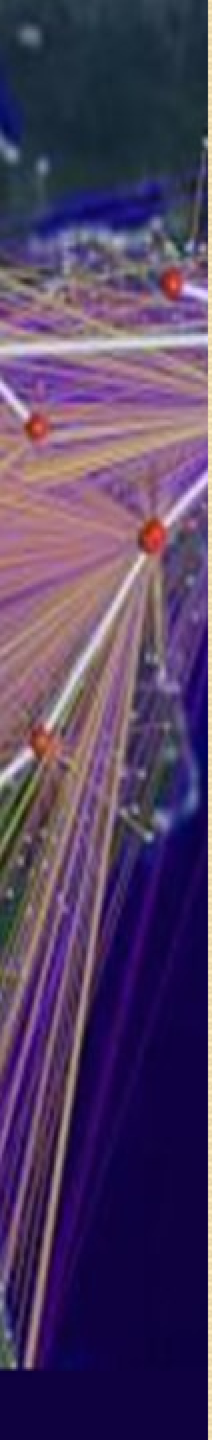


Is There an Order?



Is There an Order?





Graph Theory Approach

Degree irregularity indices

- **Collatz-Sinogowitz Index (1957)**

$$CS(G) = \lambda_1 - \langle k \rangle$$

- **Bell Index (1992)**

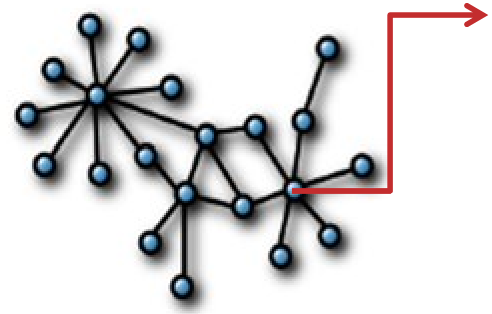
$$Var(G) = \frac{1}{n}\sum_i k_i^2 - \left(\frac{1}{n}\sum_i k_i\right)^2$$

- **Albertson Index (1997)**

$$irr(G) = \sum_{(i,j)\in E} \left|k_i - k_j\right|$$

A New Index: Degree Skewness

probability of picking at random an incident edge



$$p_i = \frac{1}{k_i}$$

Non-standardised skewness: $\rho_{ij} = \langle p_{ij} \rangle - \hat{p}_{ij}$

Arithmetic mean

$$\langle p_{ij} \rangle = \frac{1}{2}\left(\frac{1}{k_i} + \frac{1}{k_j} \right)$$

geometric mean

$$\hat{p}_{ij} = \left(\frac{1}{\sqrt{k_i k_j}} \right)$$

Estrada: *Phys. Rev. E* **82**, 066102 (2010)

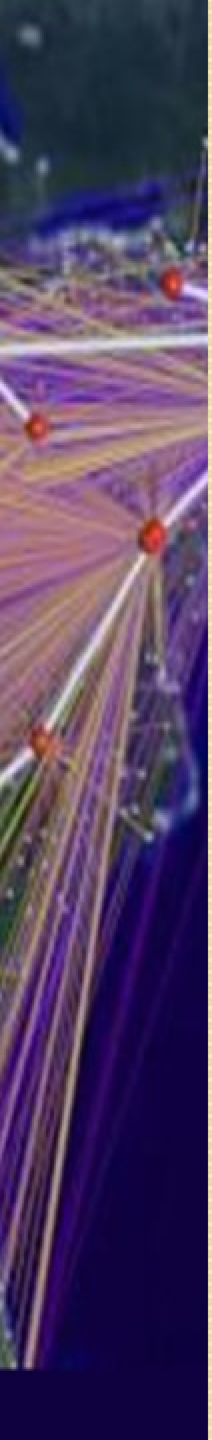
Degree Skewness

$$2\rho_{ij} = \left(\frac{1}{k_i} + \frac{1}{k_j} \right) - \frac{2}{\sqrt{k_i k_j}}$$

$$= \left(\frac{1}{\sqrt{k_i}} - \frac{1}{\sqrt{k_j}} \right)^2$$

Network degree skewness

$$\rho(G) = 2 \sum_{(ij)\in E} \rho_{ij} = \sum_{(ij)\in E} \left(\frac{1}{\sqrt{k_i}} - \frac{1}{\sqrt{k_j}} \right)^2$$



The Laplacian Encounter

Graph Laplacian

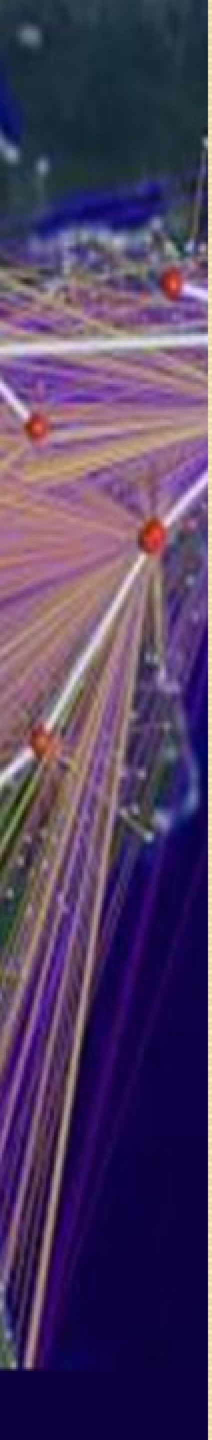
$$L = K - A$$

$$L_{ij} = \begin{cases} k_i & \text{for } i = j, \\ -1 & \text{for } i \sim j, \\ 0 & \text{otherwise,} \end{cases}$$

Normalised Graph Laplacian

$$\mathcal{L} = K^{-1/2} \cdot L \cdot K^{-1/2}$$
$$= I - K^{-1/2} \cdot A \cdot K^{-1/2}$$

$$\mathcal{L}_{ij} = \begin{cases} 1 & \text{for } i = j, \\ -\dfrac{1}{\sqrt{k_i k_j}} & \text{for } i \sim j, \\ 0 & \text{otherwise,} \end{cases}$$



The Laplacian Encounter

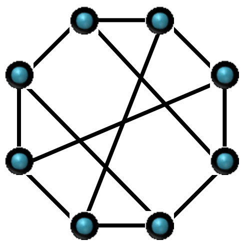
Network degree skewness

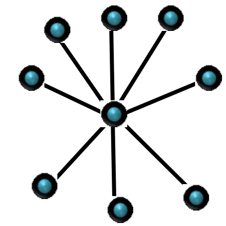
$$\rho(G) = \vec{1} \cdot \mathcal{L} \cdot \vec{1}^{T}$$
$$= n - 2 \sum_{(i,j) \in E} \left(k_i k_j\right)^{-1/2}$$
$$= n - 2\chi(G).$$

Randic index

A Nice Normalisation

$$\sqrt{n-1} \le \chi(G) \le \frac{n}{2}$$

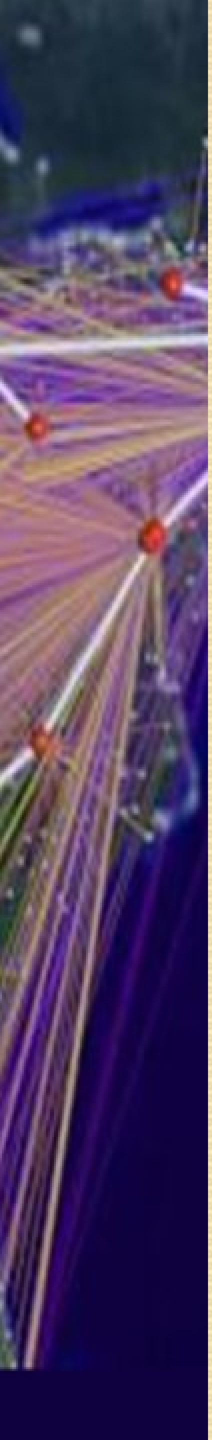




Normalised network degree skewness

$$\tilde{\rho}(G) = \frac{n - 2\chi(G)}{n - 2\sqrt{n-1}}$$

$$0 \le \tilde{\rho}(G) \le 1$$



Back to the Normalised Laplacian

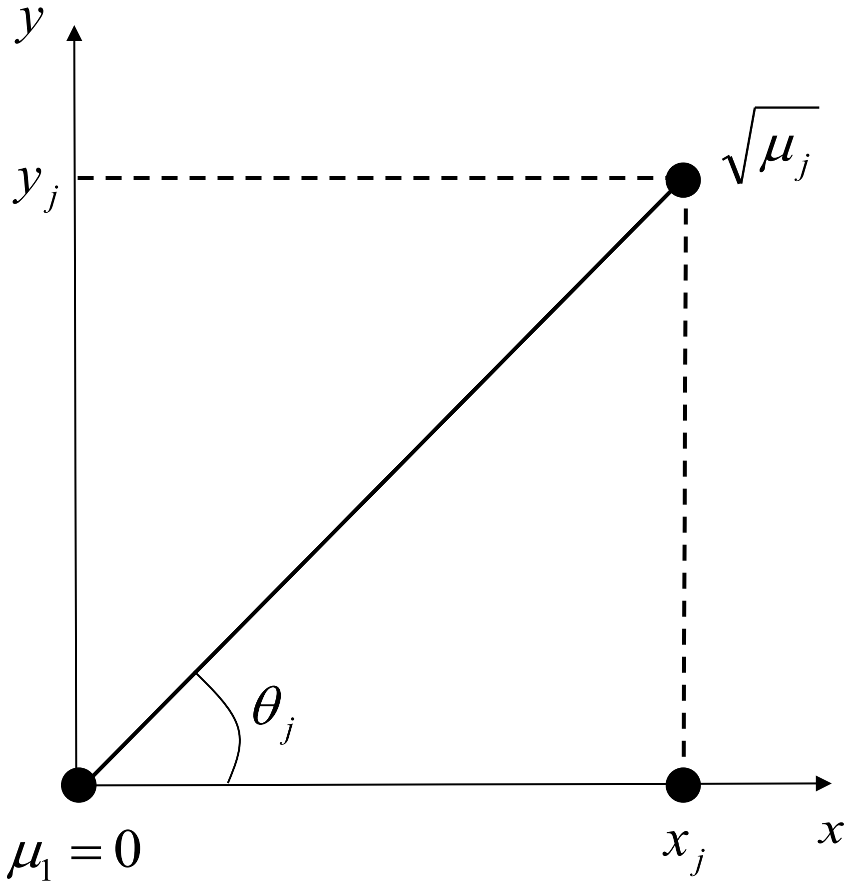
$$\rho(G) = \vec{1}^T \cdot \mathscr{L} \cdot \vec{1} = n \sum_j \mu_j \cos^2 \theta_j$$

$$0 = \mu_1 < \mu_2 \leq \cdots \leq \mu_n$$

$$\cos\theta_j = \frac{1^T \cdot \vec{\varphi}_j}{\|1\| \|\vec{\varphi}_j\|} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \varphi_j(i)$$

$$\rho(G) = \vec{1}^T \cdot \mathscr{L} \cdot \vec{1} = \sum_j \mu_j \left(\sum_i \varphi_j(i) \right)^2$$

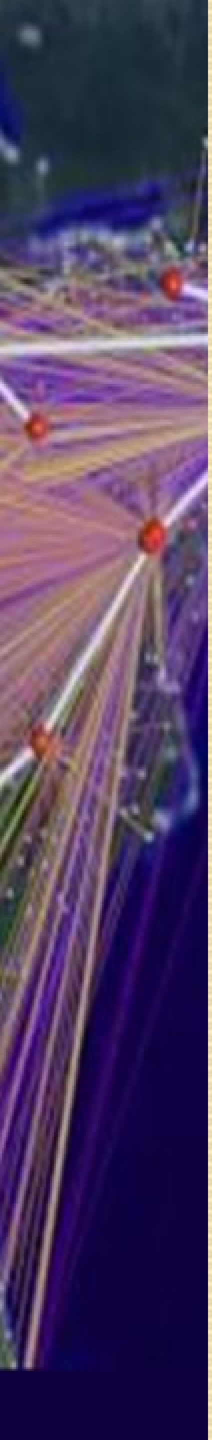
Spectral Representation



$$x_j = \sqrt{\mu_{j>1}}\cos\theta_j$$

$$y_j = \sqrt{\mu_{j>1}}\sin\theta_j$$

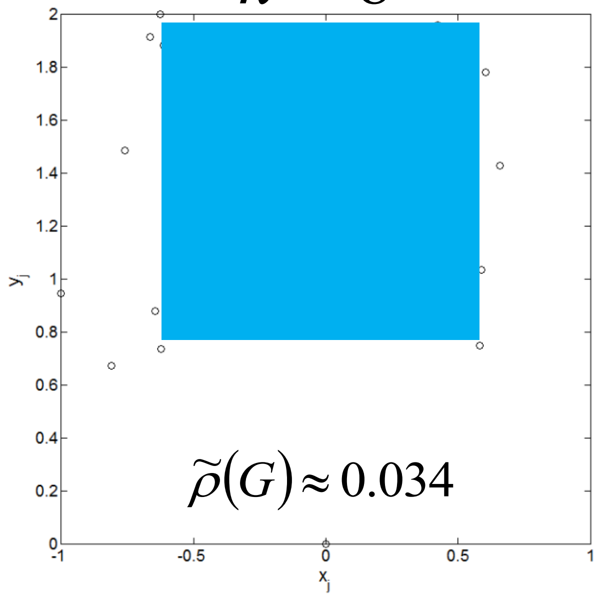
$$\tilde{\rho}(G) = \frac{n}{n-2\sqrt{n-1}}\sum_{j=1}^{n} x_j^2$$



Results

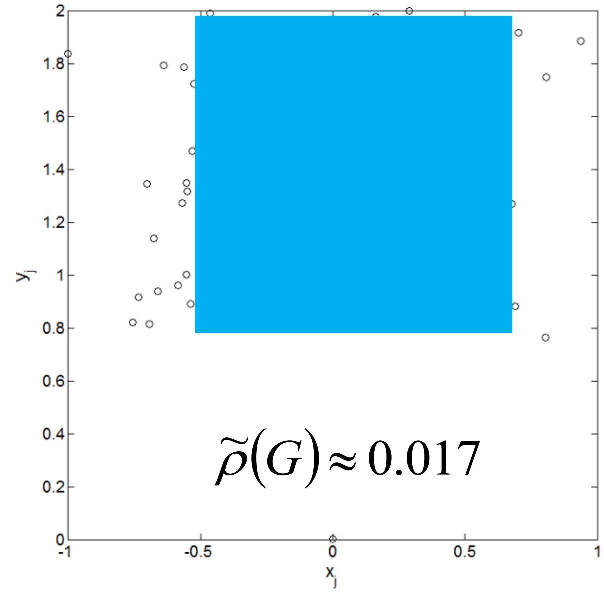
Erdös-Rényi Random Graphs $G_{n,p}$

$\bar{k} = 8$



$\tilde{\rho}(G) \approx 0.034$

$\bar{k} = 16$

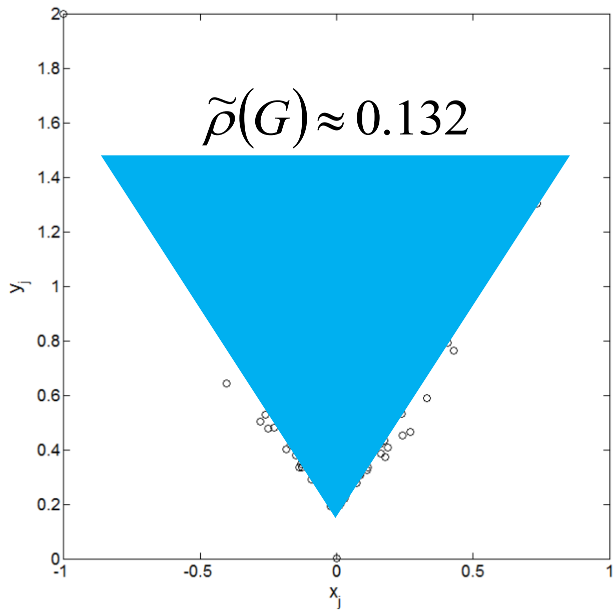


$\tilde{\rho}(G) \approx 0.017$

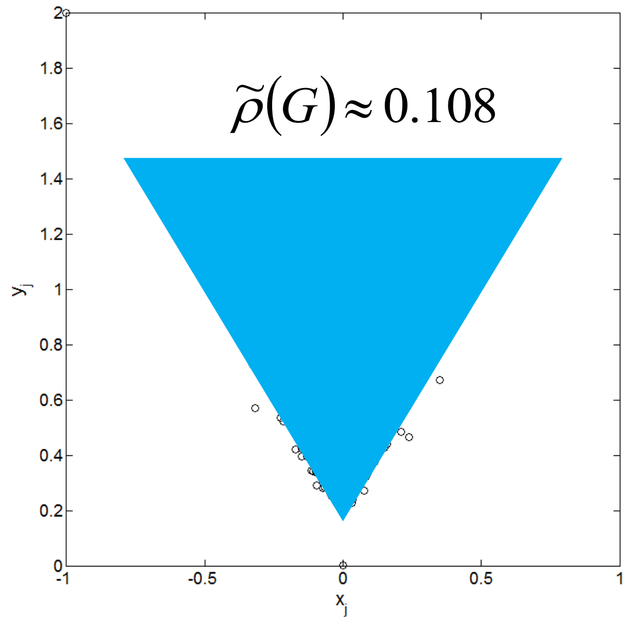
Barabási-Albert Random Graphs

$$\rho(BA) \approx \left(\frac{0.27 + \langle k \rangle^{-1.18}}{2.27 + \langle k \rangle^{-1.18}}\right)\left(\frac{n}{n - 2\sqrt{n-1}}\right)$$

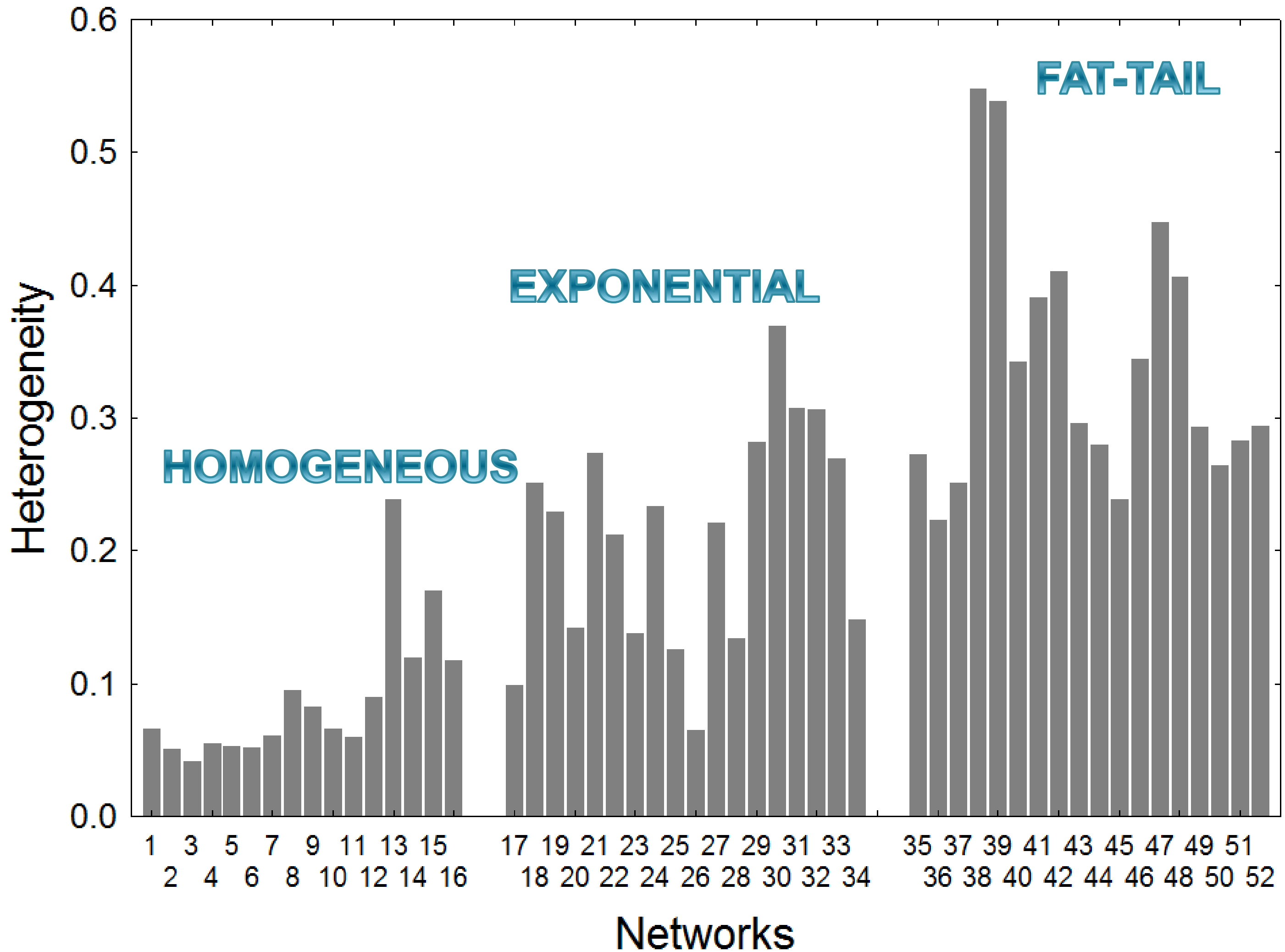
$\bar{k} = 8$

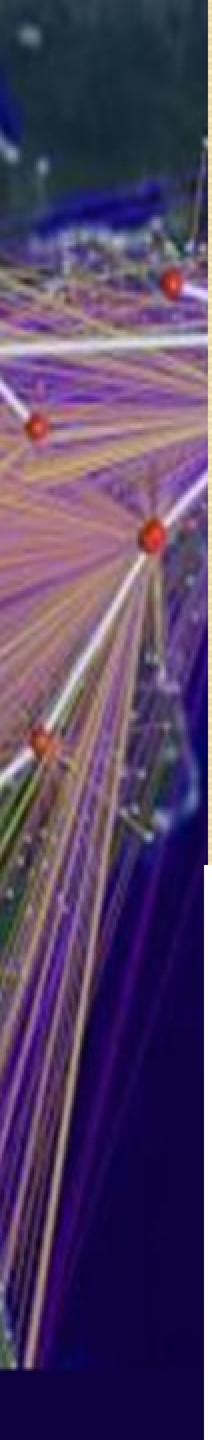


$\bar{k} = 16$



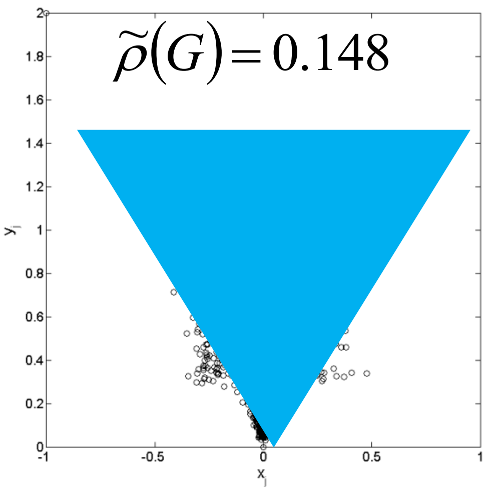
Real-World Networks



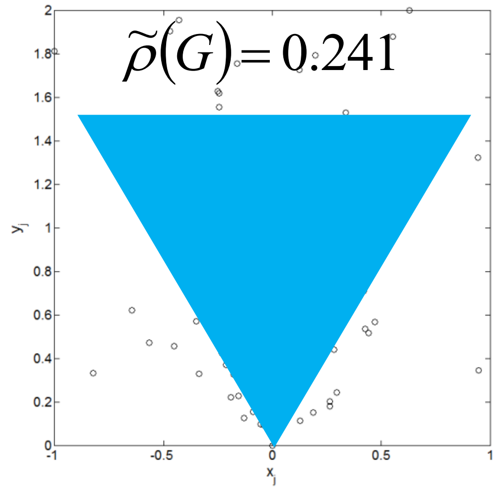


Real-World Networks

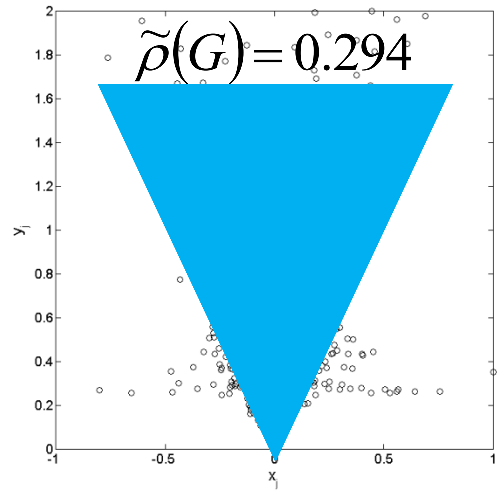
D. melanogaster



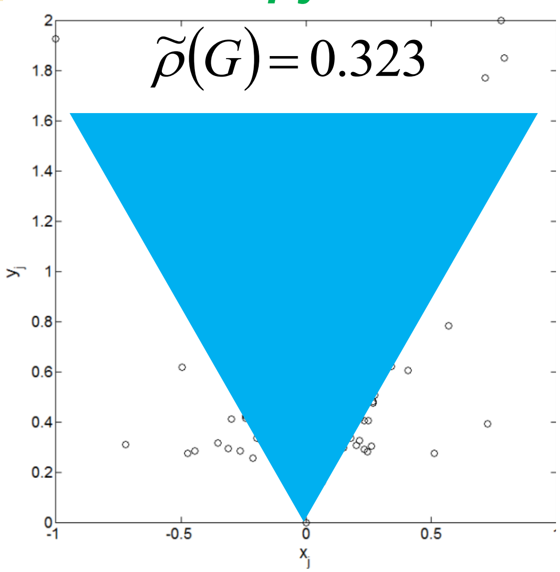
E. coli



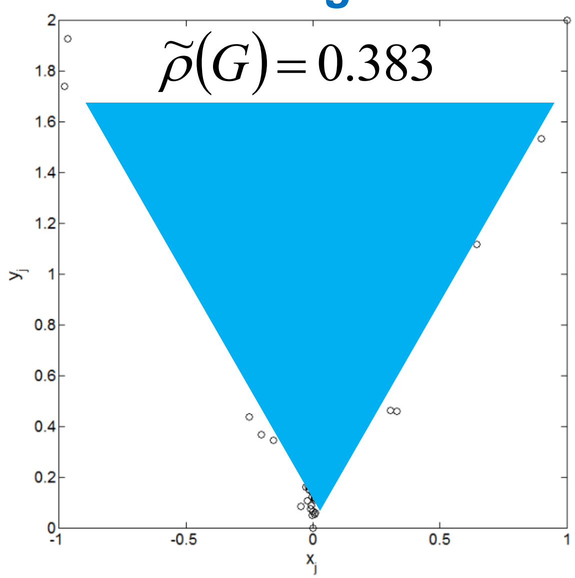
S. cereviciae



H. pylori



C. elegans



Stretched exponential

$$p(k) \sim e^{-\alpha k / \bar{k}} k^{-\gamma}$$

Log-normal

$$p(k) \sim \frac{e^{-\ln((k-\theta)/m)^2 / (2\sigma^2)}}{(k-\theta)\sigma\sqrt{2\pi}}$$

Degree Assortativity. Statistical vs. Combinatorial Approaches

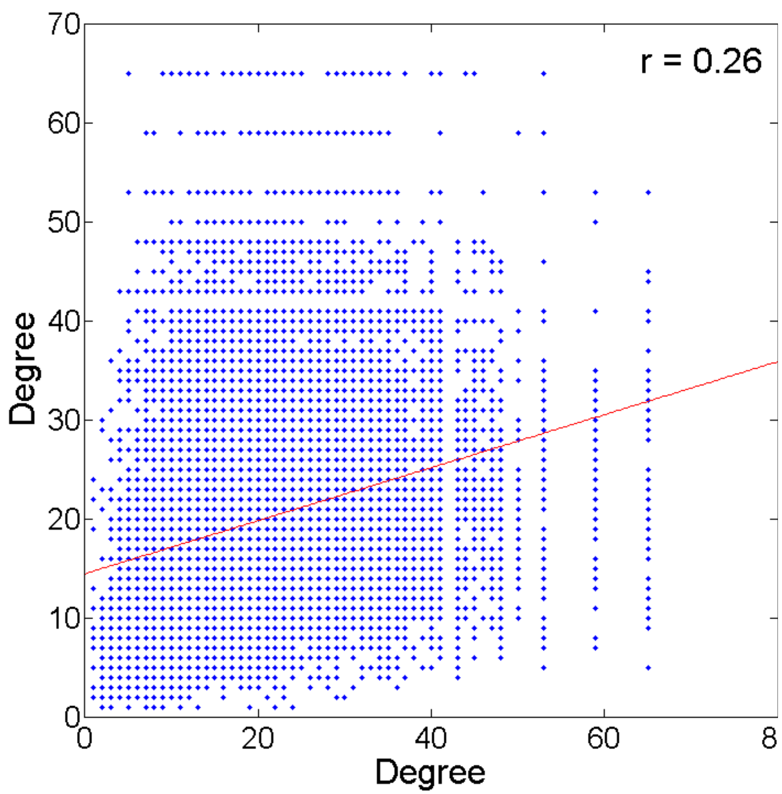
Professor Ernesto Estrada
Department of Mathematics & Statistics
University of Strathclyde
Glasgow, UK

www.estradalab.org

Degree-degree Correlation

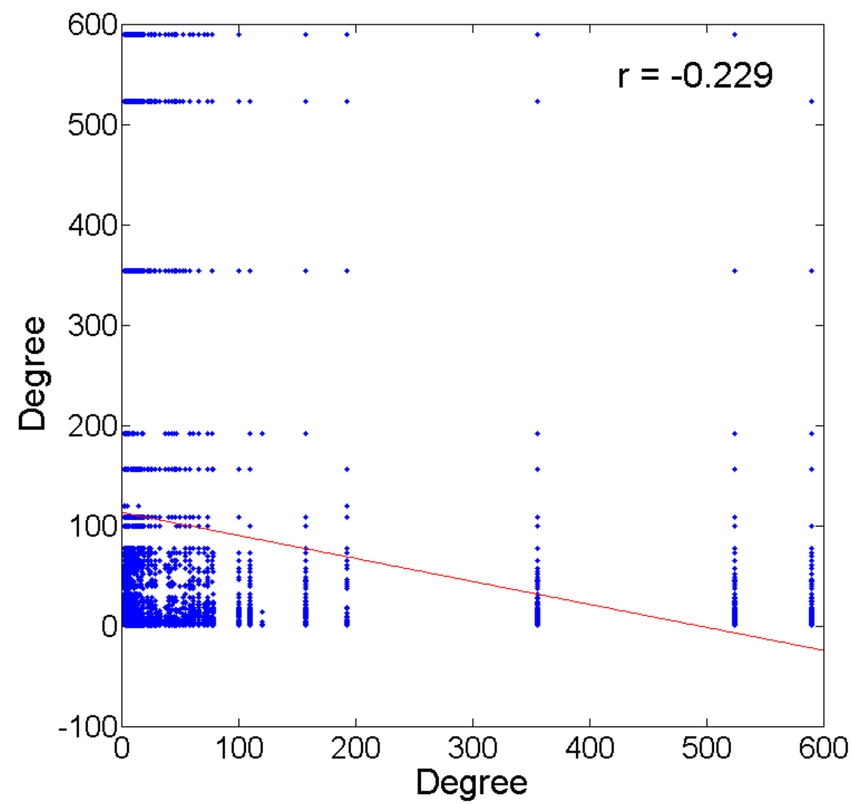
Statistical approach

SOCIAL NETWORK



Assortative

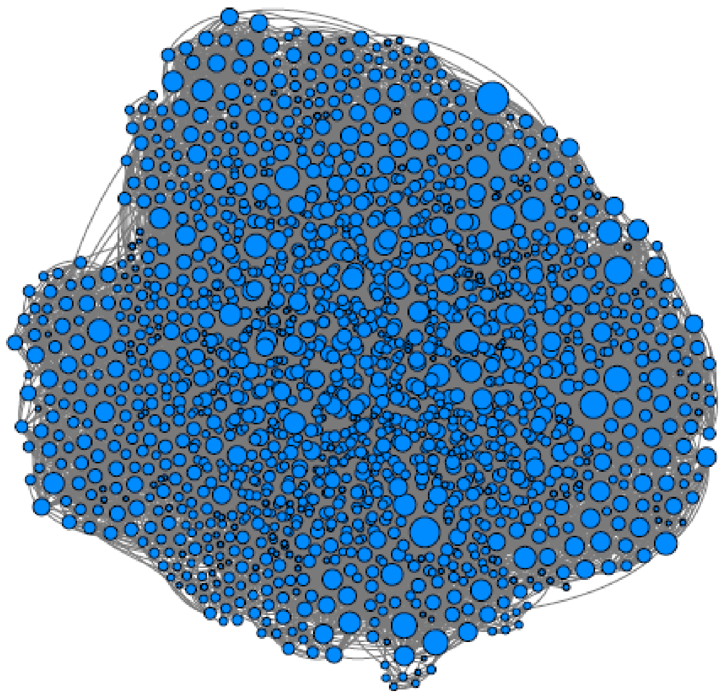
INTERNET



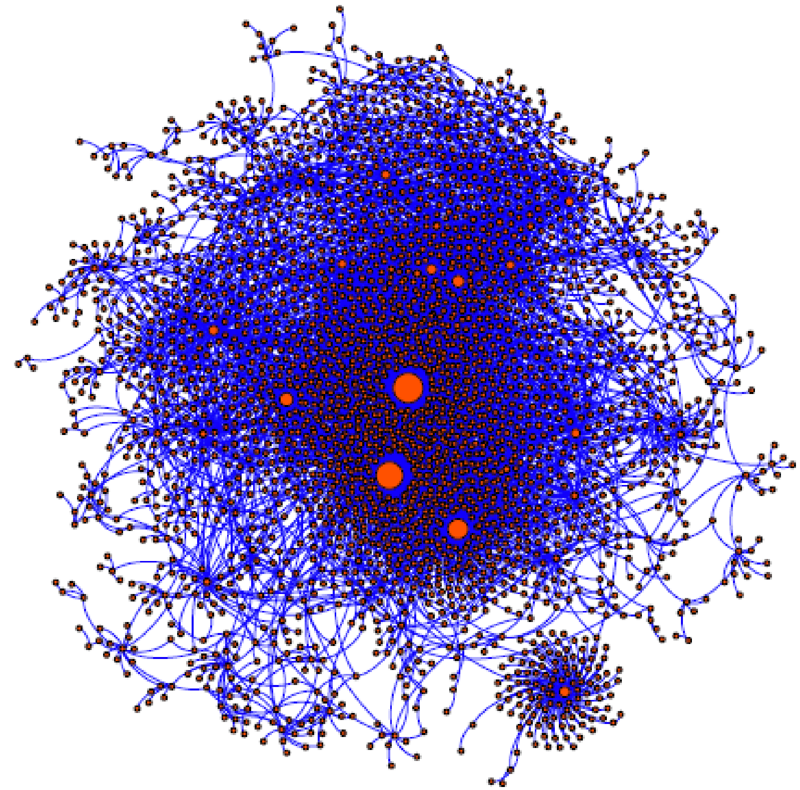
Disassortative

Degree-degree Correlation

Example

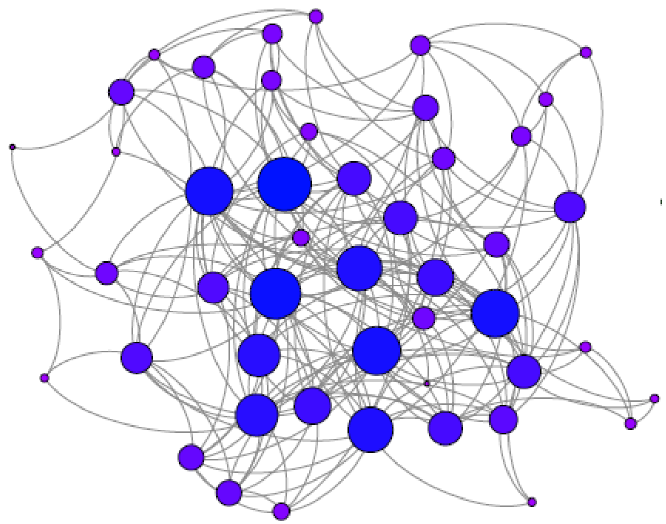


Assortative

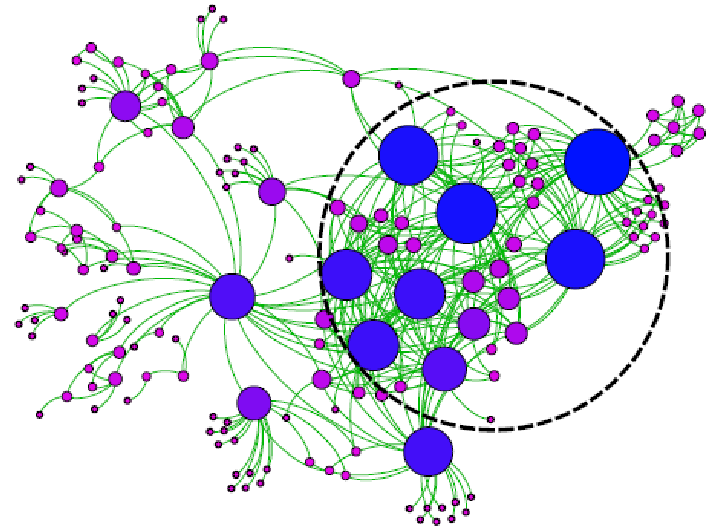


Disassortative

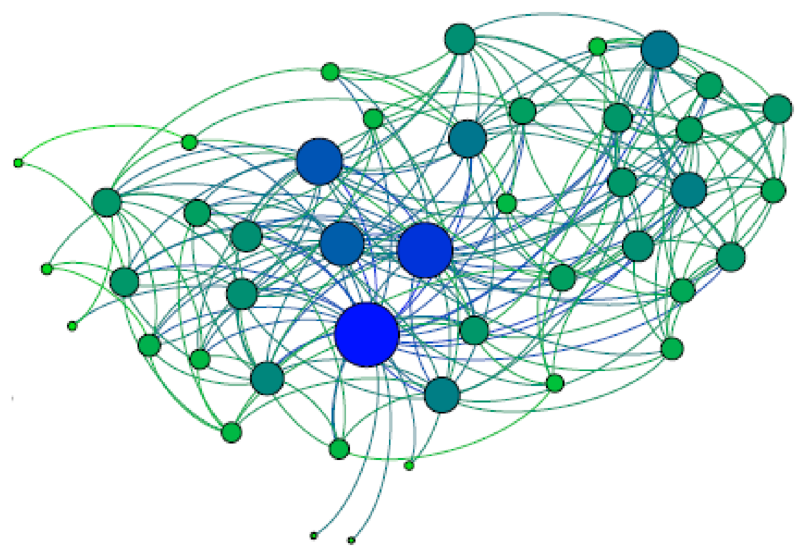
Degree-degree Correlation



$r = 0.118$



$r = -0.304$



$r = -0.153$

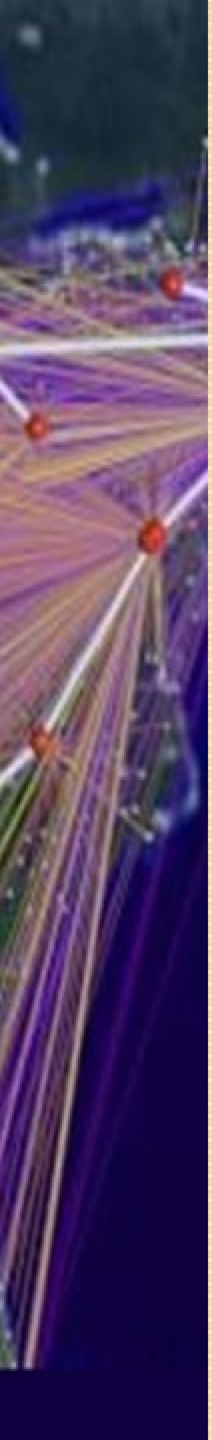
Degree-degree Correlation

Statistical approach

$$r=\frac{\frac{1}{m}\sum_{(i,j)\in E}k_ik_j-\left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i+k_j\right)\right)^2}{\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i^2+k_j^2\right)-\left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i+k_j\right)\right)^2}$$

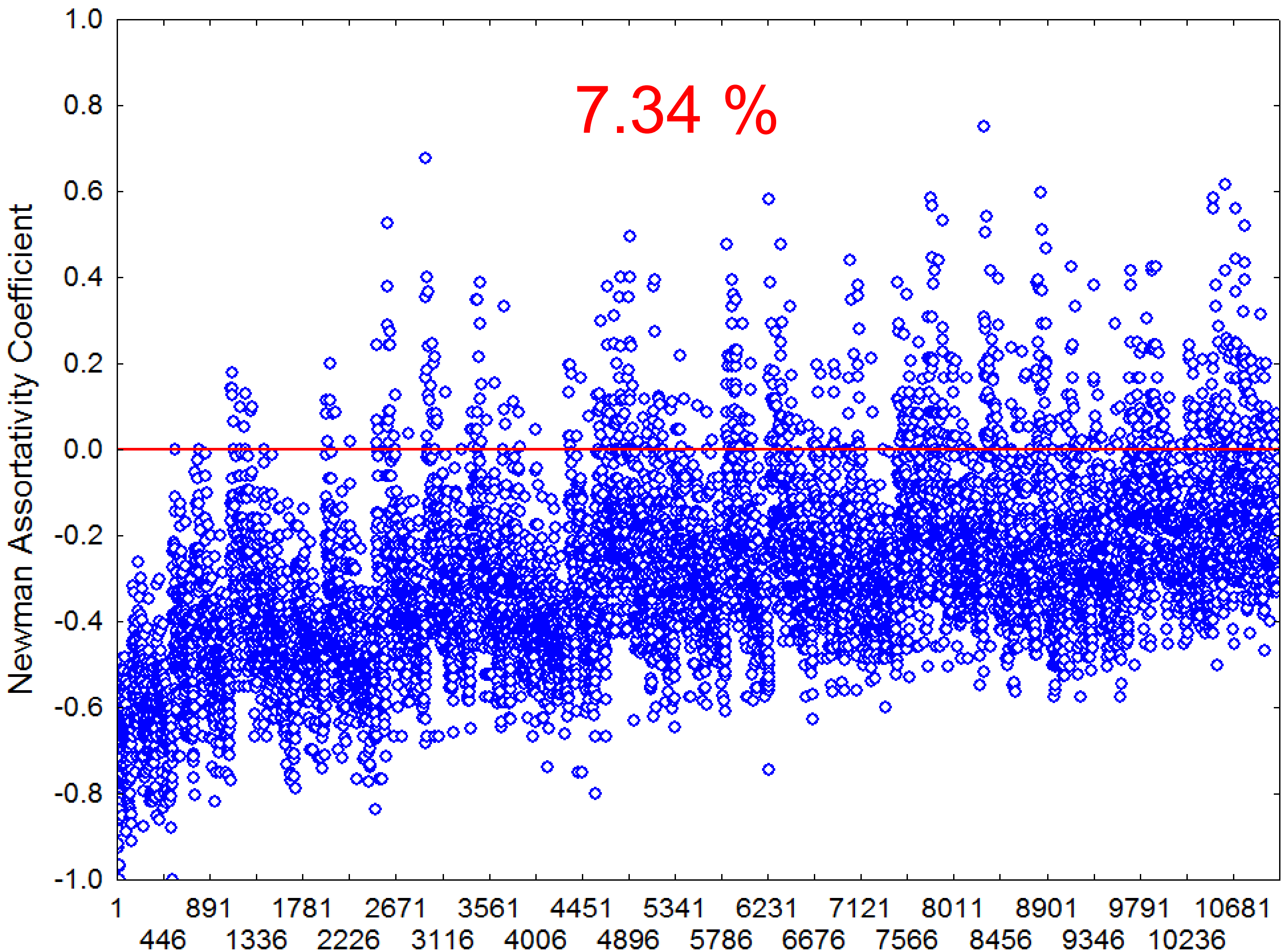
$r > 0$ **Assortative**

$r < 0$ **Disassortative**



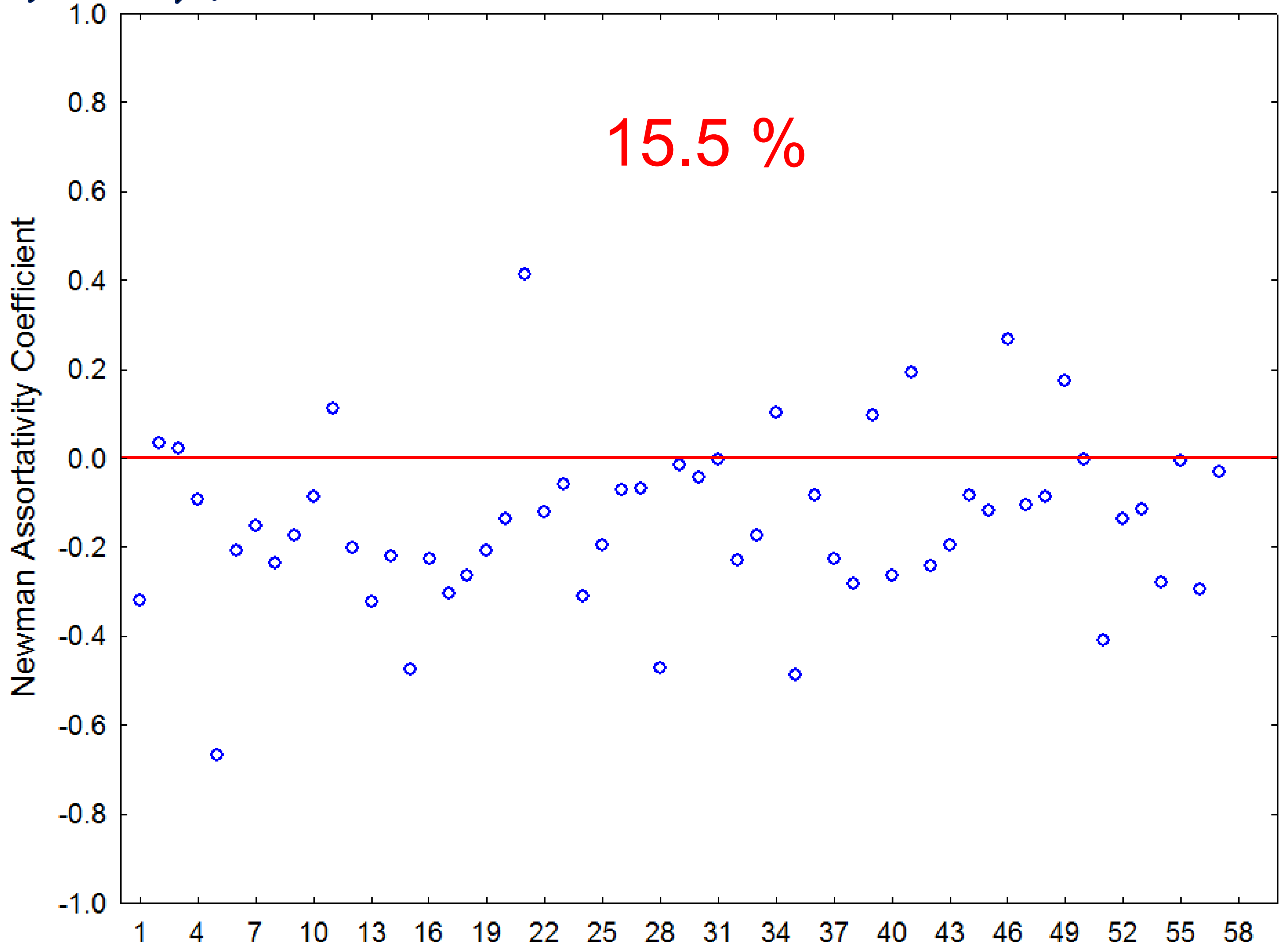
Degree-degree Correlation

Small graphs



Degree-degree Correlation

Real-world networks



Degree-degree Correlation

Statistical approach

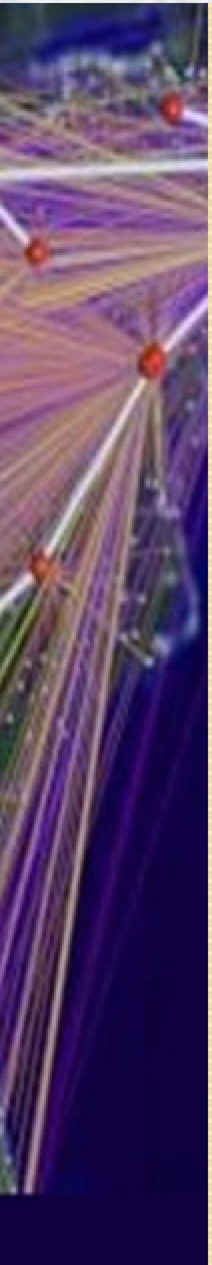
$$r=\frac{\frac{1}{m}\sum_{(i,j)\in E}k_ik_j-\left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i+k_j\right)\right)^2}{\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i^2+k_j^2\right)-\left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i+k_j\right)\right)^2}$$

$$\sum_{(i,j)\in E}\left(k_i^2+k_j^2\right)=\sum_t k_t^3$$

$$\left(\sum_{(i,j)\in E}\left(k_i+k_j\right)\right)^2=\left(\sum_t k_t^2\right)^2$$

$$\frac{1}{2m}\left[\sum_t k_t^3-\frac{1}{2m}\left(\sum_t k_t^2\right)^2\right]$$

$$\frac{1}{4m^2}\left[\left(\sum_t k_t\right)\left(\sum_t k_t^3\right)-\left(\sum_t k_t^2\right)^2\right]$$



Degree-degree Correlation

Statistical approach

$$\frac{1}{4m^2}\left[\left(\sum_t k_t\right)\left(\sum_t k_t^3\right)-\left(\sum_t k_t^2\right)^2\right]$$

$$\left(\sum_i k_i\right)\left(\sum_i k_i^3\right)-\left(\sum_i k_i^2\right)^2 \overset{?}{\geq} 0$$

$$\sum_{i,j} k_i k_j \left(k_i^2 + k_j^2\right) - 2\sum_{i,j}\left(k_i k_j\right)^2 \overset{?}{\geq} 0$$

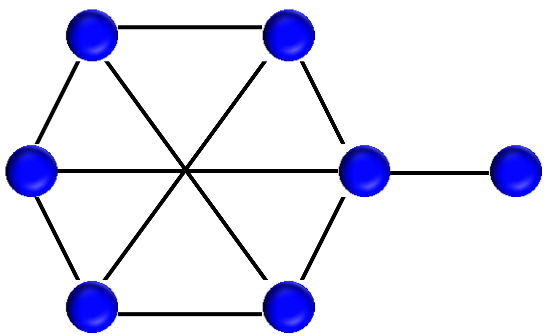
$$\sum_{i,j} k_i k_j \left(k_i^2 + k_j^2 - 2k_i k_j\right) \overset{?}{\geq} 0$$

$$\sum_{i,j} k_i k_j \left(k_i - k_j\right)^2 \geq 0.$$

Degree-degree Correlation

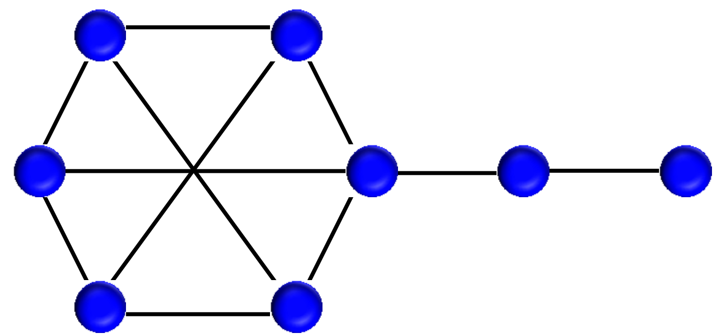
What is it?

Disassortative



r = -0.538

Assortative



r = 0.200

Degree-degree Correlation

Combinatorial approach

$$r=\frac{\frac{1}{m}\sum_{(i,j)\in E}k_ik_j-\left(\frac{1}{2m}\boxed{\sum_{(i,j)\in E}\left(k_i+k_j\right)}\right)^2}{\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i^2+k_j^2\right)-\left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i+k_j\right)\right)^2}$$

$$\sum_t k_t^2=\sum_t\left(k_t^2-k_t+k_t\right)$$

$$=\sum_t\left(k_t\left(k_t-1\right)\right)+\sum_t k_t$$

$$=2\left|P_2\right|+2m;$$

⇒

$$\sum_{(i,j)\in E}\left(k_i+k_j\right)=\sum_t k_t^2$$

$$=2\left|P_2\right|+2m$$

$$\left(\sum_{(i,j)\in E}\left(k_i+k_j\right)\right)^2=\left(2\left|P_2\right|+2m\right)^2$$

Degree-degree Correlation

Combinatorial approach

$$r=\frac{\frac{1}{m}\sum_{(i,j)\in E}k_ik_j-\left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i+k_j\right)\right)^2}{\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i^2+k_j^2\right)-\left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i+k_j\right)\right)^2}$$

$$\left|P_3\right|=\sum_{(i,j)\in E}\left(k_i-1\right)\left(k_j-1\right)-3\left|C_3\right|$$

$$=\sum_{(i,j)\in E}\left(k_ik_j\right)-\sum_{(i,j)\in E}\left(k_i+k_j\right)+m-3\left|C_3\right|$$

$$\sum_{(i,j)\in E}\left(k_ik_j\right)=\sum_{(i,j)\in E}\left(k_i+k_j\right)-m+3\left|C_3\right|+\left|P_3\right|$$

$$=m+2\left|P_2\right|+\left|P_3\right|+3\left|C_3\right|$$

Degree-degree Correlation

Combinatorial approach

$$r = \frac{\frac{1}{m}\sum_{(i,j)\in E} k_i k_j - \left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i + k_j\right)\right)^2}{\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i^2 + k_j^2\right) - \left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i + k_j\right)\right)^2}$$

$$\left|S_{1,3}\right| = \sum_t \binom{k_t}{3} = \frac{1}{6}\sum_i k_t\left(k_t - 1\right)\left(k_t - 2\right)$$

$$= \frac{1}{6}\sum_t k_t^3 - \frac{1}{2}\sum_t k_t^2 + \frac{1}{3}\sum_t k_t$$

$$\sum_t k_t^3 = 6\left|S_{1,3}\right| + 3\sum_t k_t^2 - 2\sum_t k_t$$

$$= 6\left|S_{1,3}\right| + 6\left|P_2\right| + 2m$$

$$\sum_{(i,j)\in E}\left(k_i^2 + k_j^2\right) = 6\left|S_{1,3}\right| + 6\left|P_2\right| + 2m$$

Degree-degree Correlation

Combinatorial approach

$$r = \frac{\frac{1}{m}\left(m + 2|P_2| + |P_3| + 3|C_3|\right) - \frac{1}{4m^2}\left(2|P_2| + 2m\right)^2}{\frac{1}{2m}\left(6|S_{1,3}| + 6|P_2| + 2m\right) - \frac{1}{4m^2}\left(2|P_2| + 2m\right)^2}$$

$$r = \frac{|P_3| + 3|C_3| - \frac{|P_2|^2}{m}}{3|S_{1,3}| + |P_2| - \frac{|P_2|^2}{m}}$$

$$|P_{r/s}| = |P_r| / |P_s|$$

$$r = \frac{|P_2|\left(|P_{3/2}| + \frac{3|C_3|}{|P_2|} - |P_{2/1}|\right)}{3|S_{1,3}| + |P_2|\left(1 - |P_{2/1}|\right)}$$

$$r = \frac{|P_2|\left(|P_{3/2}| + C - |P_{2/1}|\right)}{3|S_{1,3}| + |P_2|\left(1 - |P_{2/1}|\right)}$$

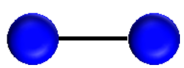
Degree-degree Correlation

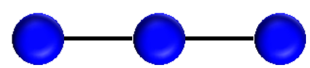
Combinatorial approach

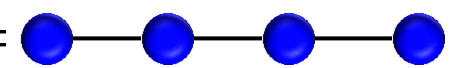
$$r = \frac{|P_2|\left(|P_{3/2}| + C - |P_{2/1}|\right)}{3|S_{1,3}| + |P_2|\left(1 - |P_{2/1}|\right)}$$

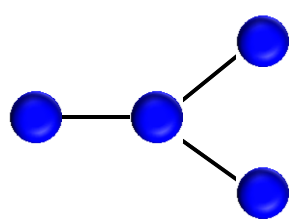
$$|P_{r/s}| = |P_r| / |P_s|$$

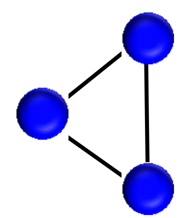
$$C = 3|C_3| / |P_2|$$

$P_1 =$ 

$P_2 =$ 

$P_3 =$ 

$S_{1,3} =$ 

$C_3 =$ 

Estrada, E., Combinatorial study of degree assortativity in networks. *Physical Review E* 84 **2011**, 047101

Degree-degree Correlation

Combinatorial approach

Remark: Because we have already proved that the denominator of the Pearson correlation coefficient is always nonnegative, the assortativity/disassortativity of a network is determined by the sign of the numerator.

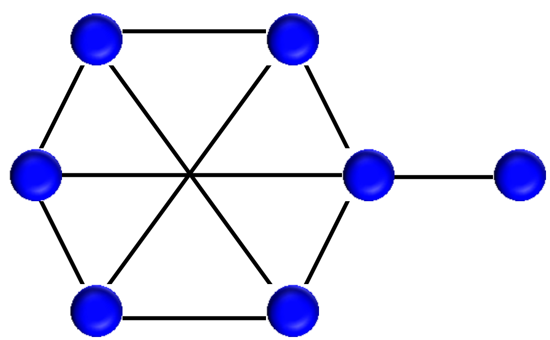
Assortative	***Disassortative***
$r > 0$	$r < 0$
$\lvert P_{2/1} \rvert < \lvert P_{3/2} \rvert + C$	$\lvert P_{2/1} \rvert > \lvert P_{3/2} \rvert + C$

Degree-degree Correlation

Combinatorial approach

DISASSORTATIVE

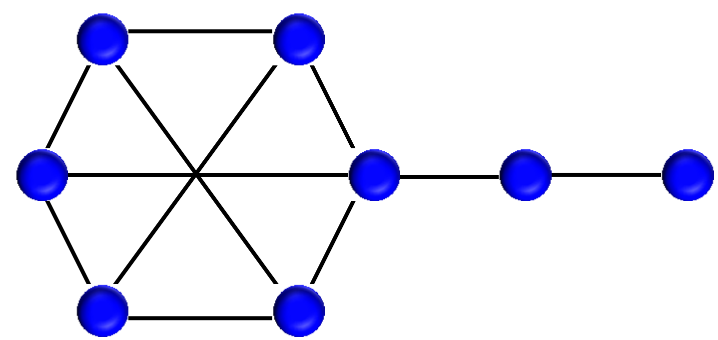
$$|P_{2/1}| > |P_{3/2}| \left(C = 0\right)$$



r = -0.538

ASSORTATIVE

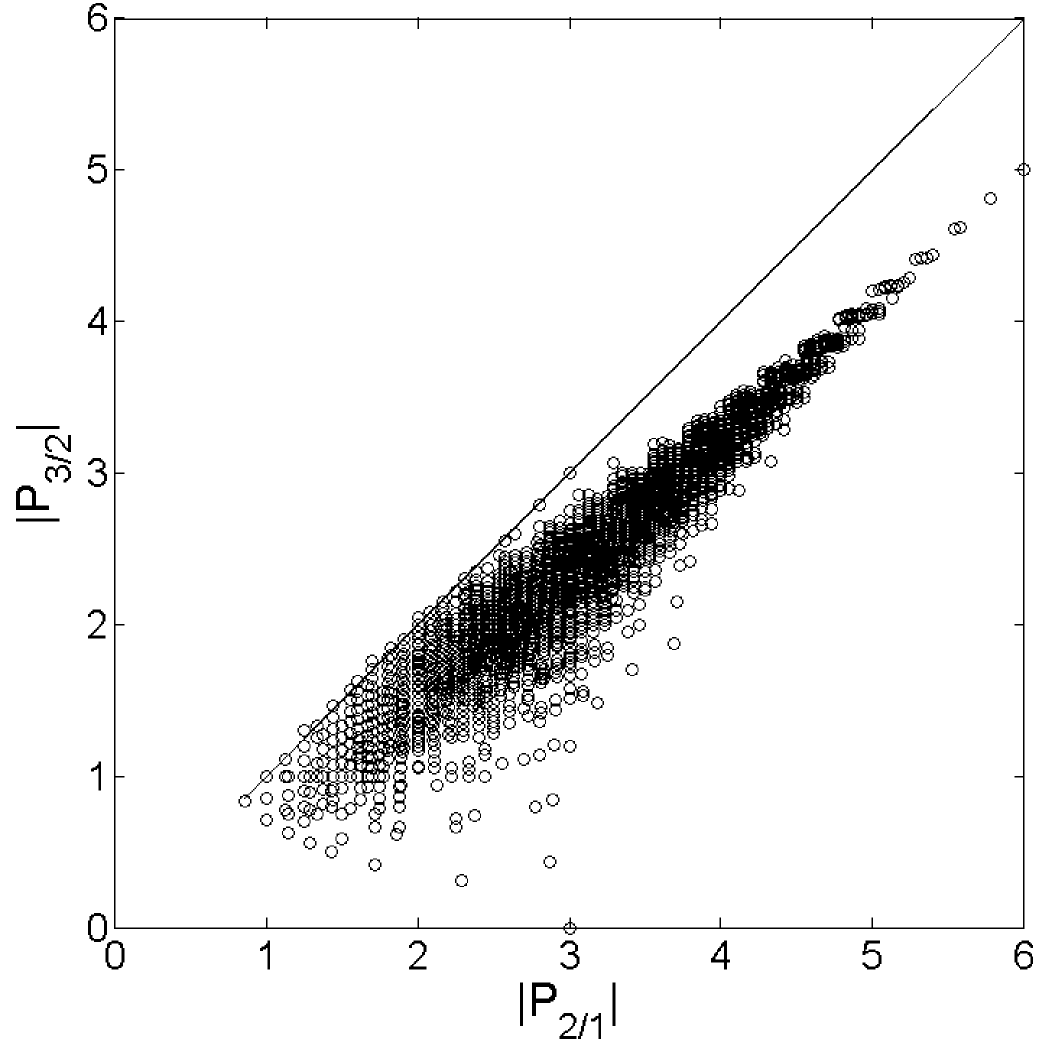
$$|P_{2/1}| < |P_{3/2}| \left(C = 0\right)$$

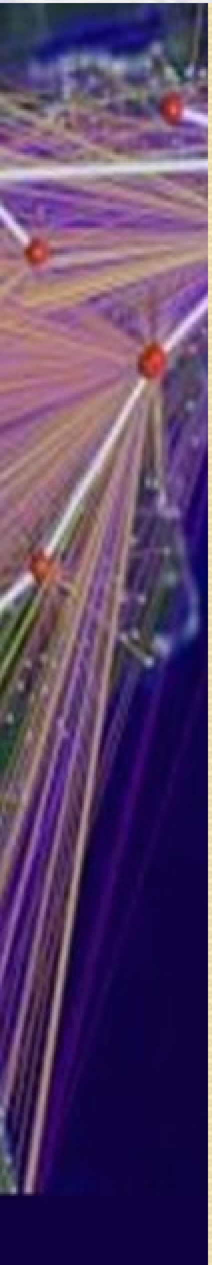


r = 0.200

Degree-degree Correlation

Combinatorial approach





Degree-degree Correlation

$$r = \frac{|P_2|\left(|P_{3/2}| + C - |P_{2/1}|\right)}{3|S_{1,3}| + |P_2|\left(1 - |P_{2/1}|\right)}$$

TABLE I. Relative branching ($|P_{2/1}|$), transitivity (C), intermodular connectivity ($|P_{3/2}|$), and assortativity coefficient for real-world networks.

Network	$\vert P_{2/1}\vert$		$\vert P_{3/2}\vert$	C	r
Prison	4.25	<	(4.09 +	0.288)	
Protein residue	4.41	<	(4.45 +	0.417)	
St. Marks	10.54	<	(10.46 +	0.291)	
Geom	17.42	<	(22.09 +	0.224)	
Corporate	19.42	<	(20.60 +	0.498)	
Roget	9.55	<	(10.08 +	0.134)	
Jazz	127.30	<	(144.84 +	0.771)	
Zachary	6.77	>	(4.49 +	0.256)	
Drugs	14.58	>	(12.84 +	0.368)	
Transcription	12.51	>	(3.01 +	0.016)	
Bridge Brook	22.42	>	(17.31 +	0.191)	
USAir97	43.36	>	(36.97 +	0.396)	
Internet	91.00	>	(11.53 +	0.015)	