Roman V. Belavkin   Panos Pardalos

Faculty of Science and Technology
Middlesex University, London NW4 4BT, UK

March 6, 2018
Summer School on Operational Research and Applications
Higher School of Economics, Nizhny Novgorod, Russia

The power-law degree sequence

Power-law and maximum entropy

Power-law exponent (inverse temperature)

Free energy and phase transition

Graphs

Definition (Graph)

A pair $G = (V, E)$, where $V = \{v_1, \dots, v_N\}$ is a set of *vertexes* (or *nodes*), and $E = \{e_1, \dots, e_M\} \subseteq V \times V$ is a set of *edges* (or *links*, *arrows*).

Example

Social contacts : vertexes represent humans or anumals, edges represent social contacrts (e.g. Zachary karate club, $N = 34$, Zebra network, $N = 27$, Madrid train bombing network, $N = 64$).

Graphs

Definition (Graph)

A pair $G = (V, E)$, where $V = \{v_1, \dots, v_N\}$ is a set of *vertexes* (or *nodes*), and $E = \{e_1, \dots, e_M\} \subseteq V \times V$ is a set of *edges* (or *links*, *arrows*).

Example

Social contacts : vertexes represent humans or anumals, edges represent social contacrts (e.g. Zachary karate club, $N = 34$, Zebra network, $N = 27$, Madrid train bombing network, $N = 64$).

Co-authorship : vertexes represent authors, endges represend coauthorship (e.g. arXiv astro-ph, $N = 19K$, DBLP, $N = 317K$).

Graphs

Definition (Graph)

A pair $G = (V, E)$, where $V = \{v_1, \dots, v_N\}$ is a set of *vertexes* (or *nodes*), and $E = \{e_1, \dots, e_M\} \subseteq V \times V$ is a set of *edges* (or *links*, *arrows*).

Example

Social contacts : vertexes represent humans or anumals, edges represent social contacrts (e.g. Zachary karate club, $N = 34$, Zebra network, $N = 27$, Madrid train bombing network, $N = 64$).

Co-authorship : vertexes represent authors, endges represend coauthorship (e.g. arXiv astro-ph, $N = 19K$, DBLP, $N = 317K$).

Infrastructure : vertexes represent physical locations, edges represent links (e.g. Euroroad, $N = 1K$, US power grid, $N = 5K$).

Graphs

Definition (Graph)

A pair $G = (V, E)$, where $V = \{v_1, \dots, v_N\}$ is a set of *vertexes* (or *nodes*), and $E = \{e_1, \dots, e_M\} \subseteq V \times V$ is a set of *edges* (or *links*, *arrows*).

Example

Social contacts : vertexes represent humans or anumals, edges represent social contacrts (e.g. Zachary karate club, $N = 34$, Zebra network, $N = 27$, Madrid train bombing network, $N = 64$).

Co-authorship : vertexes represent authors, endges represend coauthorship (e.g. arXiv astro-ph, $N = 19K$, DBLP, $N = 317K$).

Infrastructure : vertexes represent physical locations, edges represent links (e.g. Euroroad, $N = 1K$, US power grid, $N = 5K$).

Social networks : edges represent friendships (e.g. subsets of Facebook, $N = 63K$, YouTube, $N = 3M$, LiveJournal, $N = 5M$).

Some questions about graphs and networks

- Is the graph fully connected?

Some questions about graphs and networks

- Is the graph fully connected?
- If not, is there a *giant component*?

Some questions about graphs and networks

- Is the graph fully connected?
- If not, is there a *giant component*?
- What is the average degree (number of links of a node)?

Some questions about graphs and networks

- Is the graph fully connected?
- If not, is there a *giant component*?
- What is the average degree (number of links of a node)?
- What is the average distance between nodes (shortest path length)?

Some questions about graphs and networks

- Is the graph fully connected?
- If not, is there a *giant component*?
- What is the average degree (number of links of a node)?
- What is the average distance between nodes (shortest path length)?
-

Degree Sequence

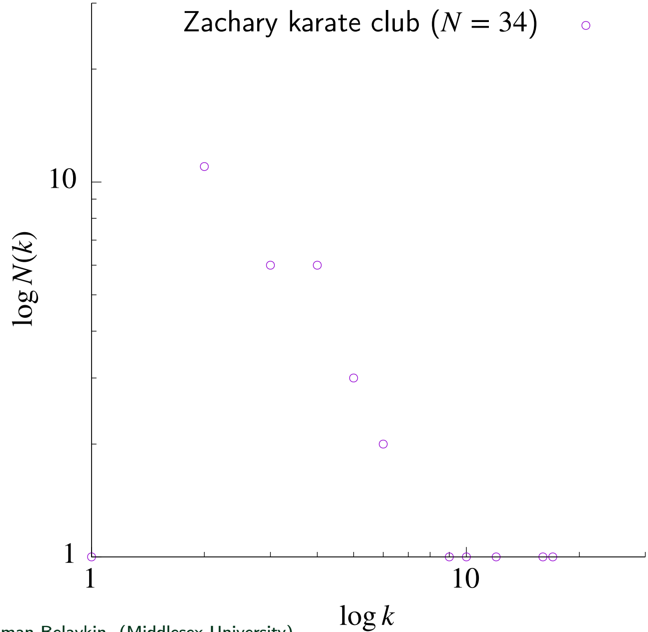
Definition (Degree sequence)

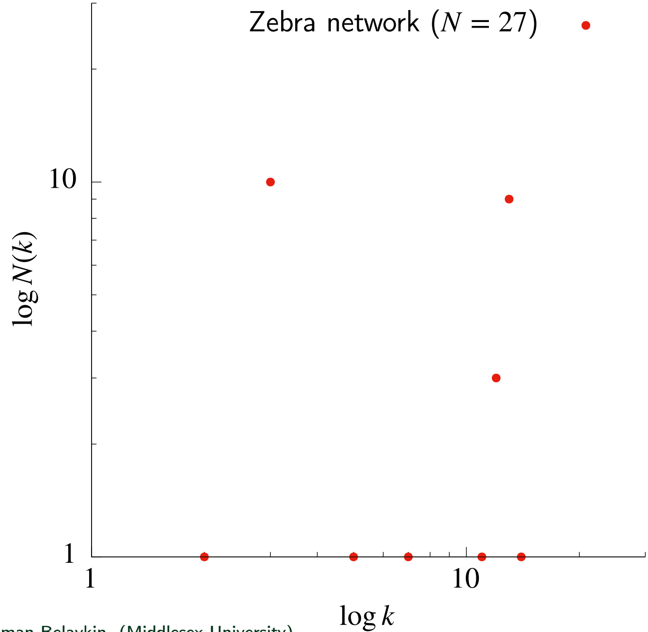
Function $N : \mathbb{N} \to \mathbb{N} \cup \{0\}$ representing the nuber $N(k)$ of vertexes $v \in V$ with degree k (number of edges $(v, \cdot)$ or $(\cdot, v) \in E \subseteq V \times V$):

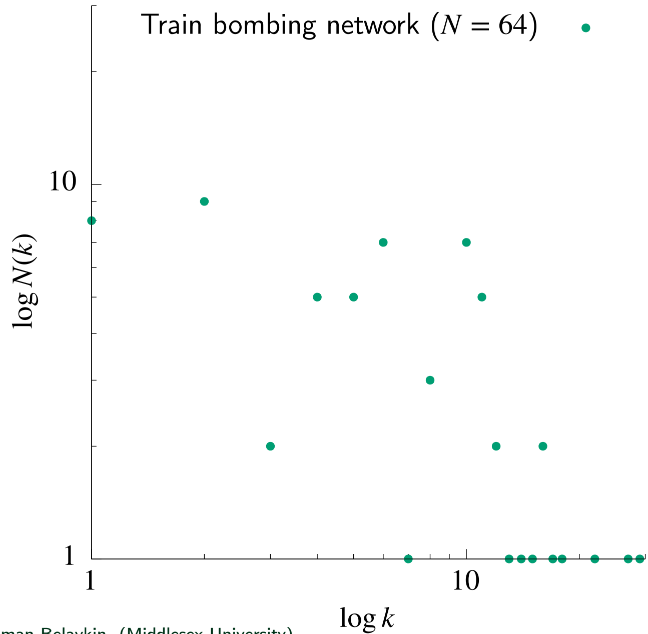
$$N(k) := |\{v \in V : \deg(v) = k\}|$$

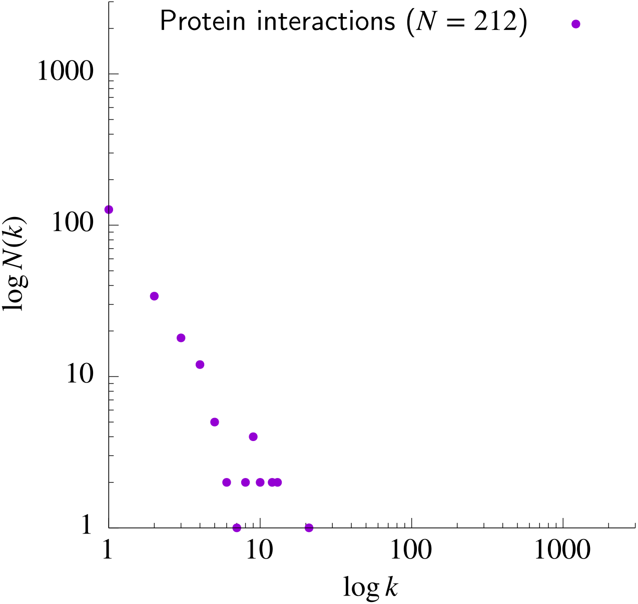
Normalized $N(k)$ is the degree *distribution*

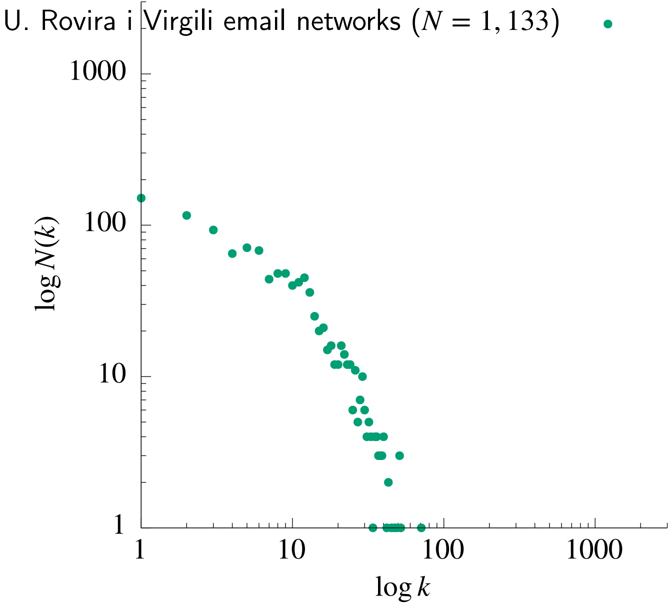
$$P(k) = \frac{N(k)}{N}$$

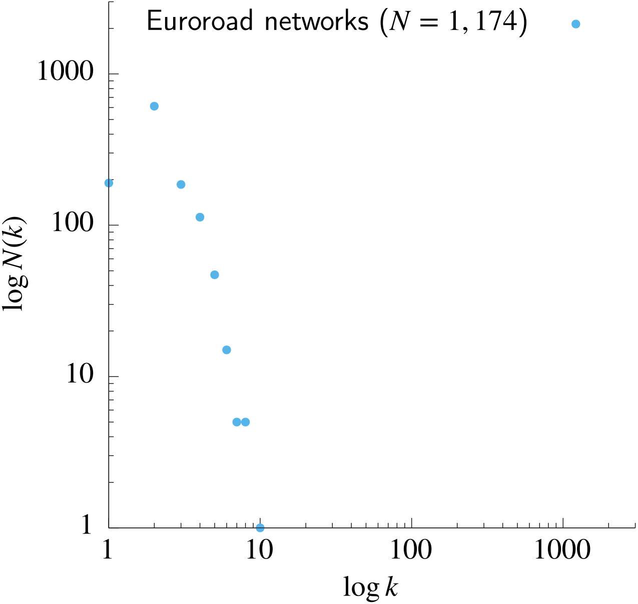

Zachary karate club ($N = 34$)
10
1
$\log N(k)$
1
10
$\log k$

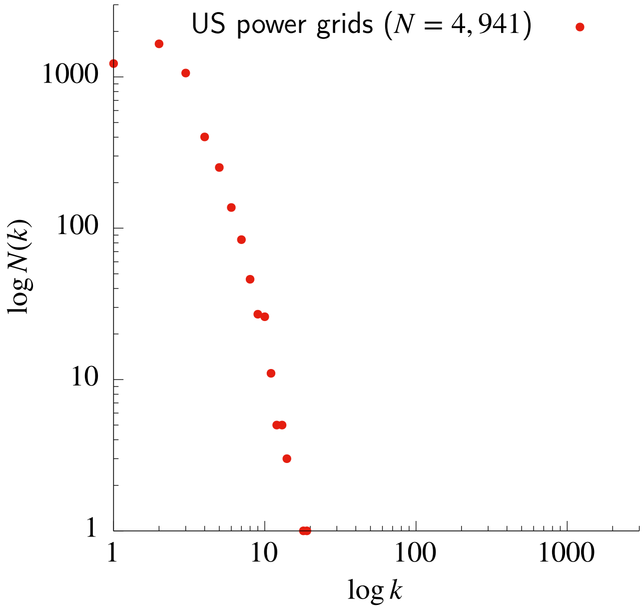

Zebra network ($N = 27$)
10
1
$\log N(k)$
1
10
$\log k$

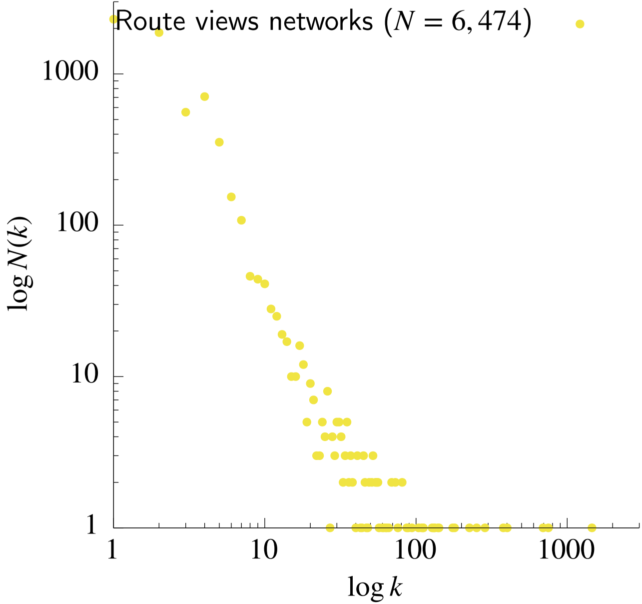

Train bombing network ($N = 64$)
10
1
$\log N(k)$
1
10
$\log k$

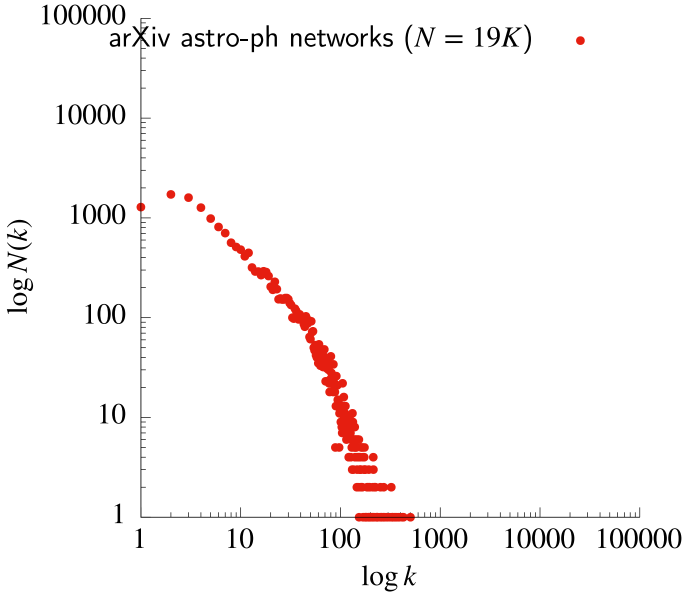

Protein interactions ($N = 212$)
1000
100
10
1
1
10
100
1000
$\log N(k)$
$\log k$

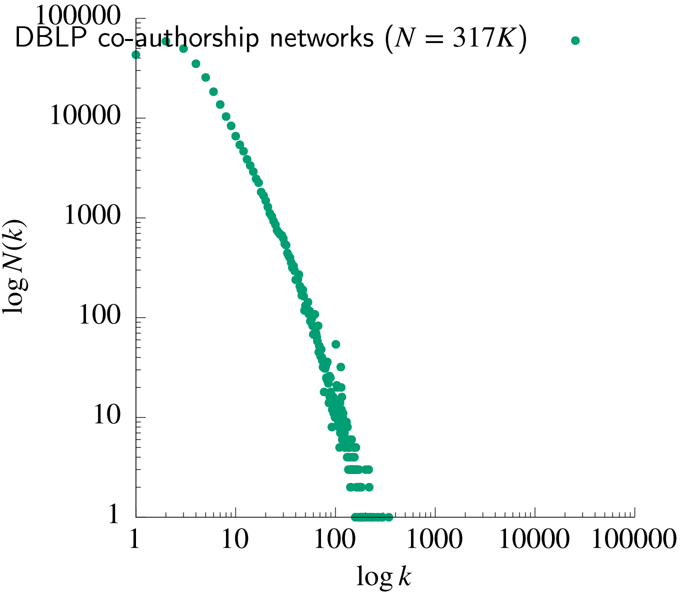

U. Rovira i Virgili email networks ($N = 1,133$)
1000
100
10
1
$\log N(k)$
1
10
100
1000
$\log k$

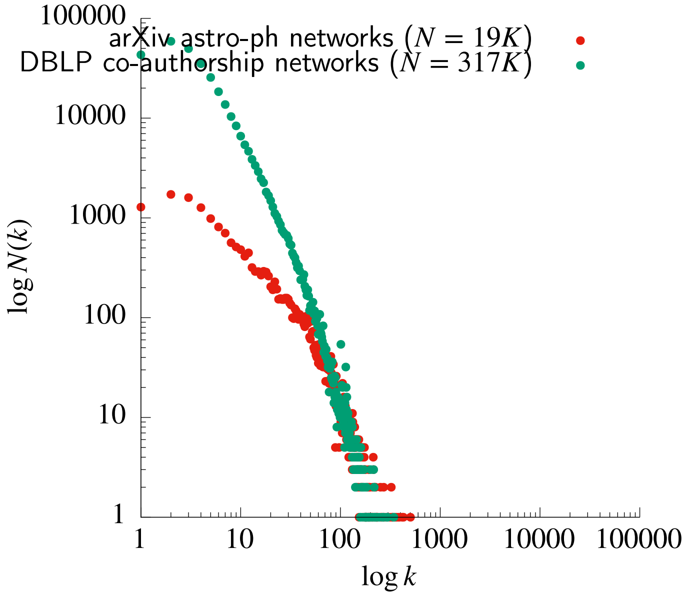

Euroroad networks $(N = 1,174)$
$\log N(k)$
1000
100
10
1
1
10
100
1000
$\log k$

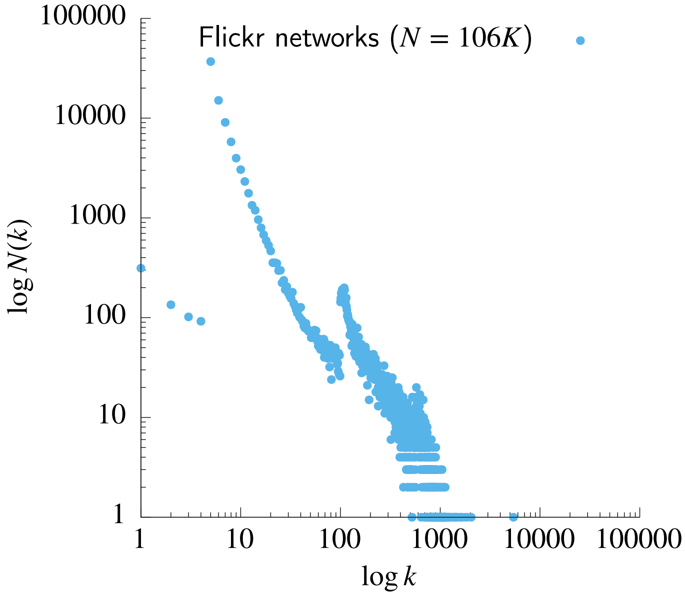

US power grids ($N = 4,941$)
1000
100
10
1
$\log N(k)$
1
10
100
1000
$\log k$

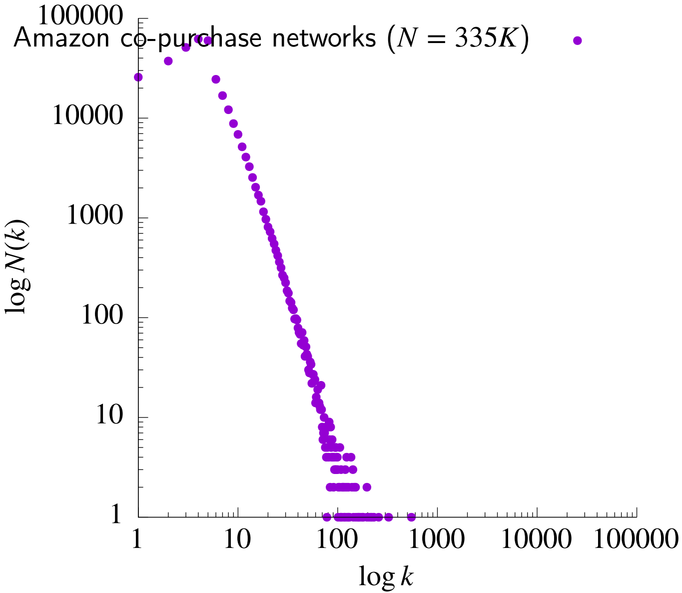

Route views networks ($N = 6,474$)
1000
100
10
1
$\log N(k)$
1
10
100
1000
$\log k$

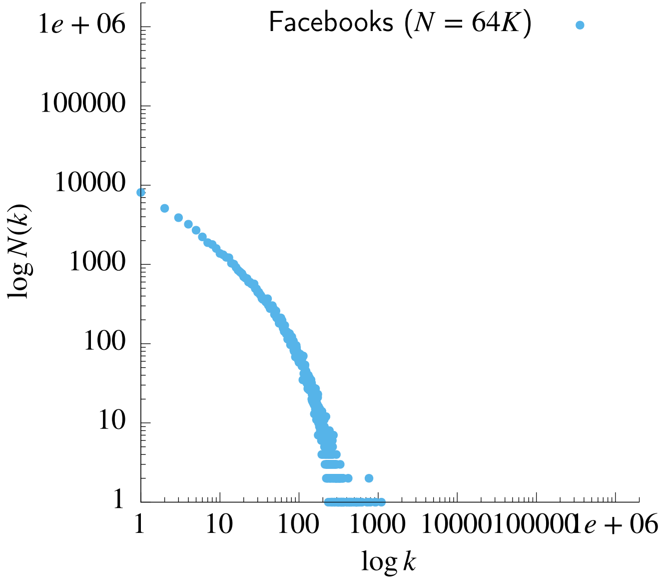

arXiv astro-ph networks ($N = 19K$)
100000
10000
1000
100
10
1
$\log N(k)$
1
10
100
1000
10000
100000
$\log k$

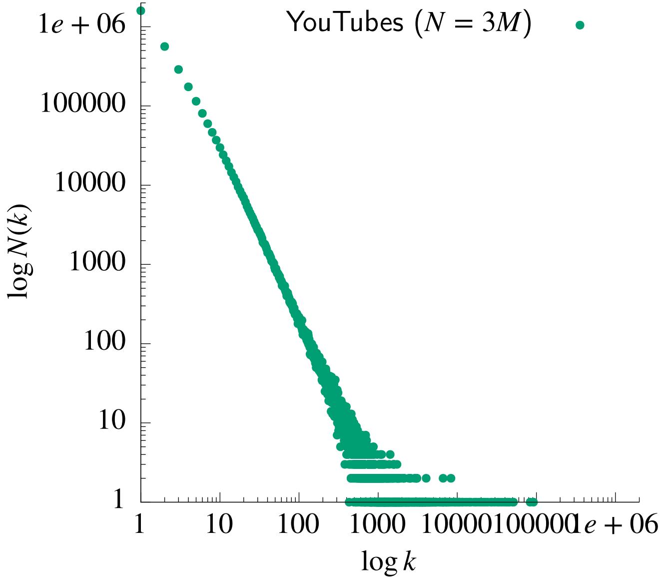

DBLP co-authorship networks ($N = 317K$)
100000
10000
1000
100
10
1
$\log N(k)$
1
10
100
1000
10000
100000
$\log k$

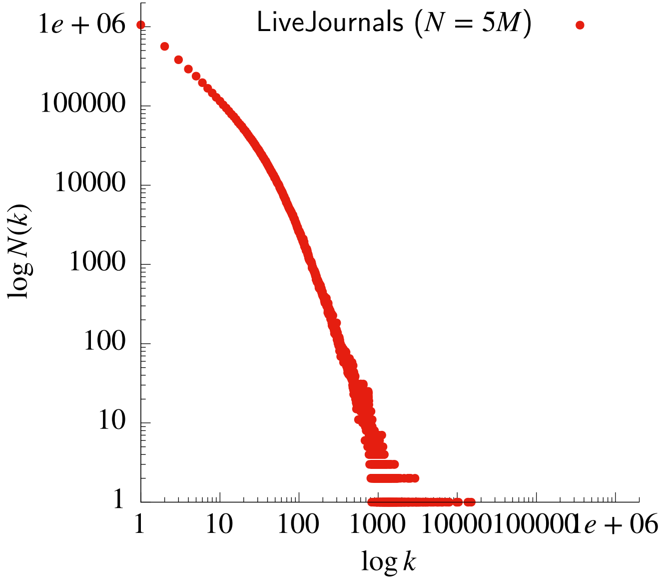

arXiv astro-ph networks ($N = 19K$)
DBLP co-authorship networks ($N = 317K$)
100000
10000
1000
100
10
1
$\log N(k)$
1
10
100
1000
10000
100000
$\log k$

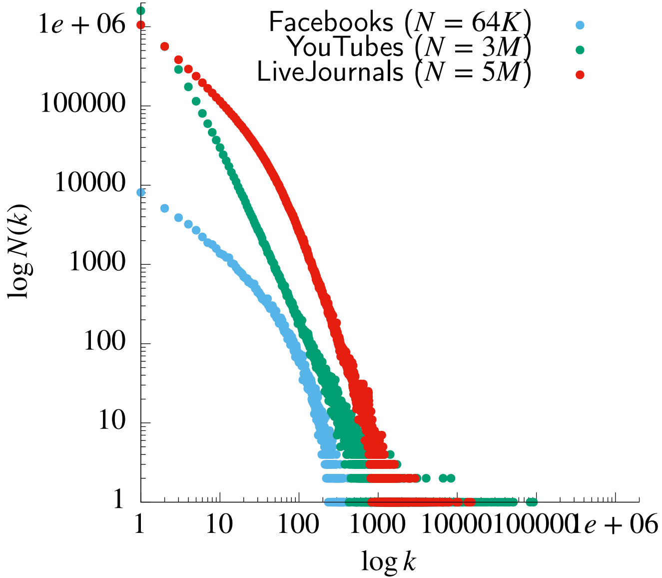

Flickr networks ($N = 106K$)
100000
10000
1000
100
10
1
$\log N(k)$
1
10
100
1000
10000
100000
$\log k$


Amazon co-purchase networks ($N = 335K$)
100000
10000
1000
100
10
1
$\log N(k)$
1
10
100
1000
10000
100000
$\log k$


Facebooks ($N = 64K$)
1e+06
100000
10000
1000
100
10
1
log N(k)
1
10
100
1000
10000
100000
1e+06
log k


YouTubes ($N = 3M$)
1e + 06
100000
10000
1000
100
10
1
log N(k)
1
10
100
1000
10000
100000
1e + 06
log k


LiveJournals ($N = 5M$)
$\log N(k)$
1e + 06
100000
10000
1000
100
10
1
$\log k$
1
10
100
1000
10000
100000
1e + 06


Facebooks ($N = 64K$)
YouTubes ($N = 3M$)
LiveJournals ($N = 5M$)
$\log N(k)$
$\log k$
1e+06
100000
10000
1000
100
10
1
1
10
100
1000
10000
100000
1e+06

Power-Law

Power-Law

$$\ln N(k) = \alpha - \beta \ln k$$

α — intercept

β — slope or *exponent*

$$P(k) = \frac{k^{-\beta}}{\sum_{k=1}^{N} k^{-\beta}}$$

- Almost surely connected for $\beta < 1$, and a.s. disconnected for $\beta > 1$ (Aiello, Chung, & Lu, 2000, 2001).

Power-Law

Power-Law

$$\ln N(k) = \alpha - \beta \ln k$$

α — intercept

β — slope or *exponent*

$$P(k) = \frac{k^{-\beta}}{\sum_{k=1}^{N} k^{-\beta}}$$

- Almost surely connected for $\beta < 1$, and a.s. disconnected for $\beta > 1$ (Aiello et al., 2000, 2001).
- A giant componnent almost surely exists if $Q := \mathbb{E}_P\{k^2\} - 2\mathbb{E}_P\{k\} \geq 0$ (Molloy & Reed, 1995).

Power-Law

Power-Law

$$\ln N(k) = \alpha - \beta \ln k$$

α — intercept

β — slope or *exponent*

$$P(k) = \frac{k^{-\beta}}{\sum_{k=1}^{N} k^{-\beta}}$$

- Almost surely connected for $\beta < 1$, and a.s. disconnected for $\beta > 1$ (Aiello et al., 2000, 2001).
- A giant componnent almost surely exists if $Q := \mathbb{E}_P\{k^2\} - 2\mathbb{E}_P\{k\} \geq 0$ (Molloy & Reed, 1995).
- For power-law with $N \to \infty$ ($\sum_{k=1}^{N} k^{-\beta} \to \zeta(\beta)$) this becomes $\zeta(\beta - 2) - 2\zeta(\beta - 1) = 0$, which guves the value $\beta_0 \approx 3.47875$ (Aiello et al., 2000, 2001).

Power-Law

Power-Law

$$\ln N(k) = \alpha - \beta \ln k$$

α — intercept
β — slope or *exponent*

$$P(k) = \frac{k^{-\beta}}{\sum_{k=1}^{N} k^{-\beta}}$$

- Almost surely connected for $\beta < 1$, and a.s. disconnected for $\beta > 1$ (Aiello et al., 2000, 2001).
- A giant componnent almost surely exists if $Q := \mathbb{E}_P\{k^2\} - 2\mathbb{E}_P\{k\} \geq 0$ (Molloy & Reed, 1995).
- For power-law with $N \to \infty$ ($\sum_{k=1}^{N} k^{-\beta} \to \zeta(\beta)$) this becomes $\zeta(\beta-2) - 2\zeta(\beta-1) = 0$, which guves the value $\beta_0 \approx 3.47875$ (Aiello et al., 2000, 2001).
- Power-law graphs can be generated by preferential attachment:

$$P[(i,j) \in E \mid k_i] = \frac{k_i^{\gamma}}{\sum_{k=1}^{k_{\max}} k_i^{\gamma}}$$

Power-Law as exponential family

- Changing $k^{-\beta} = e^{-\beta \ln k}$ let us write

$$P(k) = e^{-\beta \ln k - \Gamma(\beta)}$$

Power-Law as exponential family

- Changing $k^{-\beta} = e^{-\beta \ln k}$ let us write

$$P(k) = e^{-\beta \ln k - \Gamma(\beta)}$$

- $\Gamma(\beta) = \ln Z(\beta)$ is the *cumulant generating function*:

$$\Gamma' = m_1 = -\mathbb{E}_P\{\ln k\}$$
$$\Gamma'' = m_2 - m_1^2 = \sigma^2(\ln k)$$

where $m_n = \frac{Z^{(n)}}{Z}$ are nth moments.

Variational problems with entropy

- $P(k) = \exp\{-\beta \ln k - \Gamma(\beta)\}$ is a solution to:

$$\text{maximize} \quad H(P) \quad \text{subject to} \quad \mathbb{E}_P\{\ln k\} \leq \upsilon$$

Variational problems with entropy

- $P(k) = \exp\{-\beta \ln k - \Gamma(\beta)\}$ is a solution to:

$$\text{maximize} \quad H(P) \quad \text{subject to} \quad \mathbb{E}_P\{\ln k\} \leq \upsilon$$

- where H is the entropy of $P(k)$:

$$H(P) = -\sum_{k=1}^{N} [\ln P(k)]\, P(k)$$

Variational problems with entropy

- $P(k) = \exp\{-\beta \, \ln k - \Gamma(\beta)\}$ is a solution to:

$$\text{maximize} \quad H(P) \quad \text{subject to} \quad \mathbb{E}_P\{\ln k\} \leq \upsilon$$

- where H is the entropy of $P(k)$:

$$H(P) = -\sum_{k=1}^{N} [\ln P(k)] \, P(k)$$

- Equivalent problem:

$$\text{minimize} \quad \mathbb{E}_P\{\ln k\} \quad \text{subject to} \quad H(P) \geq \ln N - \lambda \geq 0$$

Variational problems with entropy

- $P(k) = \exp\{-\beta \ln k - \Gamma(\beta)\}$ is a solution to:

$$\text{maximize} \quad H(P) \quad \text{subject to} \quad \mathbb{E}_P\{\ln k\} \leq \upsilon$$

- where H is the entropy of $P(k)$:

$$H(P) = -\sum_{k=1}^{N} [\ln P(k)]\, P(k)$$

- Equivalent problem:

$$\text{minimize} \quad \mathbb{E}_P\{\ln k\} \quad \text{subject to} \quad H(P) \geq \ln N - \lambda \geq 0$$

- $\ln k$ plays the role of a cost to be minimized.

Solution using Lagrange multipliers

- Lagrange function

$$K(P,\beta,\gamma) = -\sum_{k=1}^{N}[\ln P(k)]\,P(k) + \beta\left[\upsilon - \sum_{k=1}^{N}(\ln k)\,P(k)\right] + \gamma\left[1 - \sum_{k=1}^{N}P(k)\right]$$

Solution using Lagrange multipliers

- Lagrange function

$$K(P,\beta,\gamma) = -\sum_{k=1}^{N} [\ln P(k)]\, P(k) + \beta \left[\upsilon - \sum_{k=1}^{N} (\ln k)\, P(k) \right] + \gamma \left[1 - \sum_{k=1}^{N} P(k) \right]$$

- Necessary and sufficient conditions:

$$\frac{\partial}{\partial P} K(P,\beta,\gamma) = -\ln P(k) - 1 - \beta \ln k + \gamma = 0 \quad \Rightarrow \quad P(k) = e^{-\beta \ln k - \Gamma(\beta)}$$

Solution using Lagrange multipliers

- Lagrange function

$$K(P, \beta, \gamma) = -\sum_{k=1}^{N} [\ln P(k)]\, P(k) + \beta \left[\upsilon - \sum_{k=1}^{N} (\ln k)\, P(k) \right] + \gamma \left[1 - \sum_{k=1}^{N} P(k) \right]$$

- Necessary and sufficient conditions:

$$\frac{\partial}{\partial P} K(P, \beta, \gamma) = -\ln P(k) - 1 - \beta \ln k + \gamma = 0 \quad \Rightarrow \quad P(k) = e^{-\beta \ln k - \Gamma(\beta)}$$

$$\frac{\partial}{\partial \beta} K(P, \beta, \gamma) = \upsilon - \sum_{k=1}^{N} (\ln k)\, P(k) = 0 \quad \Rightarrow \quad \sum_{k=1}^{N} (\ln k)\, P(k) = \upsilon$$

Solution using Lagrange multipliers

- Lagrange function

$$K(P,\beta,\gamma) = -\sum_{k=1}^{N}[\ln P(k)]\,P(k) + \beta\left[\upsilon - \sum_{k=1}^{N}(\ln k)\,P(k)\right] + \gamma\left[1 - \sum_{k=1}^{N}P(k)\right]$$

- Necessary and sufficient conditions:

$$\frac{\partial}{\partial P}K(P,\beta,\gamma) = -\ln P(k) - 1 - \beta\ln k + \gamma = 0 \quad \Rightarrow \quad P(k) = e^{-\beta\,\ln k - \Gamma(\beta)}$$

$$\frac{\partial}{\partial \beta}K(P,\beta,\gamma) = \upsilon - \sum_{k=1}^{N}(\ln k)\,P(k) = 0 \quad \Rightarrow \quad \sum_{k=1}^{N}(\ln k)\,P(k) = \upsilon$$

$$\frac{\partial}{\partial \gamma}K(P,\beta,\gamma) = 1 - \sum_{k=1}^{N}P(k) = 0 \quad \Rightarrow \quad \Gamma(\beta) = \ln\sum_{k=1}^{N}e^{-\beta\ln k}$$

Optimal communication

- Let $c(x_i, y_j)$ be some cost function for $x_i, y_j \in V$:

$$\text{minimize} \quad \mathbb{E}_P\{c(x_i, y_j)\} \quad \text{subject to} \quad I(x_i, y_j) \le \lambda$$

Optimal communication

- Let $c(x_i, y_j)$ be some cost function for $x_i, y_j \in V$:

$$\text{minimize} \quad \mathbb{E}_P\{c(x_i, y_j)\} \quad \text{subject to} \quad I(x_i, y_j) \leq \lambda$$

- where I is Shannon's mutual information:

$$I(x_i, y_j) := \sum_{(x_i, y_j) \in X \times Y} \left[\ln \frac{P(x_i, y_j)}{P(x_i)P(y_j)} \right] P(x_i, y_j)$$

$$= H\{x\} - H\{x \mid y\} = H\{y\} - H\{y \mid\}$$

Optimal communication

- Let $c(x_i, y_j)$ be some cost function for $x_i, y_j \in V$:

$$\text{minimize} \quad \mathbb{E}_P\{c(x_i, y_j)\} \quad \text{subject to} \quad I(x_i, y_j) \leq \lambda$$

- where I is Shannon's mutual information:

$$I(x_i, y_j) := \sum_{(x_i, y_j) \in X \times Y} \left[\ln \frac{P(x_i, y_j)}{P(x_i)P(y_j)} \right] P(x_i, y_j)$$
$$= H\{x\} - H\{x \mid y\} = H\{y\} - H\{y \mid\}$$

- Equivalent problem

$$\text{minimize} \quad I(x_i, y_j) \quad \text{subject to} \quad \mathbb{E}_P\{c(x_i, y_j)\} \leq \upsilon$$

Optimal communication

- Let $c(x_i, y_j)$ be some cost function for $x_i, y_j \in V$:

$$\text{minimize} \quad \mathbb{E}_P\{c(x_i, y_j)\} \quad \text{subject to} \quad I(x_i, y_j) \leq \lambda$$

- where I is Shannon's mutual information:

$$I(x_i, y_j) := \sum_{(x_i, y_j) \in X \times Y} \left[\ln \frac{P(x_i, y_j)}{P(x_i) P(y_j)} \right] P(x_i, y_j)$$

$$= H\{x\} - H\{x \mid y\} = H\{y\} - H\{y \mid \}$$

- Equivalent problem

$$\text{minimize} \quad I(x_i, y_j) \quad \text{subject to} \quad \mathbb{E}_P\{c(x_i, y_j)\} \leq \upsilon$$

- Solution

$$P(x_i, y_i) = e^{-\beta c(x_i, y_j) - \Gamma(\beta, x_i)} P(x_i) P(y_j)$$

Average distance

- Consider $c(x_i, y_j) = d(i,j) - 1$ and denote $\ell := \mathbb{E}_P\{d(i,j)\}$.

Average distance

- Consider $c(x_i, y_j) = d(i,j) - 1$ and denote $\ell := \mathbb{E}_P\{d(i,j)\}$.
- Let z_m denote the average number of neighbours at $d(i,j) = m$:

$$z_0 = 1, \quad z_1 = \mathbb{E}_P\{k\}, \dots, z_{m+1} = \frac{z_2}{z_1} z_m$$

Average distance

- Consider $c(x_i, y_j) = d(i,j) - 1$ and denote $\ell := \mathbb{E}_P\{d(i,j)\}$.
- Let z_m denote the average number of neighbours at $d(i,j) = m$:

$$z_0 = 1, \quad z_1 = \mathbb{E}_P\{k\}, \dots, z_{m+1} = \frac{z_2}{z_1} z_m$$

- For $|V| \gg z_1$ and $z_2 \gg z_1$ one can assume

$$|V| = \sum_{m=0}^{\ell} z_m = 1 + \sum_{m=1}^{\ell} \left[\frac{z_2}{z_1}\right]^{m-1} z_1 \approx 1 + \left[\frac{z_2}{z_1}\right]^{\ell-1} z_1$$

Average distance

- Consider $c(x_i, y_j) = d(i,j) - 1$ and denote $\ell := \mathbb{E}_P\{d(i,j)\}$.
- Let z_m denote the average number of neighbours at $d(i,j) = m$:

$$z_0 = 1, \quad z_1 = \mathbb{E}_P\{k\}, \dots, z_{m+1} = \frac{z_2}{z_1} z_m$$

- For $|V| \gg z_1$ and $z_2 \gg z_1$ one can assume

$$|V| = \sum_{m=0}^{\ell} z_m = 1 + \sum_{m=1}^{\ell} \left[\frac{z_2}{z_1}\right]^{m-1} z_1 \approx 1 + \left[\frac{z_2}{z_1}\right]^{\ell-1} z_1$$

- This leads to

$$\ell = \frac{\ln N - \ln z_1}{\ln(z_2/z_1)} + 1$$

Average distance

- Consider $c(x_i, y_j) = d(i,j) - 1$ and denote $\ell := \mathbb{E}_P\{d(i,j)\}$.
- Let z_m denote the average number of neighbours at $d(i,j) = m$:

$$z_0 = 1, \quad z_1 = \mathbb{E}_P\{k\}, \dots, z_{m+1} = \frac{z_2}{z_1} z_m$$

- For $|V| \gg z_1$ and $z_2 \gg z_1$ one can assume

$$|V| = \sum_{m=0}^{\ell} z_m = 1 + \sum_{m=1}^{\ell} \left[\frac{z_2}{z_1}\right]^{m-1} z_1 \approx 1 + \left[\frac{z_2}{z_1}\right]^{\ell-1} z_1$$

- This leads to

$$\ell = \frac{\ln N - \ln z_1}{\ln(z_2/z_1)} + 1$$

- Relacing $z_1 = \mathbb{E}_P\{k\}$ by the actual degree k_i, we obtain conditional $\ell(k_i) := \mathbb{E}_P\{d(i,j) \mid k_i\}$:

$$\ell(k_i) = \frac{\ln N - \ln k_i}{\ln(z_2/z_1)} + 1$$

Daul problems

- Preferential attachment

$$P[(i,j) \in E \mid k_i)] = e^{-\beta(\ell(k_i)-1)-\Gamma(\beta)} = e^{\gamma \ln k_i - \Gamma(\gamma)} = \frac{k_i^\gamma}{\sum_{k=1}^{k_{\max}} k_i^\gamma}$$

Daul problems

- Preferential attachment

$$P[(i,j) \in E \mid k_i)] = e^{-\beta(\ell(k_i)-1)-\Gamma(\beta)} = e^{\gamma \ln k_i - \Gamma(\gamma)} = \frac{k_i^\gamma}{\sum_{k=1}^{k_{\max}} k_i^\gamma}$$

- Solution to

$$\text{minimize} \quad \mathbb{E}_P\{d(i,j)-1\} \quad \text{subject to} \quad I(x_i, y_j) \leq \lambda$$

Daul problems

- Preferential attachment

$$P[(i,j) \in E \mid k_i)] = e^{-\beta(\ell(k_i)-1)-\Gamma(\beta)} = e^{\gamma \ln k_i - \Gamma(\gamma)} = \frac{k_i^\gamma}{\sum_{k=1}^{k_{\max}} k_i^\gamma}$$

- Solution to

$$\text{minimize} \quad \mathbb{E}_P\{d(i,j) - 1\} \quad \text{subject to} \quad I(x_i, y_j) \leq \lambda$$

- Power law

$$P(k) = \frac{k^{-\beta}}{\sum_{k=1}^{k_{\max}} k^{-\beta}}$$

Daul problems

- Preferential attachment

$$P[(i,j) \in E \mid k_i)] = e^{-\beta(\ell(k_i)-1)-\Gamma(\beta)} = e^{\gamma \ln k_i - \Gamma(\gamma)} = \frac{k_i^\gamma}{\sum_{k=1}^{k_{\max}} k_i^\gamma}$$

- Solution to

$$\text{minimize} \quad \mathbb{E}_P\{d(i,j) - 1\} \quad \text{subject to} \quad I(x_i, y_j) \leq \lambda$$

- Power law

$$P(k) = \frac{k^{-\beta}}{\sum_{k=1}^{k_{\max}} k^{-\beta}}$$

- Solution to

$$\text{maximize} \quad H(P) \quad \text{subject to} \quad \mathbb{E}_P\{\ln k\} \leq \upsilon$$

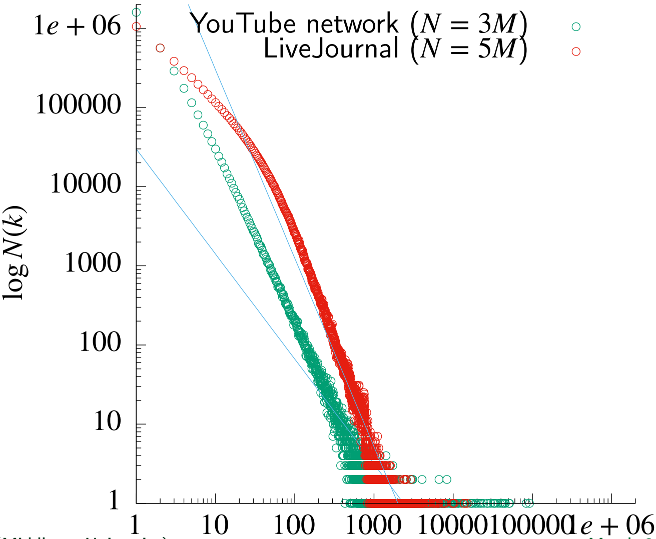
The power-law degree sequence

Power-law and maximum entropy

Power-law exponent (inverse temperature)

Free energy and phase transition

Exponent as slope



Maximum likelihood estimation

- Treating k as continuous, the m.l.e is (Newman, 2005)

$$\beta = 1 + \frac{1}{\mathbb{E}_P\{\ln k\} - \ln k_0}$$

where k_0 is the smallest degree corresponding to power-law behavious (i.e. $P(k_0) = \max P(k)$).

Maximum likelihood estimation

- Treating k as continuous, the m.l.e is (Newman, 2005)

$$\beta = 1 + \frac{1}{\mathbb{E}_P\{\ln k\} - \ln k_0}$$

where k_0 is the smallest degree corresponding to power-law behavious (i.e. $P(k_0) = \max P(k)$).

- Degree k is discrete.

Maximum likelihood estimation

- Treating k as continuous, the m.l.e is (Newman, 2005)

$$\beta = 1 + \frac{1}{\mathbb{E}_P\{\ln k\} - \ln k_0}$$

  where k_0 is the smallest degree corresponding to power-law behavious (i.e. $P(k_0) = \max P(k)$).
- Degree k is discrete.
- What about $\beta < 1$ (possible for $N < \infty$)?

Variational approach

- Recall that $P(k) = \exp\{-\beta \ln k - \Gamma(\beta)\}$ is the solution to the maximum entropy problem, where $\beta \geq 0$ is the Lagrange multiplier such that the constraint $\mathbb{E}_P\{\ln k\} \leq \upsilon$ (or $H(P) \geq \ln N - \lambda$) is satisfied with equality:

$$\mathbb{E}_P\{\ln k\} = -\Gamma'(\beta)$$
$$H(P) = \Gamma(\beta) - \beta\,\Gamma'(\beta)$$

Variational approach

- Recall that $P(k) = \exp\{-\beta \ln k - \Gamma(\beta)\}$ is the solution to the maximum entropy problem, where $\beta \geq 0$ is the Lagrange multiplier such that the constraint $\mathbb{E}_P\{\ln k\} \leq \upsilon$ (or $H(P) \geq \ln N - \lambda$) is satisfied with equality:

$$\mathbb{E}_P\{\ln k\} = -\Gamma'(\beta)$$
$$H(P) = \Gamma(\beta) - \beta\,\Gamma'(\beta)$$

- This leads to

$$\beta = \frac{H(P) - \Gamma(\beta)}{\mathbb{E}_P\{\ln k\}}$$

Variational approach

- Recall that $P(k) = \exp\{-\beta \ln k - \Gamma(\beta)\}$ is the solution to the maximum entropy problem, where $\beta \geq 0$ is the Lagrange multiplier such that the constraint $\mathbb{E}_P\{\ln k\} \leq \upsilon$ (or $H(P) \geq \ln N - \lambda$) is satisfied with equality:

$$\mathbb{E}_P\{\ln k\} = -\Gamma'(\beta)$$
$$H(P) = \Gamma(\beta) - \beta\,\Gamma'(\beta)$$

- This leads to

$$\beta = \frac{H(P) - \Gamma(\beta)}{\mathbb{E}_P\{\ln k\}}$$

- We can estimate $\mathbb{E}_P\{\ln k\}$, $H(P)$.

Variational approach

- Recall that $P(k) = \exp\{-\beta \ln k - \Gamma(\beta)\}$ is the solution to the maximum entropy problem, where $\beta \geq 0$ is the Lagrange multiplier such that the constraint $\mathbb{E}_P\{\ln k\} \leq \upsilon$ (or $H(P) \geq \ln N - \lambda$) is satisfied with equality:

$$\mathbb{E}_P\{\ln k\} = -\Gamma'(\beta)$$
$$H(P) = \Gamma(\beta) - \beta\,\Gamma'(\beta)$$

- This leads to

$$\beta = \frac{H(P) - \Gamma(\beta)}{\mathbb{E}_P\{\ln k\}}$$

- We can estimate $\mathbb{E}_P\{\ln k\}$, $H(P)$.
- Observe that $\Gamma(\beta) = -\ln P(k = 1)$.

Variational approach

- Recall that $P(k) = \exp\{-\beta \ln k - \Gamma(\beta)\}$ is the solution to the maximum entropy problem, where $\beta \geq 0$ is the Lagrange multiplier such that the constraint $\mathbb{E}_P\{\ln k\} \leq \upsilon$ (or $H(P) \geq \ln N - \lambda$) is satisfied with equality:

$$\mathbb{E}_P\{\ln k\} = -\Gamma'(\beta)$$
$$H(P) = \Gamma(\beta) - \beta\,\Gamma'(\beta)$$

- This leads to

$$\beta = \frac{H(P) - \Gamma(\beta)}{\mathbb{E}_P\{\ln k\}}$$

- We can estimate $\mathbb{E}_P\{\ln k\}$, $H(P)$.
- Observe that $\Gamma(\beta) = -\ln P(k = 1)$.
- Making the transofrmation $k \mapsto k/k_0$ leads to

$$\beta = \frac{H(P) + \ln P(k_0)}{\mathbb{E}_P\{\ln k\} - \ln k_0}$$

Exponent (inverse temperature)

- Recall the Lagrangian

$$K(P, \beta, \gamma) = H(P) + \beta[\upsilon - \mathbb{E}_P\{\ln k\}] + \gamma\left[1 - \sum P\right]$$

Exponent (inverse temperature)

- Recall the Lagrangian

$$K(P, \beta, \gamma) = H(P) + \beta[\upsilon - \mathbb{E}_P\{\ln k\}] + \gamma\left[1 - \sum P\right] = H(\upsilon)$$

- Its value at $P(k) = \exp\{-\beta \ln k - \Gamma(\beta)\}$ is the maximum entropy, which depends on the constraint $\upsilon = \mathbb{E}_P\{\ln k\}$.

Exponent (inverse temperature)

- Recall the Lagrangian

$$K(P, \beta, \gamma) = H(P) + \beta[\upsilon - \mathbb{E}_P\{\ln k\}] + \gamma \left[1 - \sum P\right] = H(\upsilon)$$

- Its value at $P(k) = \exp\{-\beta \ln k - \Gamma(\beta)\}$ is the maximum entropy, which depends on the constraint $\upsilon = \mathbb{E}_P\{\ln k\}$.
- Differentiation gives:

$$\beta = \frac{\partial H(\upsilon)}{\partial \upsilon}$$

Exponent (inverse temperature)

- Recall the Lagrangian

$$K(P, \beta, \gamma) = H(P) + \beta[\upsilon - \mathbb{E}_P\{\ln k\}] + \gamma \left[1 - \sum P\right] = H(\upsilon)$$

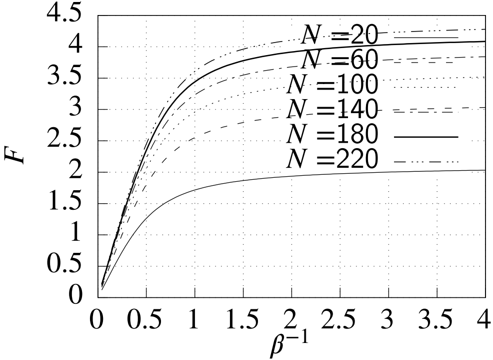
- Its value at $P(k) = \exp\{-\beta \ln k - \Gamma(\beta)\}$ is the maximum entropy, which depends on the constraint $\upsilon = \mathbb{E}_P\{\ln k\}$.
- Differentiation gives:

$$\beta = \frac{\partial H(\upsilon)}{\partial \upsilon}$$

- Compare with our formula

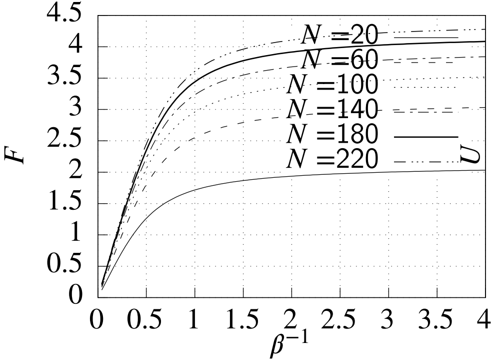
$$\beta = \frac{H(P) - \ln P^{-1}(k_0)}{\mathbb{E}_P\{\ln k\} - \ln k_0} = \frac{\Delta H}{\Delta \upsilon}$$

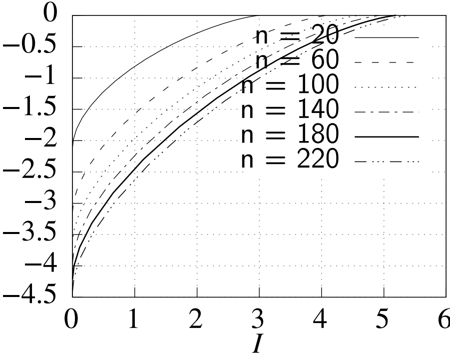
Free energy



$$F(\beta^{-1}) := -\beta^{-1}\Gamma(\beta)$$

Free energy

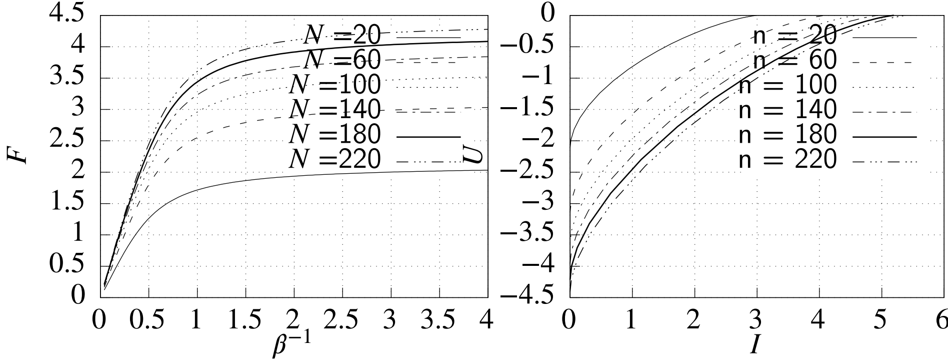




$F(\beta^{-1}) := -\beta^{-1}\Gamma(\beta)$

$F^*(I) := \inf\{\beta^{-1}I - F(\beta^{-1})\}$

Free energy

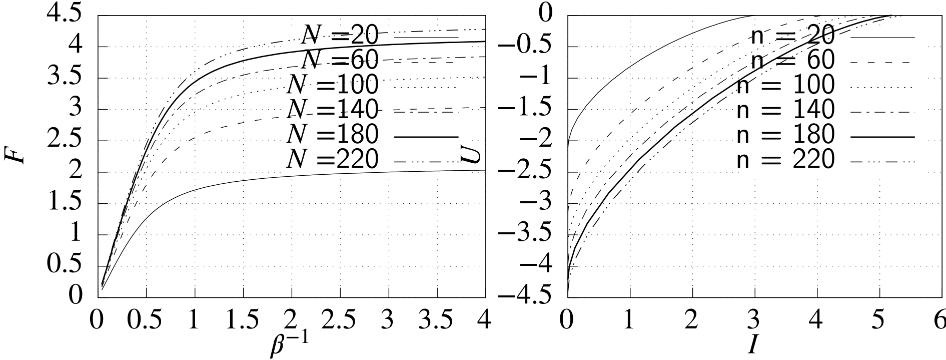


$$F(\beta^{-1}) := -\beta^{-1}\Gamma(\beta)$$

$$F^*(I) := \inf\{\beta^{-1}I - F(\beta^{-1})\}$$

$$I = F'(\beta^{-1})$$

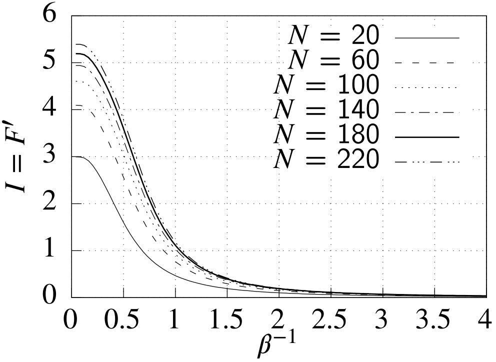
Free energy



$$F(\beta^{-1}) := -\beta^{-1}\Gamma(\beta) \qquad F^*(I) := \inf\{\beta^{-1}I - F(\beta^{-1})\}$$

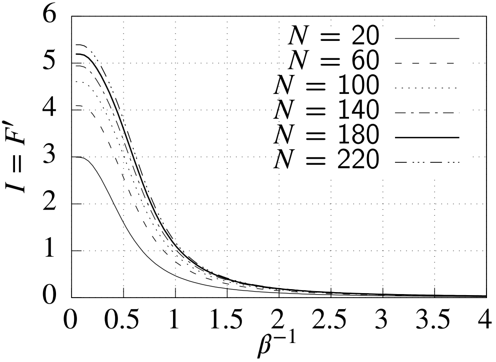
$$F^{*\prime}(I) = \beta^{-1} \qquad I = F'(\beta^{-1})$$

Information and entropy at $\beta = 1$



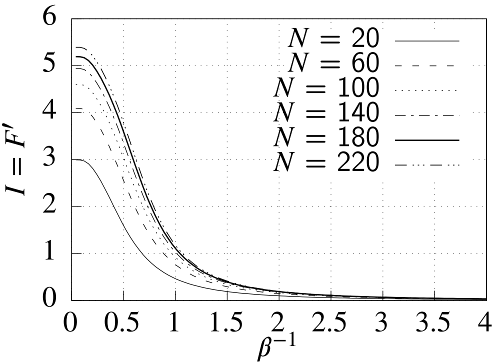
- $H(\beta) = \ln N - I(\beta)$

Information and entropy at $\beta = 1$



- $H(\beta) = \ln N - I(\beta)$
- $I(\beta) = \ln N - H(\beta)$

Information and entropy at $\beta = 1$

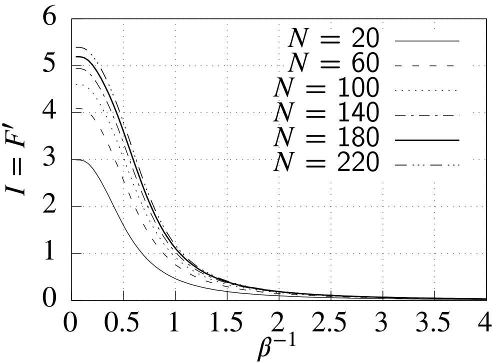


- $H(\beta) = \ln N - I(\beta)$
- $I(\beta) = \ln N - H(\beta)$

Question

- *What happens to $F'(\beta^{-1}) = I$ in the limit $N \to \infty$ and $\beta \to 1$?*

Information and entropy at $\beta = 1$

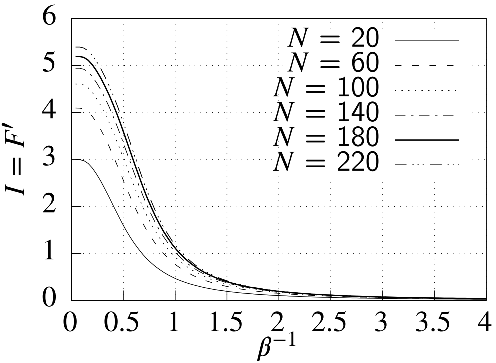


- $H(\beta) = \ln N - I(\beta)$
- $I(\beta) = \ln N - H(\beta)$

Question

- *What happens to* $F'(\beta^{-1}) = I$ *in the limit* $N \to \infty$ *and* $\beta \to 1$*?*
- $I = \beta\,\Gamma'(\beta) - \Gamma(\beta)$

Information and entropy at $\beta = 1$



- $H(\beta) = \ln N - I(\beta)$
- $I(\beta) = \ln N - H(\beta)$

Question

- *What happens to $F'(\beta^{-1}) = I$ in the limit $N \to \infty$ and $\beta \to 1$?*
- $I = \beta\,\Gamma'(\beta) - \Gamma(\beta)$
- $\Gamma'(\beta) = -\mathbb{E}_P\{\ln k\}$, $\Gamma''(\beta) = \sigma^2(\ln k)$.

Approximations at $\beta = 1$ and $N < \infty$

- nth cumulants $\Gamma^{(n)} = (\ln Z)^{(n)}$:

$$\Gamma' = m_1$$
$$\Gamma'' = m_2 - m_1^2$$
$$\Gamma^{(3)} = m_3 - 3m_1 m_2 + 2m_1^3$$
$$\Gamma^{(4)} = m_4 - 4m_3 m_1 - 3m_2^2 + 12 m_2 m_1^2 - 6m_1^4$$

Approximations at $\beta = 1$ and $N < \infty$

- nth cumulants $\Gamma^{(n)} = (\ln Z)^{(n)}$:

$$\begin{aligned}
\Gamma' &= m_1 \\
\Gamma'' &= m_2 - m_1^2 \\
\Gamma^{(3)} &= m_3 - 3m_1 m_2 + 2m_1^3 \\
\Gamma^{(4)} &= m_4 - 4m_3 m_1 - 3m_2^2 + 12 m_2 m_1^2 - 6m_1^4
\end{aligned}$$

- where m_n are nth moments of $-\ln k$:

$$m_n(\beta) = \frac{Z^{(n)}(\beta)}{Z(\beta)}, \qquad Z^{(n)}(\beta) = (-1)^n \sum_{k=1}^{N} (\ln k)^n k^{-\beta}$$

Approximations at $\beta = 1$ and $N < \infty$

- nth cumulants $\Gamma^{(n)} = (\ln Z)^{(n)}$:

$$\Gamma' = m_1$$
$$\Gamma'' = m_2 - m_1^2$$
$$\Gamma^{(3)} = m_3 - 3m_1 m_2 + 2m_1^3$$
$$\Gamma^{(4)} = m_4 - 4m_3 m_1 - 3m_2^2 + 12m_2 m_1^2 - 6m_1^4$$

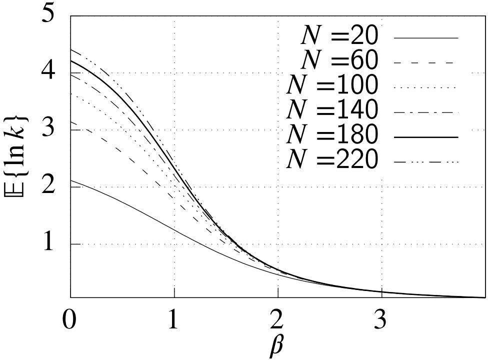
- where m_n are nth moments of $-\ln k$:

$$m_n(\beta) = \frac{Z^{(n)}(\beta)}{Z(\beta)}, \qquad Z^{(n)}(\beta) = (-1)^n \sum_{k=1}^{N} (\ln k)^n k^{-\beta}$$

- Using $\int_1^N \frac{dx}{x} = \ln N$:

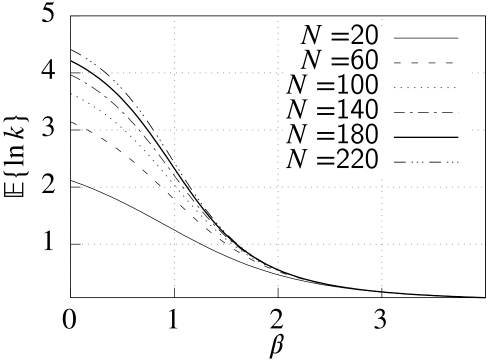
$$Z(\beta) = \sum_{k=1}^{N} \frac{1}{k^\beta}\bigg|_{\beta=1} \approx \ln N, \quad Z^{(n)}(\beta)\Big|_{\beta=1} \approx \frac{(-1)^n}{n+1} (\ln N)^{n+1}$$

Expectation of $\ln k$ at $\beta = 1$



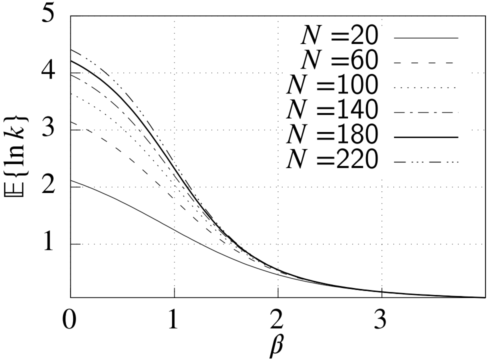
$$\mathbb{E}_P\{\ln k\} = -\Gamma'(\beta)$$

Expectation of $\ln k$ at $\beta = 1$



$$\mathbb{E}_P\{\ln k\} = -\Gamma'(\beta) = -\left.\frac{Z'(\beta)}{Z(\beta)}\right|_{\beta=1}$$

Expectation of $\ln k$ at $\beta = 1$

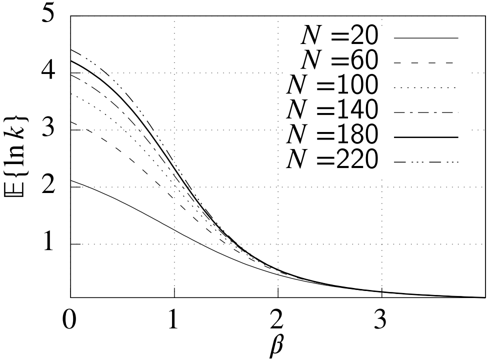


$$\mathbb{E}_P\{\ln k\} = -\Gamma'(\beta)$$

$$= -\left.\frac{Z'(\beta)}{Z(\beta)}\right|_{\beta=1}$$

$$\approx \frac{1}{2}\ln N$$

Expectation of $\ln k$ at $\beta = 1$

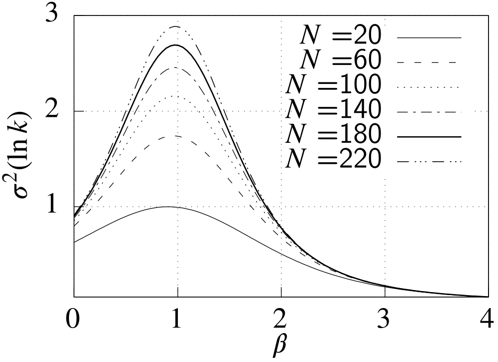


$$\mathbb{E}_P\{\ln k\} = -\Gamma'(\beta) = -\left.\frac{Z'(\beta)}{Z(\beta)}\right|_{\beta=1} \approx \frac{1}{2}\ln N$$

Remark

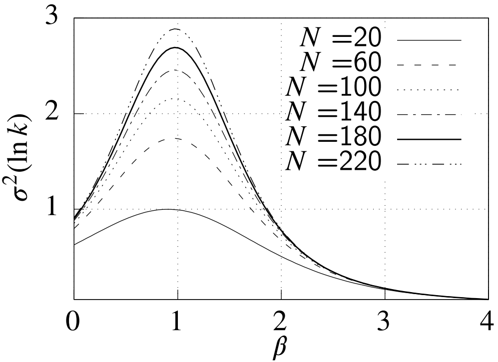
Using Jenssen's inequality $\ln \mathbb{E}_P\{k\} \geq \mathbb{E}_P\{\ln k\}$ *we also have* $\mathbb{E}_P\{k\} \geq \sqrt{N}$.

Variance of $\ln k$ at $\beta = 1$



$$\sigma^2(\ln k) = \Gamma''(\beta)\big|_{\beta=1} \approx \frac{1}{12}(\ln N)^2$$

Variance of $\ln k$ at $\beta = 1$

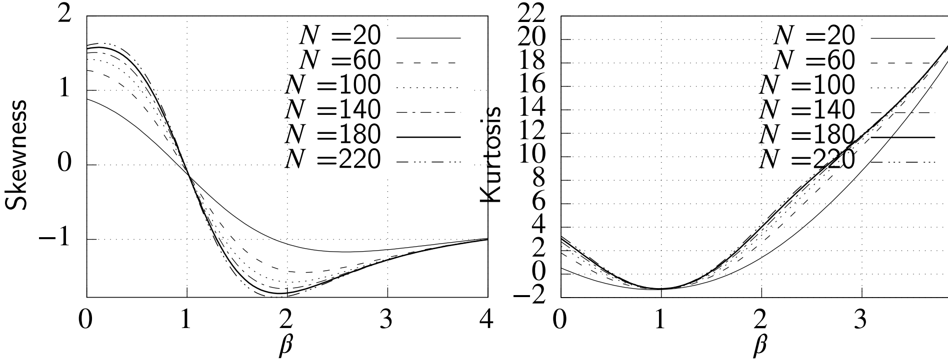


$$\sigma^2(\ln k) = \Gamma''(\beta)\big|_{\beta=1}$$
$$\approx \frac{1}{12}(\ln N)^2$$

Remark (Phase transition)

The derivative $F'(\beta^{-1}) = I(\beta) = \beta\,\Gamma'(\beta) - \Gamma(\beta)$ *is not differentiable at* $\beta = 1$ *in the limit* $N \to \infty$*, because* $\Gamma''(\beta) \to \infty$*.*

Skewness and kurtosis at $\beta = 1$



$$\frac{\Gamma^{(3)}}{(\Gamma'')^{3/2}} \qquad \frac{\Gamma^{(4)}}{(\Gamma'')^{2}}$$

Giant component

- General condition of existance (Molloy & Reed, 1995; Albert & Barabási, 2002)

$$Q := \mathbb{E}_P\{k^2\} - 2\mathbb{E}_P\{k\} \geq 0$$

Giant component

- General condition of existance (Molloy & Reed, 1995; Albert & Barabási, 2002)

$$Q := \mathbb{E}_P\{k^2\} - 2\mathbb{E}_P\{k\} \geq 0$$

- Observe that $Q = \mathbb{E}_P\{(k-1)^2\} - 1$, which means

$$\mathbb{E}_P\{(k-1)^2\} \geq 1$$

Giant component

- General condition of existance (Molloy & Reed, 1995; Albert & Barabási, 2002)

$$Q := \mathbb{E}_P\{k^2\} - 2\mathbb{E}_P\{k\} \geq 0$$

- Observe that $Q = \mathbb{E}_P\{(k-1)^2\} - 1$, which means

$$\mathbb{E}_P\{(k-1)^2\} \geq 1$$

- On the other hand

$$\inf_a \mathbb{E}_P\{(k-a)^2\} = \mathbb{E}_P\{(k - \mathbb{E}_P\{k\})^2\} =: \sigma^2(k)$$

Giant component

- General condition of existance (Molloy & Reed, 1995; Albert & Barabási, 2002)

$$Q := \mathbb{E}_P\{k^2\} - 2\mathbb{E}_P\{k\} \geq 0$$

- Observe that $Q = \mathbb{E}_P\{(k-1)^2\} - 1$, which means

$$\mathbb{E}_P\{(k-1)^2\} \geq 1$$

- On the other hand

$$\inf_a \mathbb{E}_P\{(k-a)^2\} = \mathbb{E}_P\{(k - \mathbb{E}_P\{k\})^2\} =: \sigma^2(k)$$

- Therefore, $\sigma^2(k) \geq 1$ implies $Q \geq 0$.

Giant component

- General condition of existance (Molloy & Reed, 1995; Albert & Barabási, 2002)
$$Q := \mathbb{E}_P\{k^2\} - 2\mathbb{E}_P\{k\} \geq 0$$
- Observe that $Q = \mathbb{E}_P\{(k-1)^2\} - 1$, which means
$$\mathbb{E}_P\{(k-1)^2\} \geq 1$$
- On the other hand
$$\inf_a \mathbb{E}_P\{(k-a)^2\} = \mathbb{E}_P\{(k - \mathbb{E}_P\{k\})^2\} =: \sigma^2(k)$$
- Therefore, $\sigma^2(k) \geq 1$ implies $Q \geq 0$.
- $k - 1 \geq \ln k$, and $k - 1$ approximates $\ln k$ near $k = 1$.

Giant component

- General condition of existance (Molloy & Reed, 1995; Albert & Barabási, 2002)
$$Q := \mathbb{E}_P\{k^2\} - 2\mathbb{E}_P\{k\} \geq 0$$
- Observe that $Q = \mathbb{E}_P\{(k-1)^2\} - 1$, which means
$$\mathbb{E}_P\{(k-1)^2\} \geq 1$$
- On the other hand
$$\inf_a \mathbb{E}_P\{(k-a)^2\} = \mathbb{E}_P\{(k - \mathbb{E}_P\{k\})^2\} =: \sigma^2(k)$$
- Therefore, $\sigma^2(k) \geq 1$ implies $Q \geq 0$.
- $k - 1 \geq \ln k$, and $k - 1$ approximates $\ln k$ near $k = 1$.
- $\sigma^2(k) = 1$ gives $\beta_0 \approx 3.466407\ldots$, and $\Gamma'(\beta) = \zeta'(\beta)/\zeta(\beta) = 1$ gives $\beta_0 \approx 1.6042\ldots$.

Network	$\lvert V\rvert$	LCC	$\max k$	$\mathbb{E}\{k\}$	$\mathbb{E}\{\ln k\}$	$\sigma(k)$	H
Zebra	27	23	14	8.2222	1.8819	22.7	49%
Karate club	34	34	17	4.5882	1.2805	14.6	57%
Train bombing	64	64	29	7.5938	1.6586	38.0	64%
Protein	212	161	21	2.3019	0.4918	7.0	26%
Email	1,133	1,133	71	9.6222	1.7822	87.2	45%
Euroroad	1,174	1,039	10	2.4140	0.7753	1.4	20%
US power	4,941	4,941	19	2.6691	0.8021	3.2	58%
Routes	6,474	6,474	1,459	4.2926	0.7948	628.9	22%
arXiv-ph	18,771	17,903	504	21.102	2.2768	934.2	40%
Facebook	63,731	63,392	1,098	25.640	2.3096	1599.4	37%
Flickr	105,938	105,722	5,425	43.742	2.4556	13,359.7	38%
DBLP-authorship	317,080	317,080	343	6.6221	1.3889	100.2	21%
Amazon-purchase	334,863	334,863	549	5.5299	1.4467	33.2	19%
YouTube	3,223,589	3,216,075	91,751	5.8167	0.7474	16,435	17%
LiveJournal	5,204,176	5,189,808	15,016	18.721	1.8562	2557.9	37%

Conclusions

- We show how the power-law graphs emerge as the solutions to variational problem maximizing entropy with a constraint on $\mathbb{E}\{\ln k\}$.

Conclusions

- We show how the power-law graphs emerge as the solutions to variational problem maximizing entropy with a constraint on $\mathbb{E}\{\ln k\}$.
- The negative log-degree $-\ln k$ is related to average distance $\ell(k)$, and preferential attachment emerges as The dual problem of the solution to variational problem on miniming average distnances with a constraint on mutual information.

Conclusions

- We show how the power-law graphs emerge as the solutions to variational problem maximizing entropy with a constraint on $\mathbb{E}\{\ln k\}$.
- The negative log-degree $-\ln k$ is related to average distance $\ell(k)$, and preferential attachment emerges as The dual problem of the solution to variational problem on miniming average distnances with a constraint on mutual information.
- The exponent parameter can be estimated as inverse temperature using variational principle.

Conclusions

- We show how the power-law graphs emerge as the solutions to variational problem maximizing entropy with a constraint on $\mathbb{E}\{\ln k\}$.
- The negative log-degree $-\ln k$ is related to average distance $\ell(k)$, and preferential attachment emerges as The dual problem of the solution to variational problem on miniming average distnances with a constraint on mutual information.
- The exponent parameter can be estimated as inverse temperature using variational principle.
- Power-law graphs undergo a phase transition for finite value $\beta \in [1, \beta_0]$:

$$\underbrace{\text{connected}}_{\beta<1}\underbrace{\text{disconnected, giant component}}_{[1,\beta_0]}\underbrace{\text{no giant component}}_{\beta_0<\beta}$$

The power-law degree sequence

Power-law and maximum entropy

Power-law exponent (inverse temperature)

Free energy and phase transition

Aiello, W., Chung, F., & Lu, L. (2000). A random graph model for massive graphs. In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing* (pp. 171–180). Portland, OR: ACM Press.

Aiello, W., Chung, F., & Lu, L. (2001). A random graph model for power law graphs. *Experimental Mathematics*, *10*(1), 53–66.

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*, 47–97.

Molloy, M., & Reed, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, *6*, 161–180.

Newman, M. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, *46*(5), 323–351.