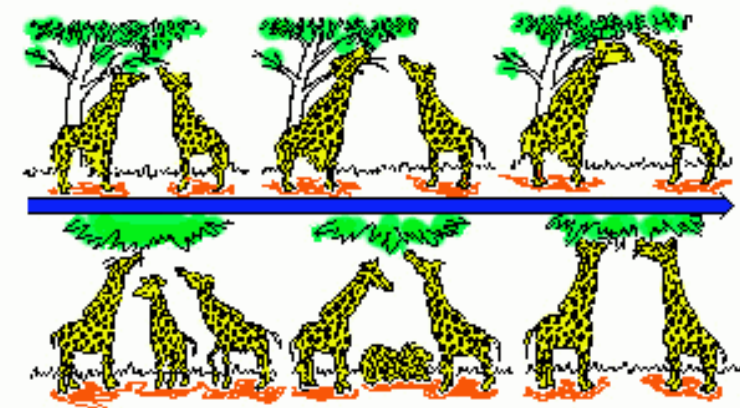


# 複雑系というシステム

## 複雑系という知能

東京大学大学院  
情報理工学系科学研究科  
電気情報学専攻  
伊庭齊志

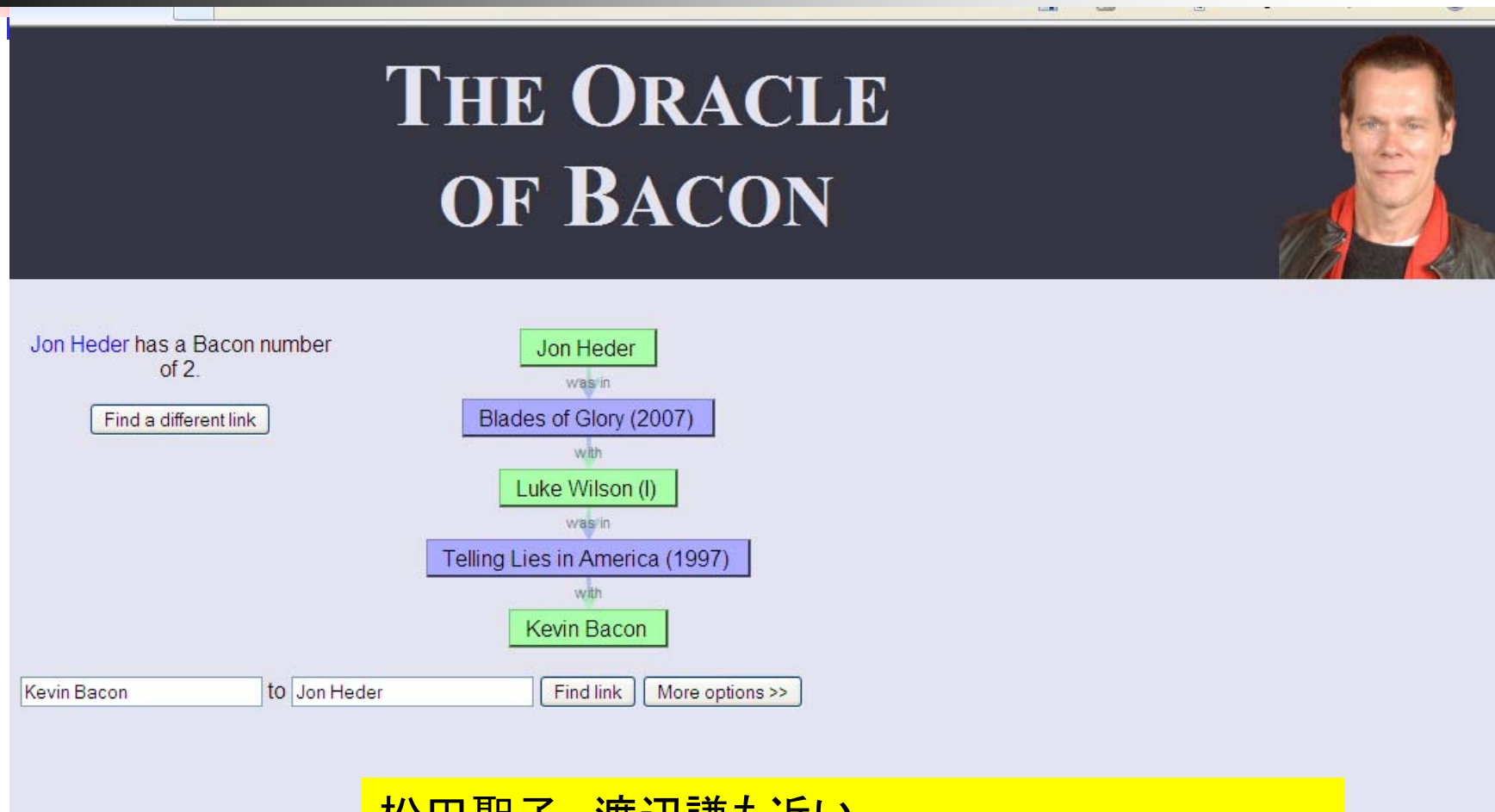




# ネットワークの例

Network	Nodes	Edges
Societies	People	Friendships
Hollywood	Actors	Movies
Internet (virtual)	Web Pages	Links
Internet (Physical)	Routers	Communication Lines
Research	Papers	Citations
Cellular Metabolism	Molecules	Biochemical Reactions
Epidemics	Hosts	Infections
Health Disorders	Diseases	Genes
Nervous Systems	Neurons	Synapses
Economic Systems	Business Entities	Loans

# The Oracle of Kevin Bacon



**THE ORACLE OF BACON**

Jon Heder has a Bacon number of 2.

```
graph TD; A[Jon Heder] -- was in --> B[Blades of Glory (2007)]; B -- with --> C[Luke Wilson (I)]; C -- was in --> D[Telling Lies in America (1997)]; D -- with --> E[Kevin Bacon];
```

Kevin Bacon to Jon Heder

松田聖子、渡辺謙も近い

<http://oracleofbacon.org/movielinks.php>

# Erdős Number : エルデス数 (ベーコンゲームの天才版)



My Brain is Open—20世紀数学界の異才ポール・エルデシュ放浪記, ポール ホフマン, 草思社

Baconと違って, Erdősは彼のネットワークの中心として良いポジション

論文の共著関係により、Erdős と研究者をつなげるの必要なリンク数

Erdős は 1500以上の論文を、507人の共著者と書いた

Jerry Grossman (Oakland Univ.)のwebsiteから 数学者のErdős数を計算できる

<http://www.oakland.edu/enp/>

数学者の間のみでのパス長:

- 平均 4.65
- 最大値 13

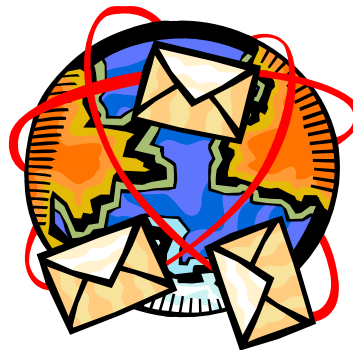
# Six Degrees of Separation

## 6次の隔たり

Stanley Milgram (1967)

### 実験:

- ネブラスカ州オハマに住む160人の住人をランダムに選択し、それぞれの知人へ手紙を転送しながら、ボストン在住のある人へその手紙を届けるよう依頼した



ほとんどの手紙が5段階か6段階の知人を経ることでゴールのあて先に届いていることが分かった。

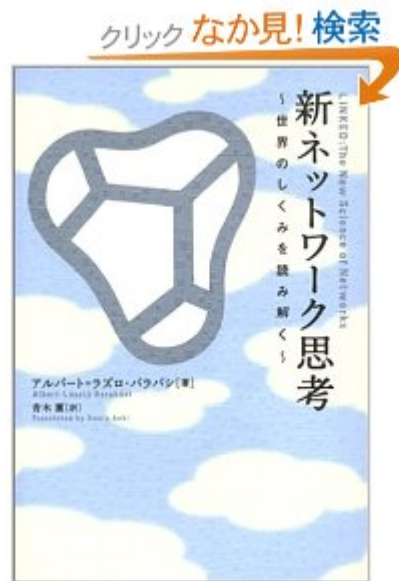


Stanley Milgram (1933-1984)

しかしながらこの実験にはいくつかの疑問点がある

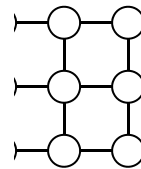
# 正規ネットワーク

## STAR

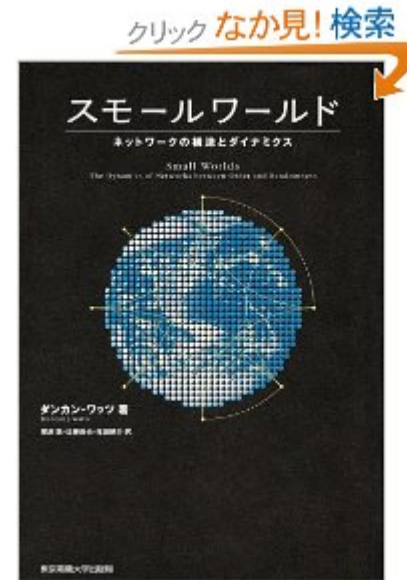


新ネットワーク思考—世界のしくみを読み解く,  
アルバート・ラズロ・バラバシ, NHK出版

## ID



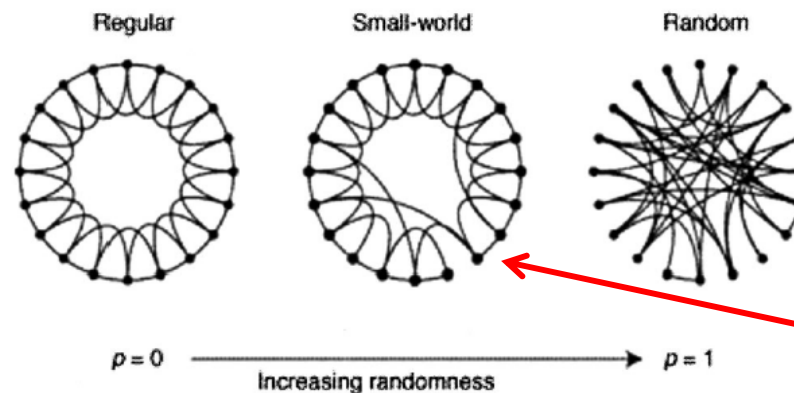
## TREE



スモールワールド—ネットワークの構造とダイナミクス,  
ダンカン ワッツ, 東京電機大学出版局

# スモールワールド・ネットワーク

- 正規ネットワークのランダムな書き換え (by **Watts and Strogatz**)
  - 確率  $p$  (or  $\beta$ ) で、正規グラフの各リンクをランダムに選ばれたノードにつなぎかえる
  - 結果としてできるグラフは正規性とランダム性の両方の性質を持つ
    - 高いクラスタ性を示し、かつ短いパス長
  - FreeNetはスモールワールドグラフであることが示されている
    - JAVA仮想マシン上で実行できるP2Pソフトウェア



**Small-world性**をもつ

- 長距離リンクは少ない
- ノード総数に比べて平均最短距離が小さい

クラスタを形成しやすい特徴をもつ



# スモールワールド性とは?

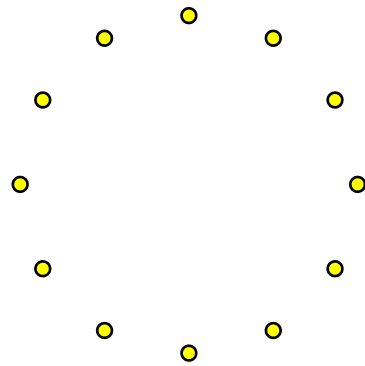
---

- 仮説: 進化には2つの相反する圧力が働く
  - システム内で迅速に情報を伝達する
  - 遠距離の結合を維持する
- このためにはスモールワールドが適している
  
- しかし、正規ネットワークをランダムに付け加えていくモデルでは、現実世界のネットワークの次数分布を説明できないことがわかった
  
- スケールフリーネットワークの登場
  - スモールワールドネットワークの特殊形態

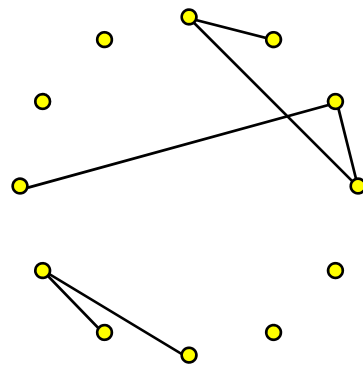


# ランダムグラフ

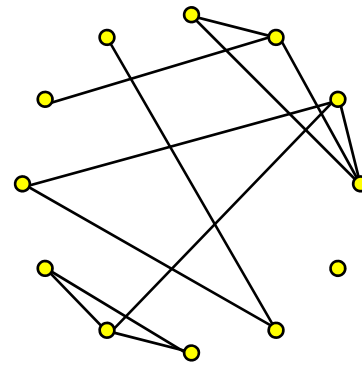
Erdős and Renyi (1959)



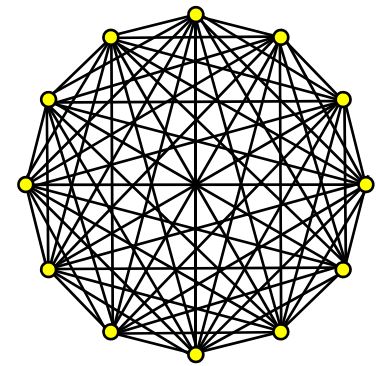
$p = 0.0 ; k = 0$



$p = 0.045 ; k = 0.5$



$p = 0.09 ; k = 1$



$p = 1.0 ; k \approx N$

最大クラスタのサイズ

1

5

11

12

最大クラスタの直径

0

4

7

1

ノード間の平均パス長

0.0

2.0

4.2

1.0

# ランダムグラフ

Erdős and Renyi (1959)

If  $k < 1$ :

- 小さく、孤立したクラスタ
- 小さな直径
- 短いパス長

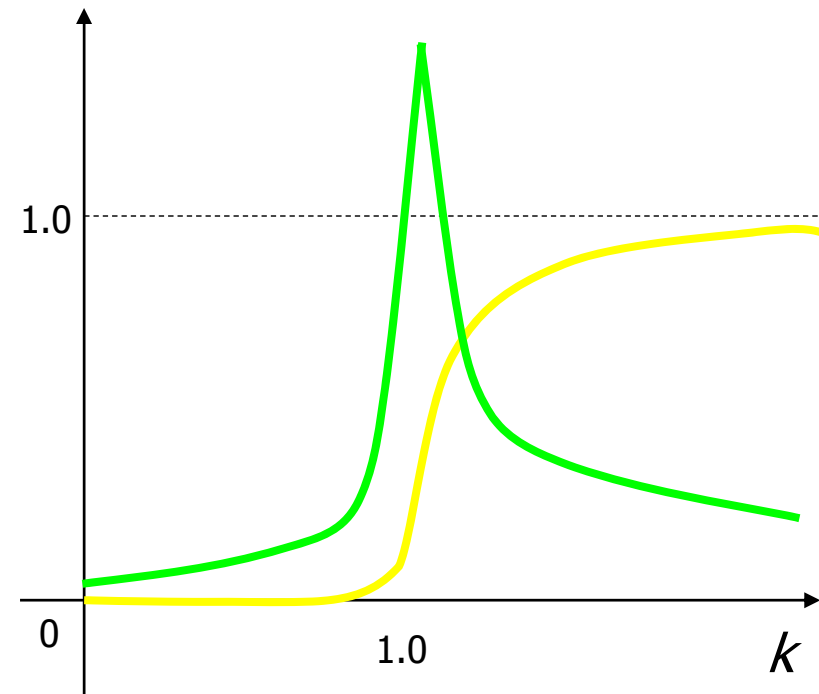
At  $k = 1$ :

- 巨大な要素の出現
- 直径のピーク
- パス長は大きい

For  $k > 1$ :

- ほとんどすべてのノードが連結する
- 直径が縮小
- パス長は短くなる

最大クラスタのノードの割合  
最大クラスタの直径(縮尺率は一定ではない)



↑  
相転移

# スケールフリー性

人口はスケールフリー性を持っていることはすでに知られていた!

1955年に、Herbert Simonはすでに多くのシステムがべき乗則の分布に従うことを示していた。そのためこの結果は新規でも唯一でもない。

- 地震の大きさ
- 月のクレータ
- 太陽フレア
- コンピュータのファイル



- 戦争
- 論文の引用回数
- Webページのヒット数
- 人々の年間収入

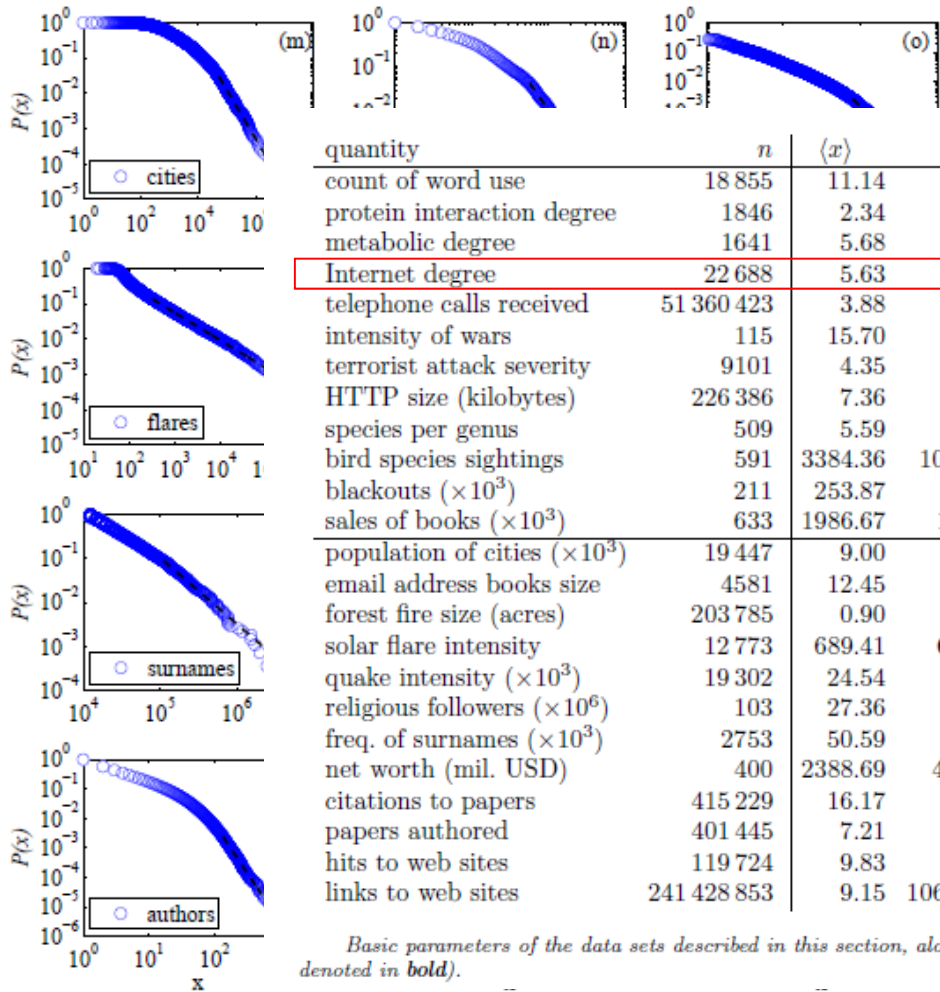


ノードの次数分布は多くのネットワークでべき乗則に従うことがわかっている。つまり多くのネットワークはスケールフリーである。

重要なことはべき乗則を見出すことではない(むしろありふれている)

なぜそうなるのか、他のどんな構造的パラメータがあるのかを理解することである。

# スケールフリー性



"POWER-LAW DISTRIBUTIONS IN EMPIRICAL DATA"  
 AARON CLAUSET, COSMA ROHILLA SHALIZI, AND M.E.J. NEWMAN

quantity	$n$	$\langle x \rangle$	$\sigma$	$x_{\max}$	$\hat{x}_{\min}$	$\hat{\alpha}$	$n_{\text{tail}}$	$p$
count of word use	18 855	11.14	148.33	14 086	$7 \pm 2$	1.95(2)	$2958 \pm 987$	<b>0.49</b>
protein interaction degree	1846	2.34	3.05	56	$5 \pm 2$	3.1(3)	$204 \pm 263$	<b>0.31</b>
metabolic degree	1641	5.68	17.81	468	$4 \pm 1$	2.8(1)	$748 \pm 136$	0.00
Internet degree	22 688	5.63	37.83	2583	$21 \pm 9$	2.12(9)	$770 \pm 1124$	<b>0.29</b>
telephone calls received	51 360 423	3.88	179.09	375 746	$120 \pm 49$	2.09(1)	$102 592 \pm 210 147$	<b>0.63</b>
intensity of wars	115	15.70	49.97	382	$2.1 \pm 3.5$	1.7(2)	$70 \pm 14$	<b>0.20</b>
terrorist attack severity	9101	4.35	31.58	2749	$12 \pm 4$	2.4(2)	$547 \pm 1663$	<b>0.68</b>
HTTP size (kilobytes)	226 386	7.36	57.94	10 971	$36.25 \pm 22.74$	2.48(5)	$6794 \pm 2232$	0.00
species per genus	509	5.59	6.94	56	$4 \pm 2$	2.4(2)	$233 \pm 138$	<b>0.10</b>
bird species sightings	591	3384.36	10 952.34	138 705	$6679 \pm 2463$	2.1(2)	$66 \pm 41$	<b>0.55</b>
blackouts ( $\times 10^3$ )	211	253.87	610.31	7500	$230 \pm 90$	2.3(3)	$59 \pm 35$	<b>0.62</b>
sales of books ( $\times 10^3$ )	633	1986.67	1396.60	19 077	$2400 \pm 430$	3.7(3)	$139 \pm 115$	<b>0.66</b>
population of cities ( $\times 10^3$ )	19 447	9.00	77.83	8 009	$52.46 \pm 11.88$	2.37(8)	$580 \pm 177$	<b>0.76</b>
email address books size	4581	12.45	21.49	333	$57 \pm 21$	3.5(6)	$196 \pm 449$	<b>0.16</b>
forest fire size (acres)	203 785	0.90	20.99	4121	$6324 \pm 3487$	2.2(3)	$521 \pm 6801$	0.05
solar flare intensity	12 773	689.41	6520.59	231 300	$323 \pm 89$	1.79(2)	$1711 \pm 384$	<b>1.00</b>
quake intensity ( $\times 10^3$ )	19 302	24.54	563.83	63 096	$0.794 \pm 80.198$	1.64(4)	$11 697 \pm 2159$	0.00
religious followers ( $\times 10^6$ )	103	27.36	136.64	1050	$3.85 \pm 1.60$	1.8(1)	$39 \pm 26$	<b>0.42</b>
freq. of surnames ( $\times 10^3$ )	2753	50.59	113.99	2502	$111.92 \pm 40.67$	2.5(2)	$239 \pm 215$	<b>0.20</b>
net worth (mil. USD)	400	2388.69	4 167.35	46 000	$900 \pm 364$	2.3(1)	$302 \pm 77$	0.00
citations to papers	415 229	16.17	44.02	8904	$160 \pm 35$	3.16(6)	$3455 \pm 1859$	<b>0.20</b>
papers authored	401 445	7.21	16.52	1416	$133 \pm 13$	4.3(1)	$988 \pm 377$	<b>0.90</b>
hits to web sites	119 724	9.83	392.52	129 641	$2 \pm 13$	1.81(8)	$50 981 \pm 16 898$	0.00
links to web sites	241 428 853	9.15	106 871.65	1 199 466	$3684 \pm 151$	2.336(9)	$28 986 \pm 1560$	0.00

TABLE 6.1

Basic parameters of the data sets described in this section, along with their power-law fits and the corresponding  $p$ -value (statistically significant values are denoted in **bold**).

# スケールフリー性

"POWER-LAW DISTRIBUTIONS IN EMPIRICAL DATA"  
AARON CLAUSET, COSMA ROHILLA SHALIZI, AND M.E.J.  
NEWMAN

data set	power law	log-normal		exponential		stretched exp.		power law + cut-off		support for power law
	$p$	LR	$p$	LR	$p$	LR	$p$	LR	$p$	
birds	<b>0.55</b>	-0.850	0.40	1.87	<b>0.06</b>	-0.882	0.38	-1.24	0.12	moderate
blackouts	<b>0.62</b>	-0.412	0.68	1.21	0.23	-0.417	0.68	-0.382	0.38	moderate
book sales	<b>0.66</b>	-0.267	0.79	2.70	<b>0.01</b>	3.885	<b>0.00</b>	-0.140	0.60	moderate
cities	<b>0.76</b>	-0.090	0.93	3.65	<b>0.00</b>	0.204	0.84	-0.123	0.62	moderate
fires	0.05	-1.78	<b>0.08</b>	4.00	<b>0.00</b>	-1.82	<b>0.07</b>	-5.02	<b>0.00</b>	with cut-off
flares	<b>1.00</b>	-0.803	0.42	13.7	<b>0.00</b>	-0.546	0.59	-4.52	<b>0.00</b>	with cut-off
HTTP	0.00	1.77	<b>0.08</b>	11.8	<b>0.00</b>	2.65	<b>0.01</b>	0.000	1.00	none
quakes	0.00	-7.14	<b>0.00</b>	11.6	<b>0.00</b>	-7.09	<b>0.00</b>	-24.4	<b>0.00</b>	with cut-off
religions	<b>0.42</b>	-0.073	0.94	1.59	0.11	1.75	<b>0.08</b>	-0.167	0.56	moderate
surnames	<b>0.20</b>	-0.836	0.40	2.89	<b>0.00</b>	-0.844	0.40	-1.36	<b>0.10</b>	with cut-off
wars	<b>0.20</b>	-0.737	0.46	3.68	<b>0.00</b>	-0.767	0.44	-0.847	0.19	moderate
wealth	0.00	0.249	0.80	6.20	<b>0.00</b>	8.05	<b>0.00</b>	-0.142	0.59	none
web hits	0.00	-10.21	<b>0.00</b>	8.55	<b>0.00</b>	10.94	<b>0.00</b>	-74.66	<b>0.00</b>	with cut-off
web links	0.00	-2.24	<b>0.03</b>	25.3	<b>0.00</b>	-1.08	0.28	-21.2	<b>0.00</b>	with cut-off

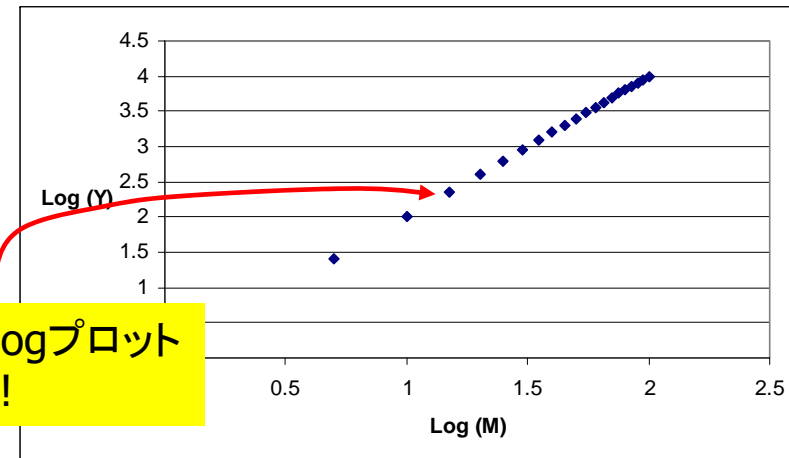
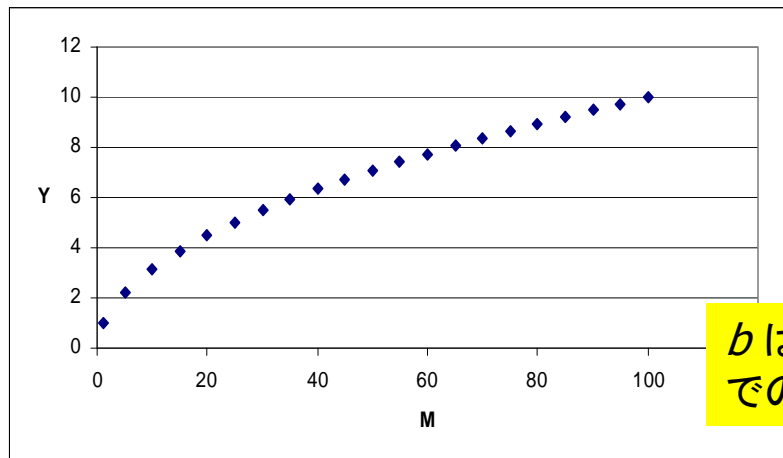
data set	$p$	Poisson		log-normal		exponential		stretched exp.		power law + cut-off		support for power law
		LR	$p$	LR	$p$	LR	$p$	LR	$p$	LR	$p$	
Internet	<b>0.29</b>	5.31	<b>0.00</b>	-0.807	0.42	6.49	<b>0.00</b>	0.493	0.62	-1.97	<b>0.05</b>	with cut-off
calls	<b>0.63</b>	17.9	<b>0.00</b>	-2.03	<b>0.04</b>	35.0	<b>0.00</b>	14.3	<b>0.00</b>	-30.2	<b>0.00</b>	with cut-off
citations	<b>0.20</b>	6.54	<b>0.00</b>	-0.141	0.89	5.91	<b>0.00</b>	1.72	<b>0.09</b>	-0.007	0.91	moderate
email	<b>0.16</b>	4.65	<b>0.00</b>	-1.10	0.27	0.639	0.52	-1.13	0.26	-1.89	<b>0.05</b>	with cut-off
metabolic	0.00	3.53	<b>0.00</b>	-1.05	0.29	5.59	<b>0.00</b>	3.66	<b>0.00</b>	0.000	1.00	none
papers	<b>0.90</b>	5.71	<b>0.00</b>	-0.091	0.93	3.08	<b>0.00</b>	0.709	0.48	-0.016	0.86	moderate
proteins	<b>0.31</b>	3.05	<b>0.00</b>	-0.456	0.65	2.21	<b>0.03</b>	0.055	0.96	-0.414	0.36	moderate
species	<b>0.10</b>	5.04	<b>0.00</b>	-1.63	0.10	2.39	<b>0.02</b>	-1.59	0.11	-3.80	<b>0.01</b>	with cut-off
terrorism	<b>0.68</b>	1.81	<b>0.07</b>	-0.278	0.78	2.457	<b>0.01</b>	0.772	0.44	-0.077	0.70	moderate
words	<b>0.49</b>	4.43	<b>0.00</b>	0.395	0.69	9.09	<b>0.00</b>	4.13	<b>0.00</b>	-0.899	0.18	good

# スケールフリー性

- Log-logプロットと通常のプロット

$$Y = Y_0 M^b$$

$$\text{Log } Y = \text{Log } Y_0 + b \text{ log } M$$



**b** はlog-logプロットでの傾き!

## なぜハブがあるのか？

- ランダムネットワークモデルはすべてのノードがネットワーク生成の最初から存在することを仮定している。これは通常ではありえない。
- スケールフリーネットワークでは、古いノードはリンクを得る可能性がより大きい。
- 優先的選択： 富める者はますます富む





## ネットワークの頑強さ

- ランダムに選ばれたノードへのアタックに対する強さ
- 80%のノードが死んだ場合でもネットワークはまだ機能する
- ただしハブへの計画的な攻撃には弱い
- 5-15%のハブが死んだだけでもネットワーククラッシュとなる

ハブの割合は全体のノードに比べて少ない!!