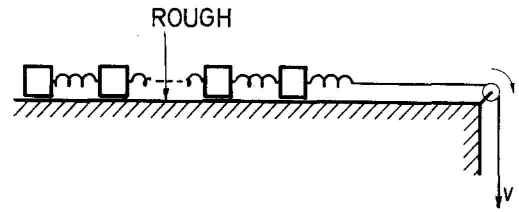
Large Fluctuations and Extreme Events, Dresden, October 2015

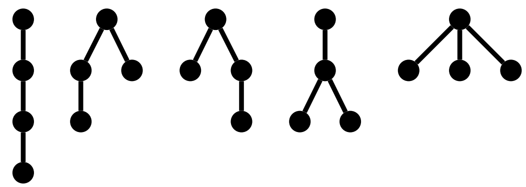
Power Laws and Criticality

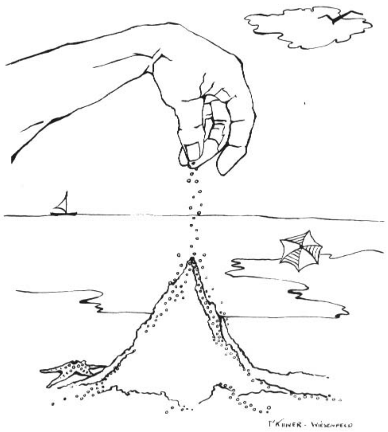
Álvaro Corral

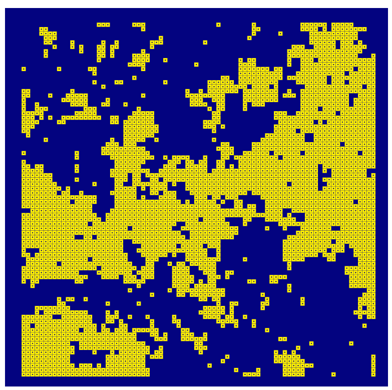
CENTRE DE RECERCA MATEMÀTICA

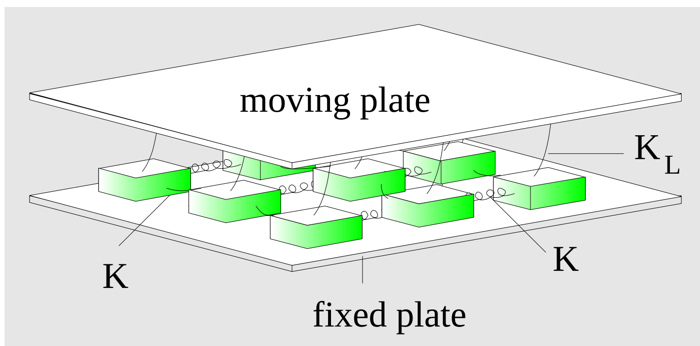
Univ. Autònoma de Barcelona
Spain











1. **The size of earthquakes** (and other natural hazards)

2. **Properties of power-law distributions**
   - Scale invariance. Divergence of moments

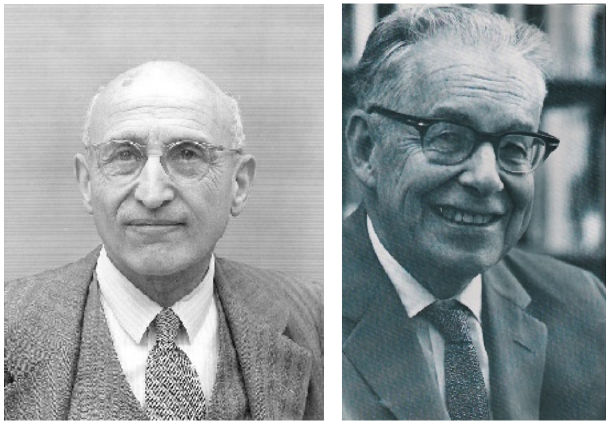
3. **Models for criticality**
   - Galton-Watson model
   - Extinction probability
   - Size distribution

4. **Self-organization towards criticality**
   - Self-organized branching process
   - Manna model, Bak-Tang-Wiesenfeld sandpile model
   - Spring-block earthquake models

5. **Fitting and goodness-of-fit testing of power-law distributions**

Only fools and charlatans predict earthquakes.
C. F. RICHTER

Gutenberg-Richter Law (1941)



seismo.berkeley.edu

- Most important law of statistical seismology and a paradigm of complex-systems geophysics

  For each earthquake with magnitude $m \geq 4$ there are about

  - 0.1 with $m \geq 5$
  - 0.01 with $m \geq 6$, etc...

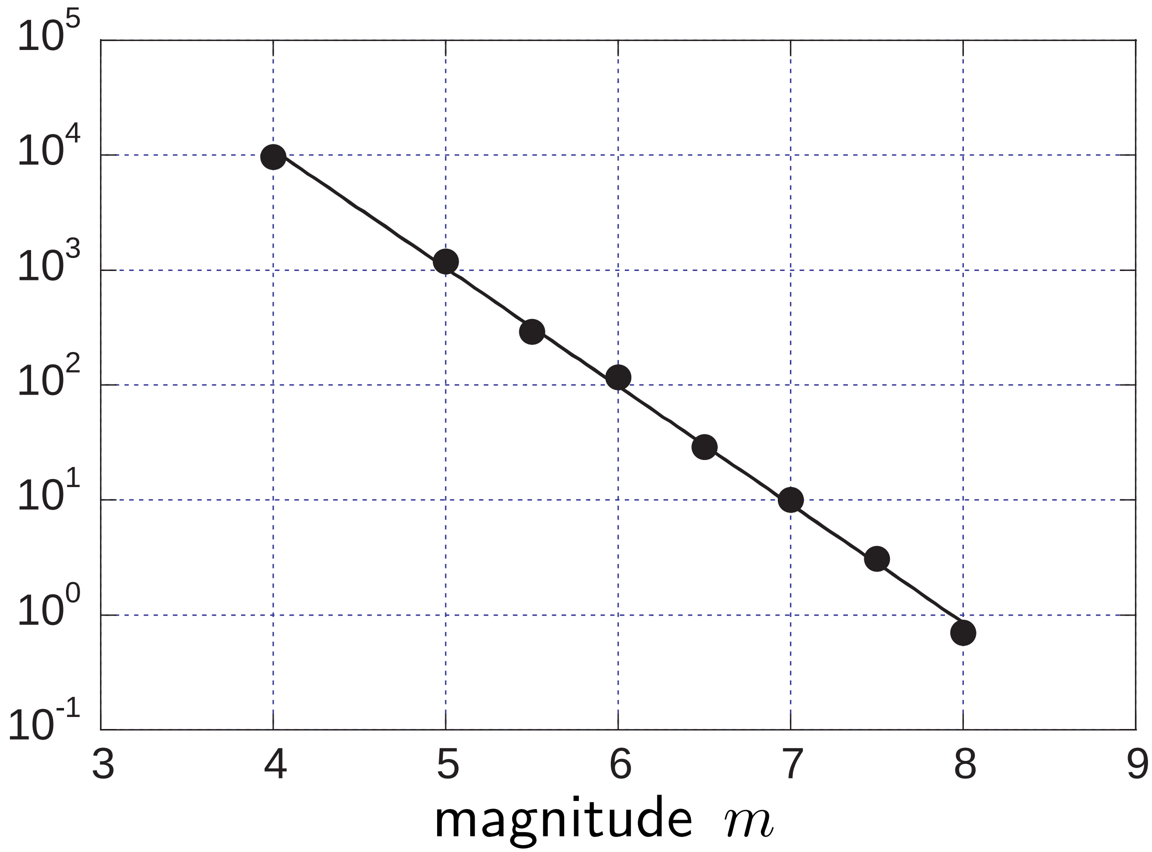
Gutenberg & Richter, BSSA 1944

Number of earthquakes with magnitude $\geq m$

$$N(m) \propto 10^{-bm}, \text{ with } b \simeq 1$$

Good news! $\Rightarrow$ Many small earthquakes, few big ones

- Example: worldwide earthquakes (one-year average)



Kanamori & Brodsky, *Rep Prog Phys* 2004

$$N(m) \propto 10^{-bm} \Rightarrow \log N(m) = \text{constant} - bm$$

Exponential Distribution of Earthquake Magnitudes

- Complementary cumulative distribution (survivor) function

$$S_m(m) = P\{\text{magnitude} \geq m\} \quad \Rightarrow \quad S_m(m) \propto N(m) \propto 10^{-bm}$$

- Probability density

$$D_m(m) = \frac{P\{m \leq \text{magnitude} < m + dm\}}{dm} = -\frac{dS_m(m)}{dm}$$

verifies $\int_0^\infty D_m(m)dm = 1$ and usually has units! $\Rightarrow$ It is not a probability

$$\text{Gutenberg-Richter law} \quad \Rightarrow \quad D_m(m) \propto 10^{-bm}$$

$S_m(m)$ and $D_m(m)$ are "the same" only for the exponential distribution

A statistician would stop here, wouldn't she?

Which is the Meaning of the Gutenberg-Richter Law?

- It depends (of course) on the meaning of magnitude...
  But magnitude is not a proper physical variable (it has no units)!

  Moreover: *magnitudes reflect radiation only from subportions of the rupture, and they saturate above certain size, rather than giving a physical characterization of the entire earthquake source* Ben-Zion, *Rev Geophys* 2008

- Radiated energy is supposed to be an exponential function of magnitude

$$E \propto 10^{3m/2}$$

  (with proportionality factor around 60 kJ)

  An increase of 1 unit in m leads to a factor $\sqrt{10^3} \simeq 32$ in E

$\Rightarrow$ An earthquake with $m = 9$ is “equivalent” to 1000 of $m = 7$

- Then, the Gutenberg-Richter law, in terms of $E \propto 10^{3m/2}$:

$$S_m(m) \propto 10^{-bm} \quad \Rightarrow \quad S_E(E) \propto \frac{1}{E^{2b/3}} \simeq \frac{1}{E^{0.7}}$$

Do you know how to perform the change of variables for the density?

$$D_m(m) \propto 10^{-bm} \quad \Rightarrow \quad D_E(E) \propto \frac{1}{E^{\beta}}$$

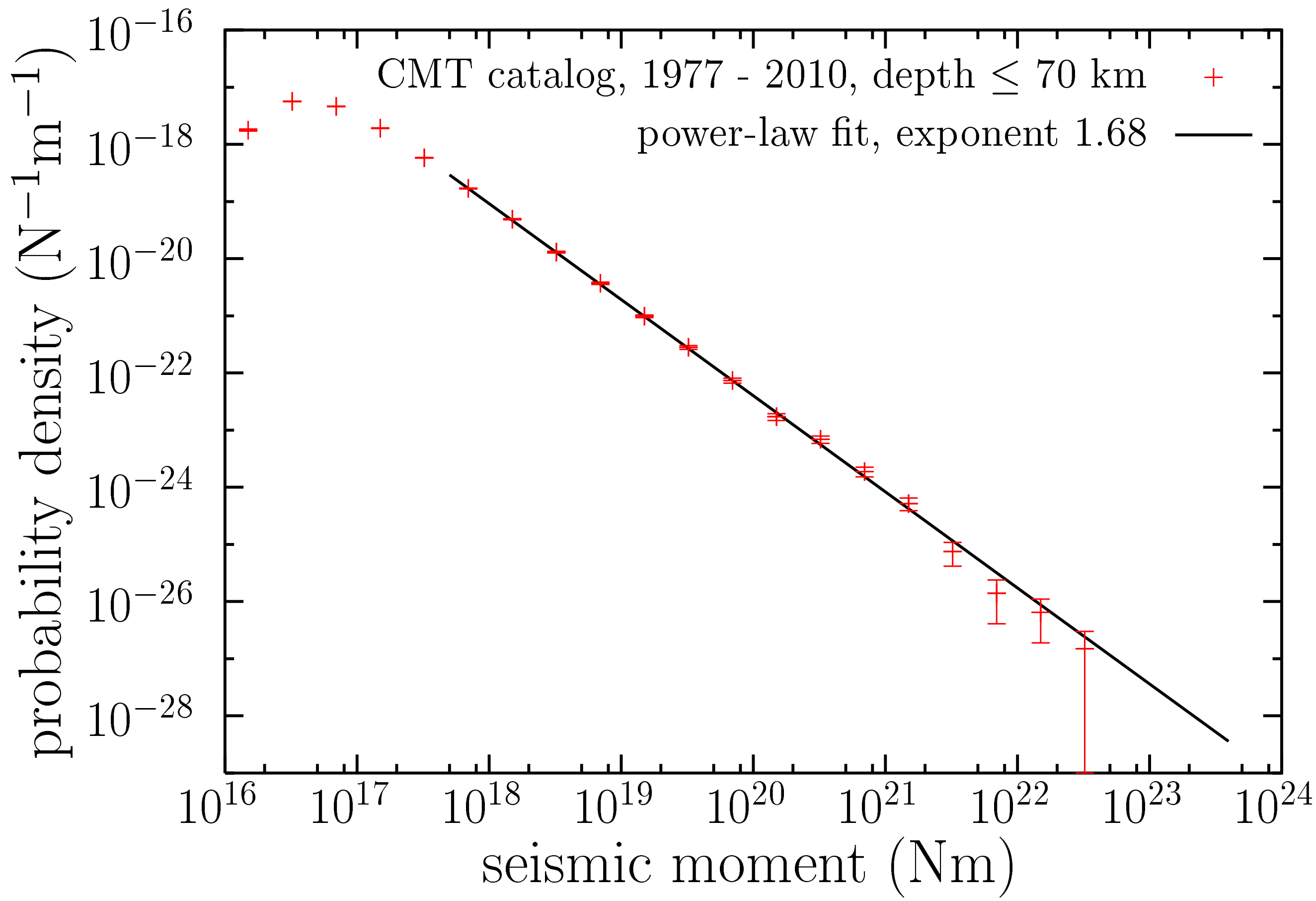
with

$$\beta = 1 + \frac{2b}{3} \simeq 1.7$$

$\Rightarrow$ Earthquake energy is power-law distributed — Wadati 1932, see Utsu, PAGEOPH 1999

$\Rightarrow$ Power-law fit cannot be rejected — Main et al. Nature Geosci. 2008

- Shallow worldwide earthquakes (seismic moment $\sim$ energy):



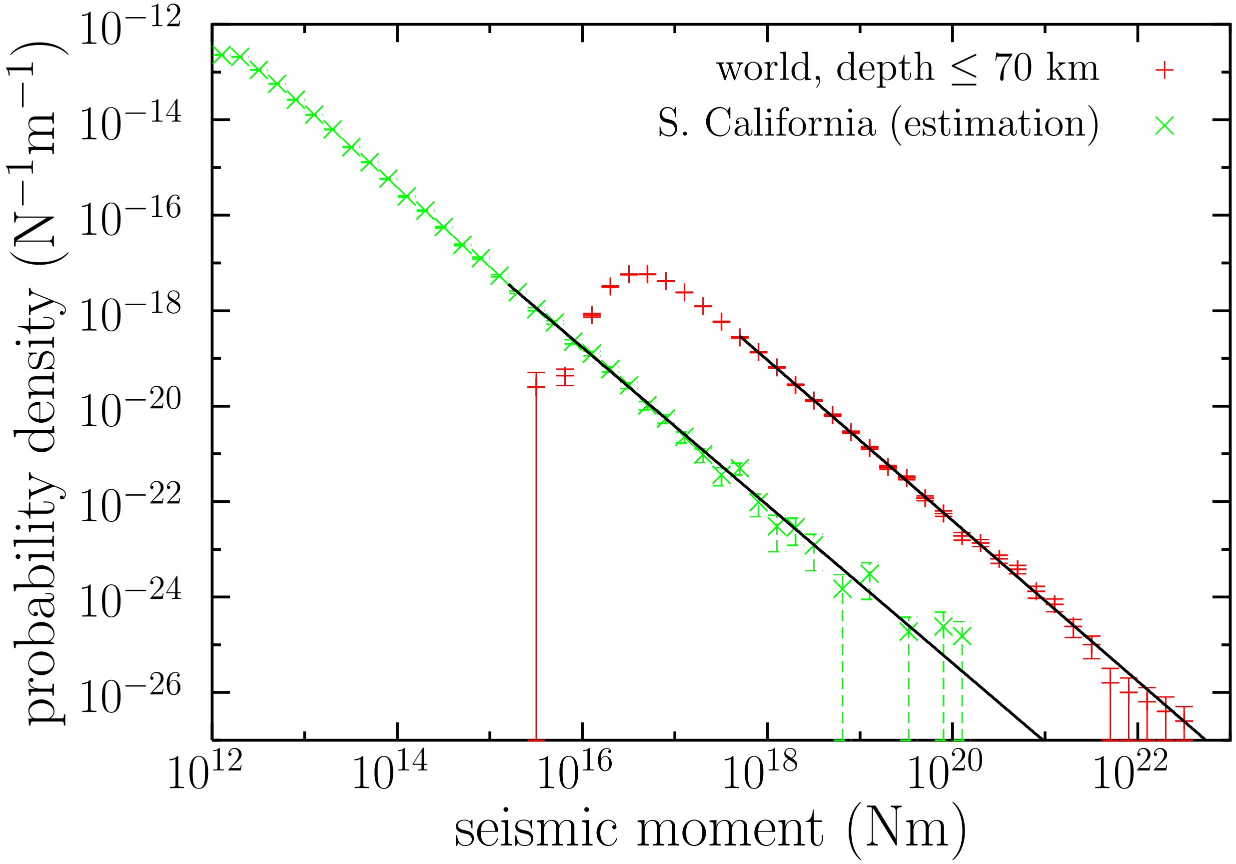
after Kagan

Geophys J Int 2002

Tectonophys 2010

$$D(E) \propto 1/E^{\beta} \quad \Rightarrow \quad \ln D(E) = \text{constant} - \beta \ln E$$

- Compare world with Southern California

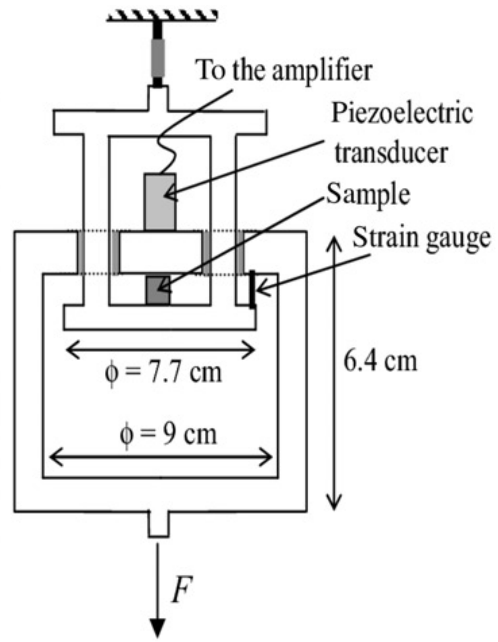


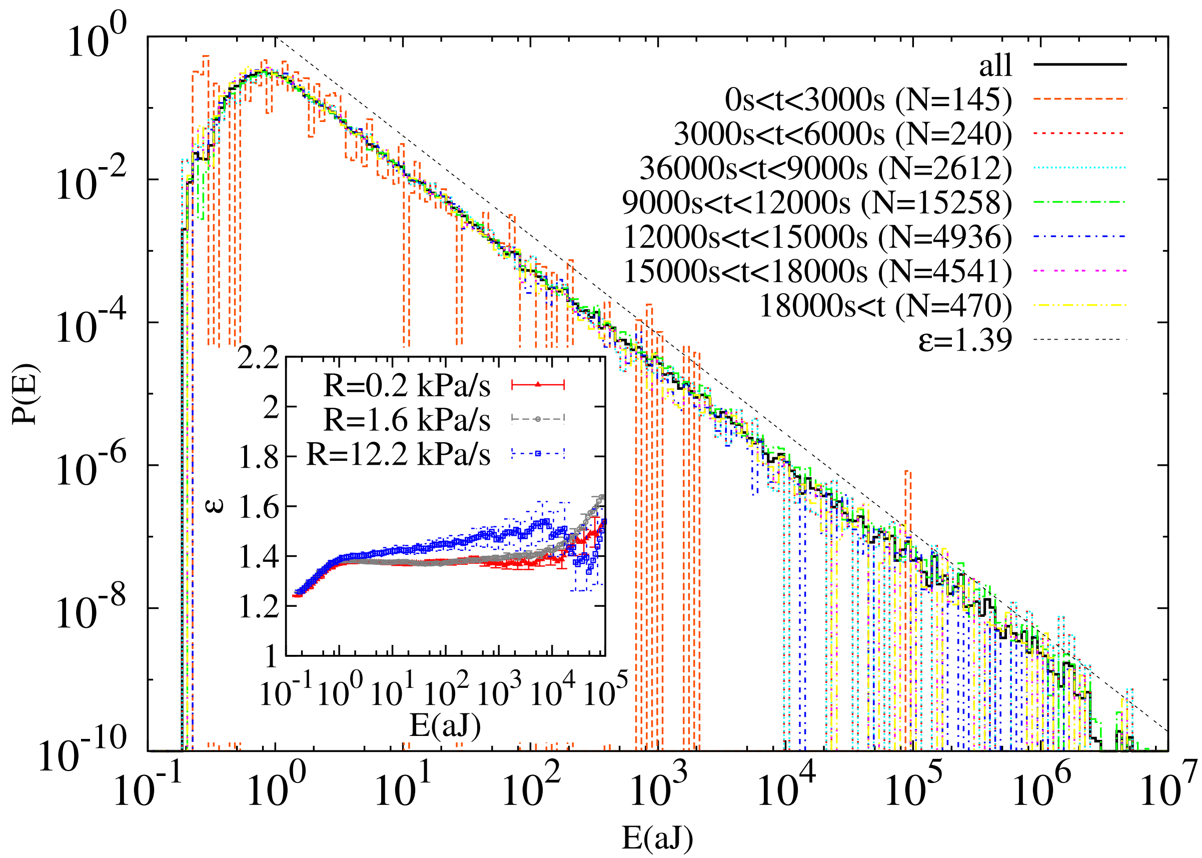
Valid up to $m \simeq -4$ in very small regions

Kwiatek et al., *Bull Seis Soc Am* 2010

- Even valid for fractures in the lab

Baró et al., *Phys Rev Lett* 2013





In nanofractures valid up to $m \simeq -13$

Åström et al. *Phys Lett A* 2006

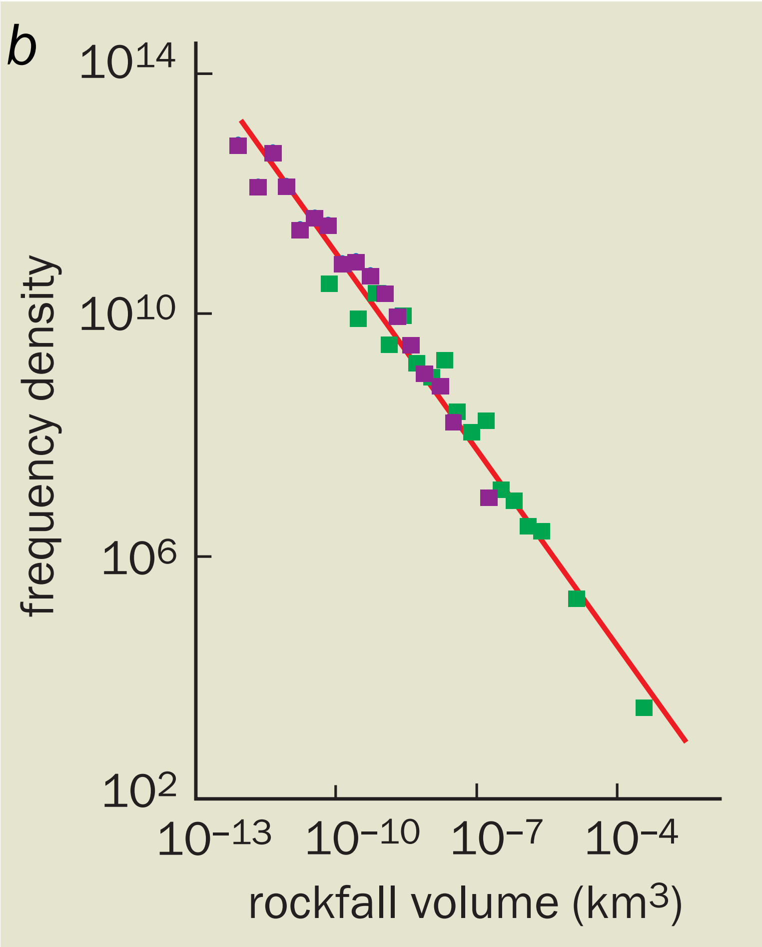
⇒ Enormous range of validity of the Gutenberg-Richter law

- *This law is amazing! How can the dynamics of all the elements of a system as complicated as the crust of the earth, with mountains, valleys, lakes, and geological structures of enormous diversity, conspire, as by magic, to produce a law with such extreme simplicity?*

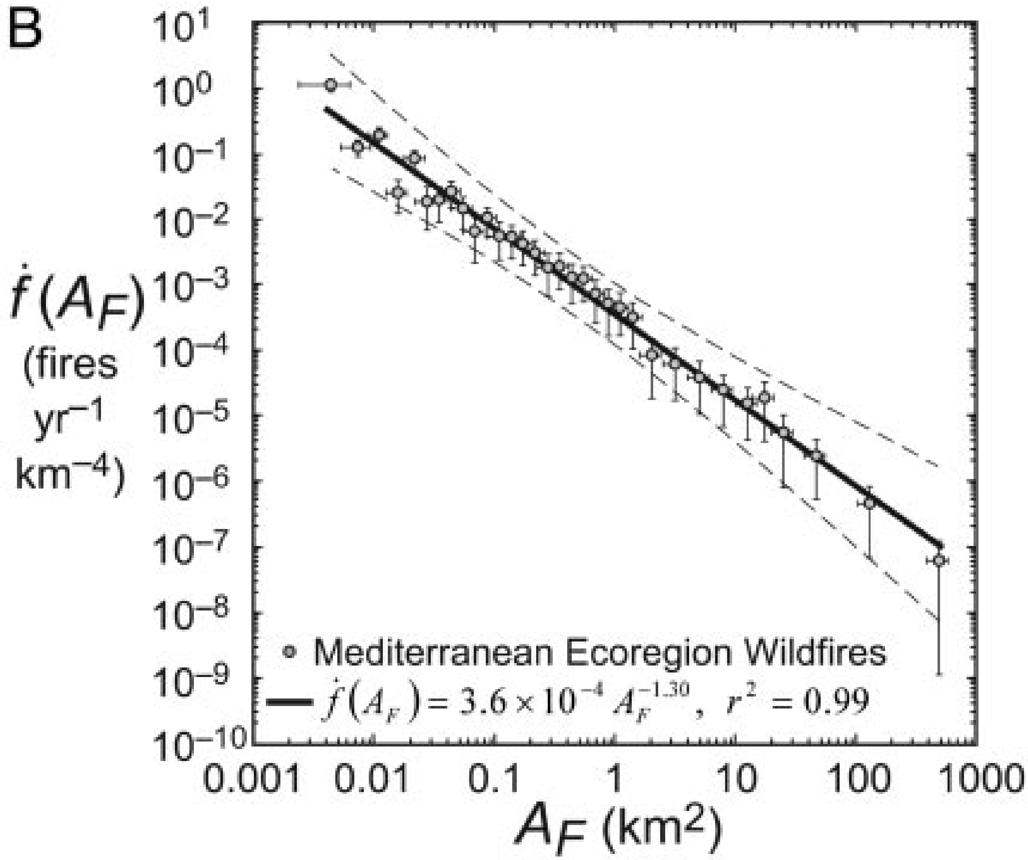
P. Bak, 1996

- Other examples of power-law distributions in natural hazards

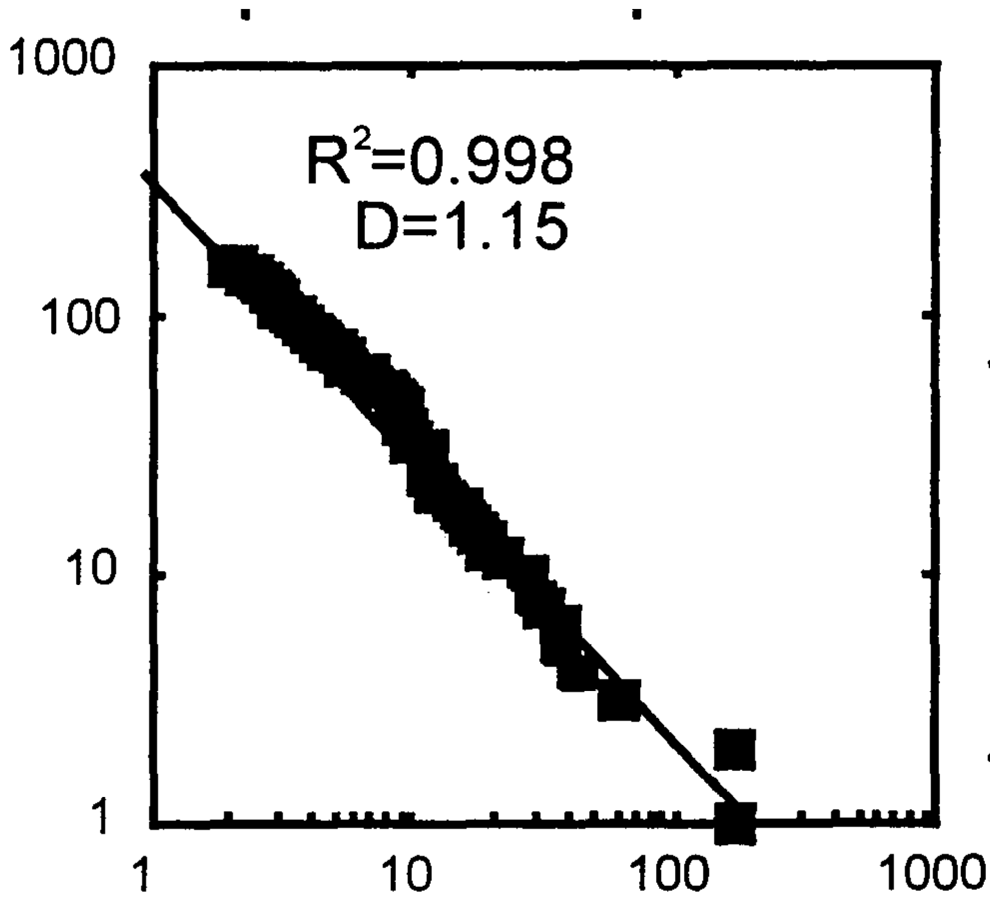
Rockfalls, Malamud, *Phys World* 2004



Forest fires, Malamud et al. *Science* 1998; *PNAS* 2005



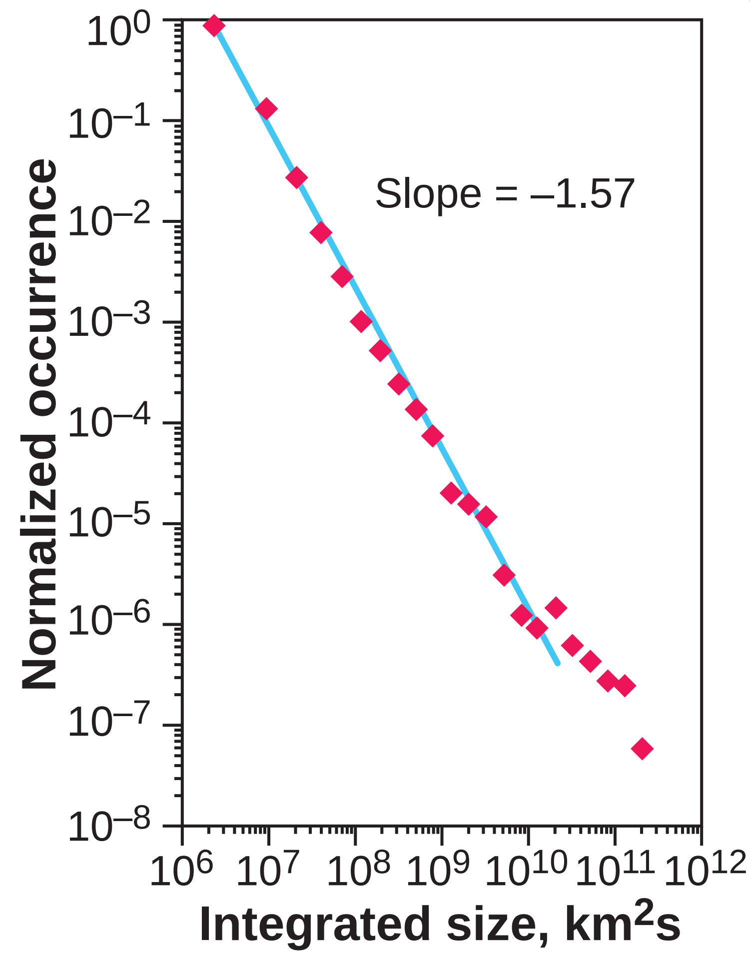
- Volcanic eruptions, Lahaie & Grasso, *J Geophys Res* 1998



Area covered by lava (10^6 m^2)

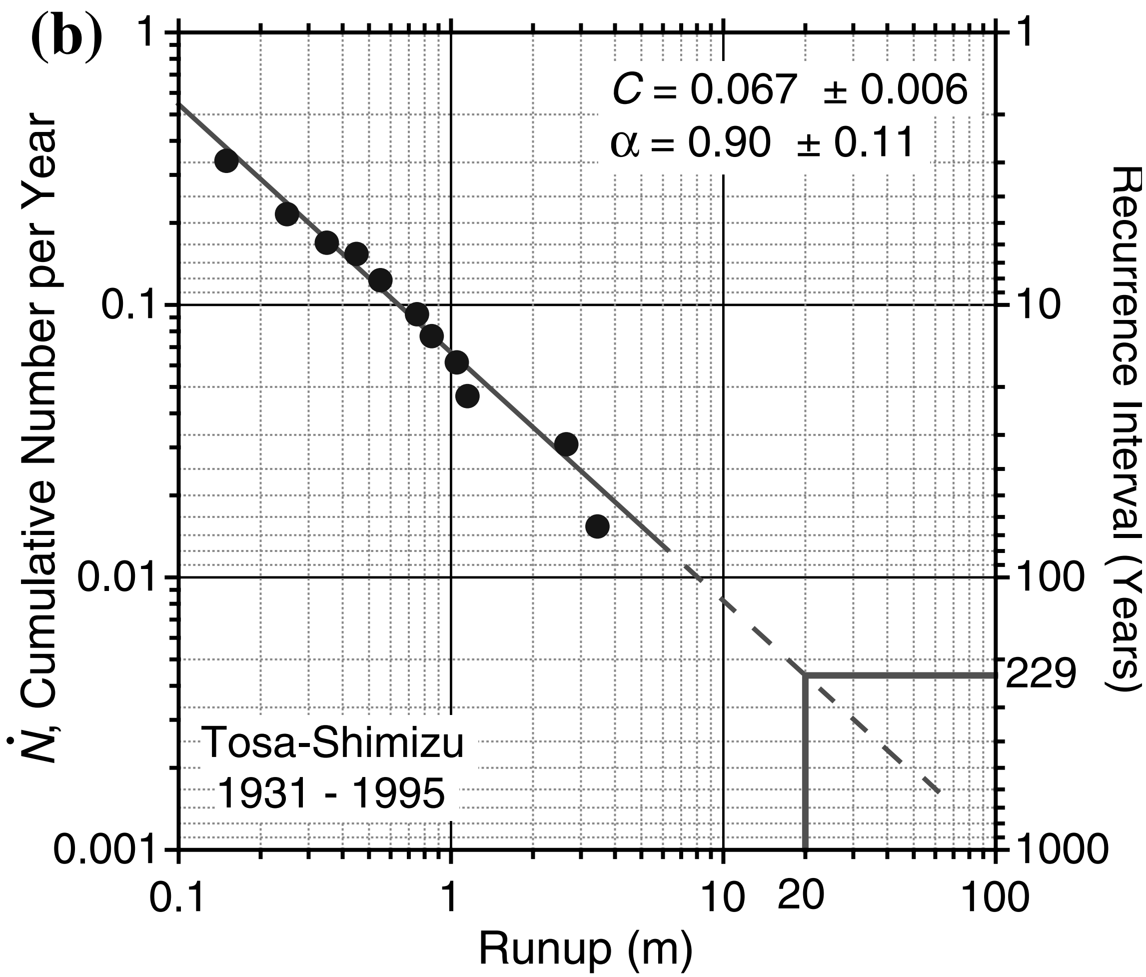
Auroras, Uritsky et al. *J Geophys Res* 2002

Freeman & Watkins, *Science* 2002

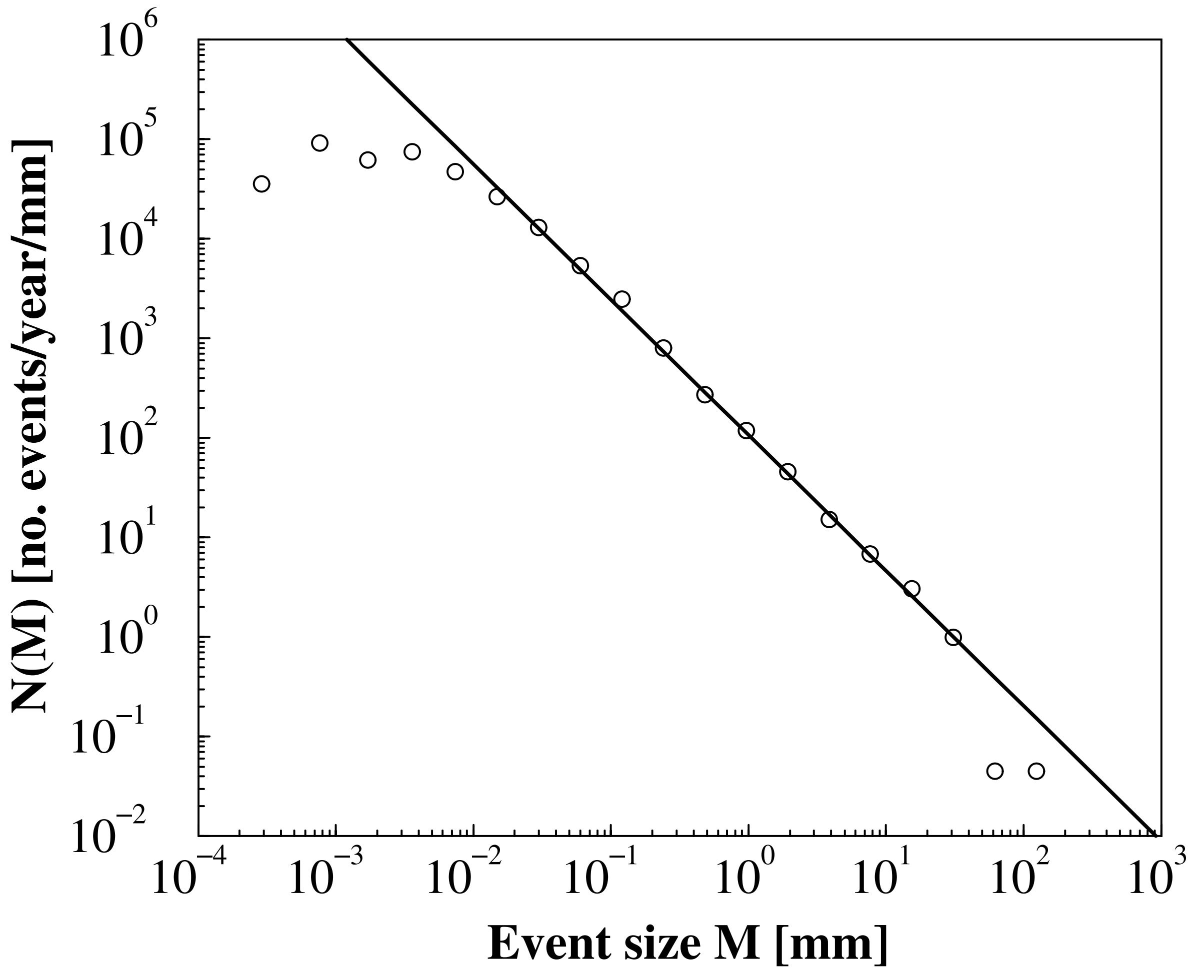


- Tsunamis,

Burroughs & Tebbens, *PAGEOPH* 2005



- Rainfall: flow of water in one point along duration of rain

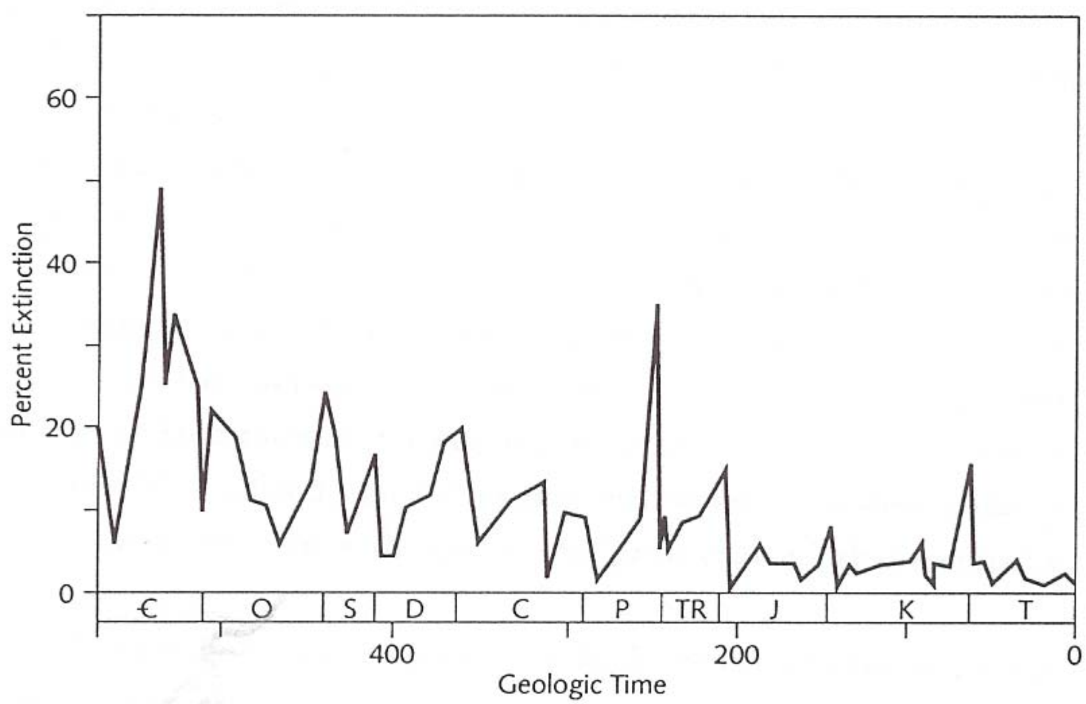


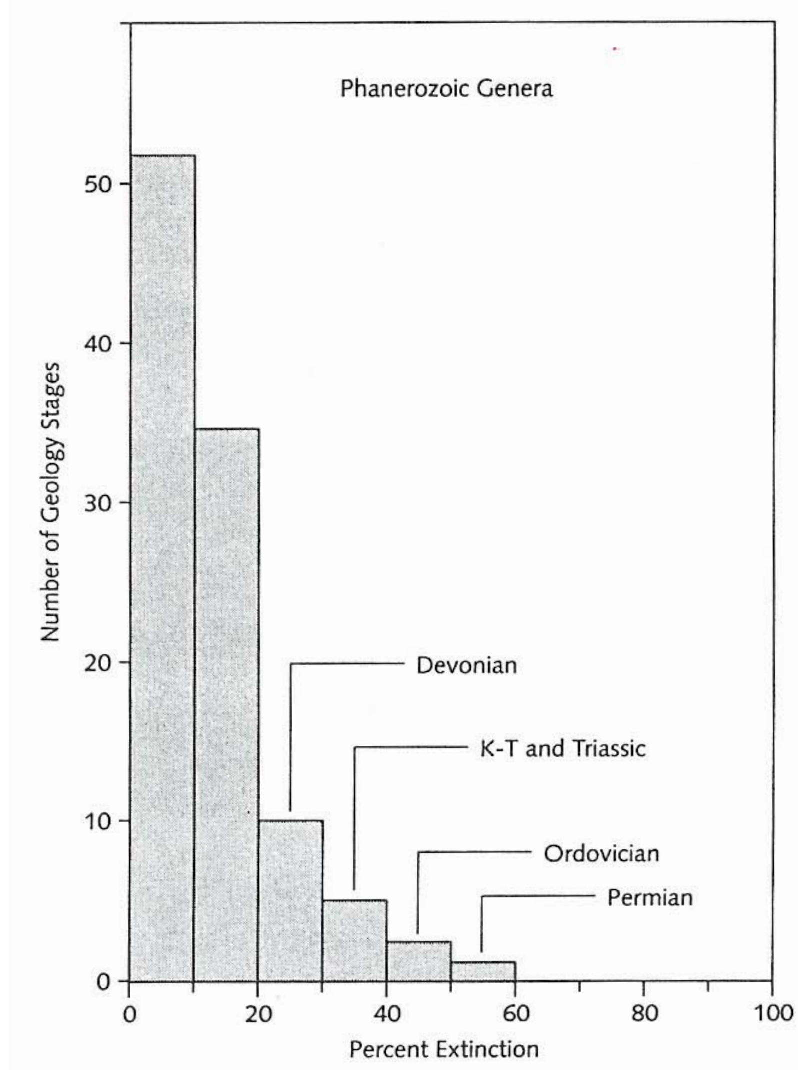
Peters et al. *Phys Rev Lett* 2002; *J Stat Mech* 2010

- Biological extinctions:

  Extinction measured as the percentage of extinct families in fixed periods of time (4 millions years)

  Sepkoski, Raup; after Bak 1996





Scaling laws never happen by accident

G. I. Barenblatt, 2003

Is there anything special about power-law distributions?

Scale transformation

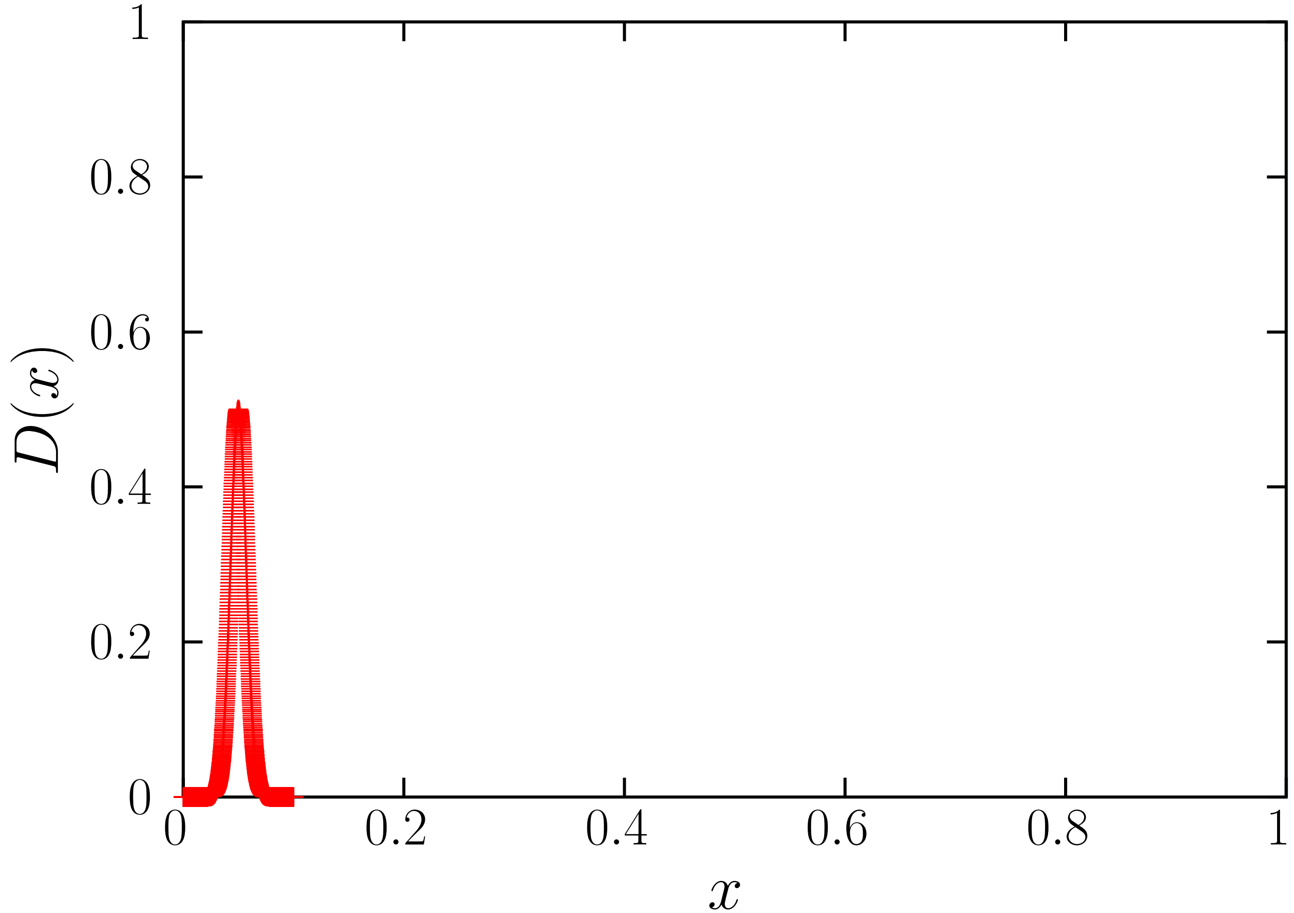
- Consider a function $D(x)$. Let us perform a linear transformation of the axes

$$\top[D(x)] = c_y D(x/c_x)$$

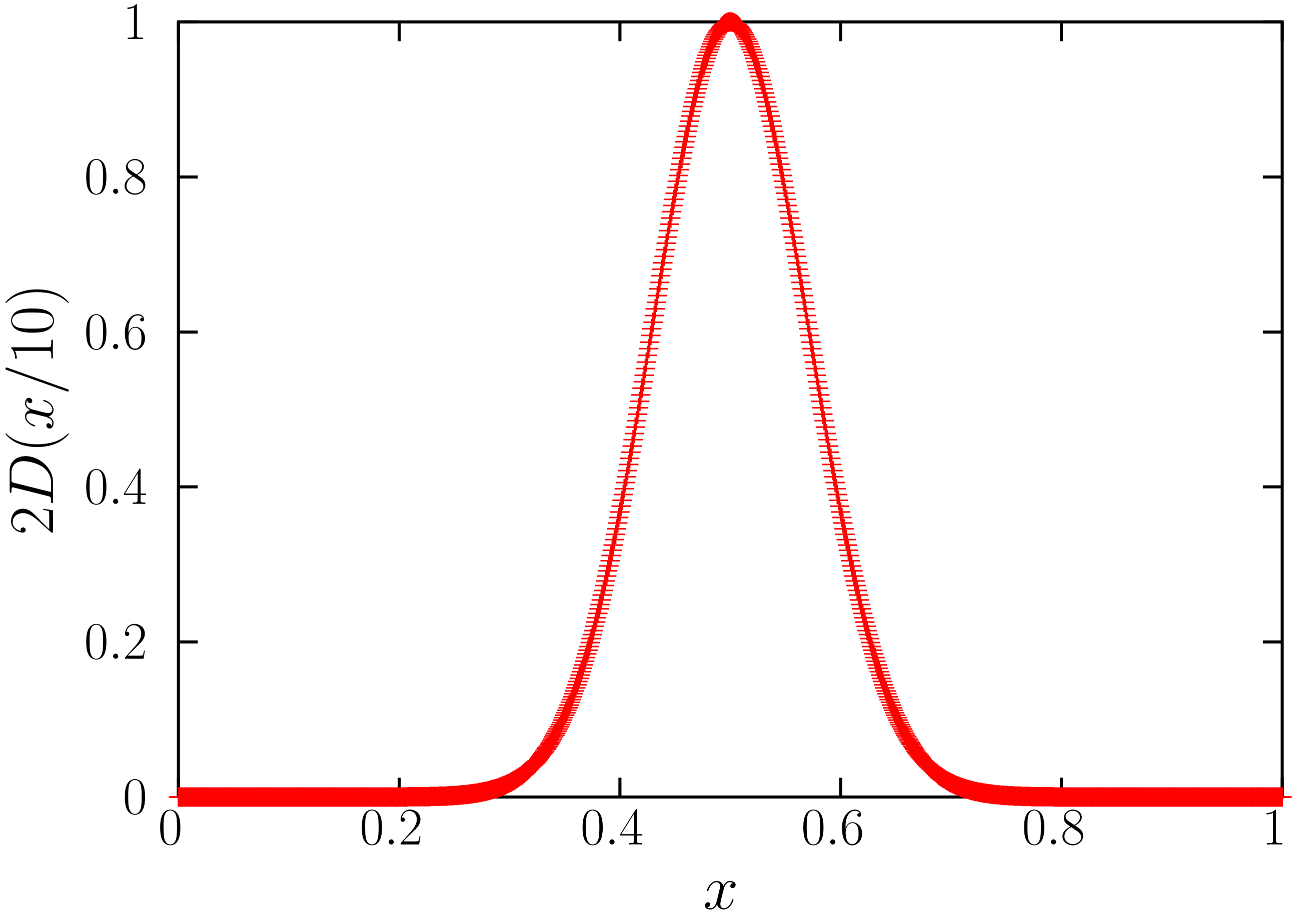
with $c_i > 0$, for $i = x, y$. If $c_i > 1$ then $\top$ acts as a mathematical microscope

- For example:
looking at $D(x)$ at the scale of m, $c_x = c_y = 100$ show $D(x)$ at the scale of cm

- Visual example: $\top[D(x)] = c_y D(x/c_x)$ with $c_x = 10$ and $c_y = 2$



- Visual example: $\top[D(x)] = c_y D(x/c_x)$ with $c_x = 10$ and $c_y = 2$



Scale invariance

- Mathematicians are allowed to ask themselves "silly" questions:

  Invariance under a scale transformation?

$$\top[D(x)] = c_y D(x/c_x) = D(x)$$

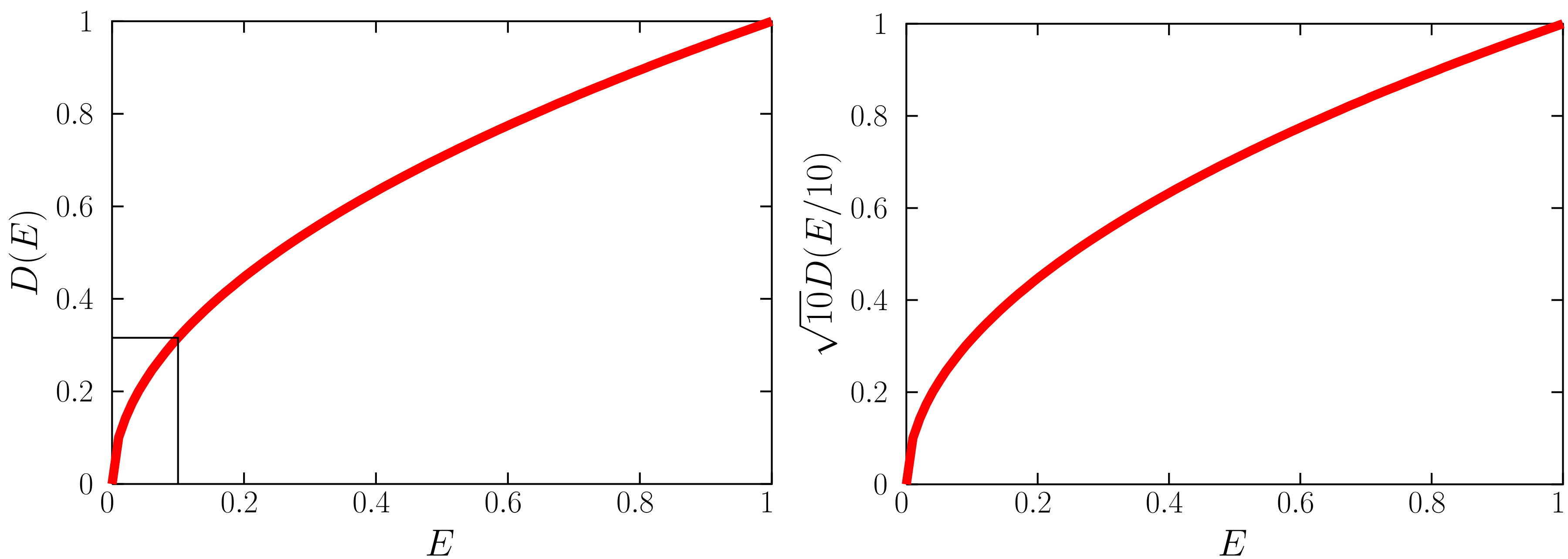
  Solution?

- The only solution of $D(x) = c_y D(x/c_x)$ for all c_x is the power law

$$D(x) \propto \frac{1}{x^\beta} \qquad \text{with } \beta = -\frac{\ln c_y}{\ln c_x} \qquad \text{i.e., } c_y = \frac{1}{c_x^\beta}$$

  Direct substitution confirms that it is a solution indeed

- Example: $D(x) \propto \sqrt{x}$ (i.e., $\beta = -1/2$). If $c_x = 10 \Rightarrow c_y = \sqrt{10}$



Difference between $\beta < 0$ (increasing power law) and $\beta > 0$ (decreasing)

⋆ if $\beta < 0$ and $c_x > 1$ then $c_y = 1/c_x^{\beta} > 1$
⋆ if $\beta > 0$ and $c_x > 1$ then $c_y = 1/c_x^{\beta} < 1$

- Demonstration

Takayasu, *Fractals* 1989; A.C. in *Carpinteri & Lacidogna* 2008

Differentiate both sides of $D(x) = c_y D(x/c_x)$ with respect x and isolate c_y

$$\frac{D'(x)}{D'(x/c_x)/c_x} = c_y = \frac{D(x)}{D(x/c_x)}$$

so, separating variables x and x/c_x and multiplying by x

$$\frac{x D'(x/c_x)}{c_x D(x/c_x)} = \frac{x D'(x)}{D(x)}$$

which has to be valid for all c_x, so, it only can be a constant ($+$, $-$, or 0),

$$\frac{x D'(x)}{D(x)} = \text{constant} = -\beta \qquad \Rightarrow \qquad D(x) \propto \frac{1}{x^\beta} \qquad \text{for } x > 0$$

Meaning of scale invariance?

- Power-law distributions do not have a characteristic scale

  One can define the time unit (or a clock) from the law of radioactive decay (which is an exponential, not a power law)

  But one cannot define a unit of distance from the law of gravitation (which is a power law)

  In the same way that one cannot built a compass from a sphere (which has rotational symmetry)

- So, earthquake energies have no characteristic scale

$\Rightarrow$ It is not possible to answer this simple question:

"How big are earthquakes in this region?"

Implications for extreme events

- We have already seen that the GR law for earthquakes implies that:

  large earthquakes do not play a special role,
  they follow the same law as small earthquakes

  $\Rightarrow$ *general theory encompassing all earthquakes, large and small* P. Bak, 1996

- But scale invariance goes beyond this fact:

  there is no unarbitrary way to separate ordinary events from extreme events

  (at least attending the statistics of event sizes)

Discrete scale invariance

- We can consider the constant β as a complex number, $\beta \to \beta - \omega i$

$$\Rightarrow \frac{1}{x^\beta} \to x^{-\beta+\omega i} = x^{-\beta} e^{i\omega \ln x} \qquad \text{and substitute in } c_y D(x/c_x) = D(x)$$

Then, if c_x is real, then $c_y = 1/c_x^{\beta - i\omega} = c_x^{-\beta} e^{i\omega \ln c_x}$ is complex (in general)

Imposing that c_y is positive real

$$c_x = \exp(2\pi n/\omega) \qquad \text{with } n = 0, \pm 1, \pm 2 \ \ldots$$

Thus, scale invariance does not hold for all c_x but for discrete values

In this case, the real part and the imaginary part are also scale invariant

$$\text{Re}[x^{-\beta+\omega i}] = \frac{1}{x^\beta}\cos(\omega \ln x) \qquad \text{or} \qquad \text{Im}[x^{-\beta+\omega i}] = \frac{1}{x^\beta}\sin(\omega \ln x)$$

Scale invariance for multivariate functions

Christensen & Moloney 2005

- Consider $D(x, y)$ and a scale transformation $\top[D(x,y)] = c_z D(x/c_x, y/c_y)$
  The scale-invariance condition $D(x,y) = c_z D(x/c_x, y/c_y)$ has a unique solution

$$D(x,y) = x^{-\beta} F(y/x^{\alpha}) \qquad \text{for all } c_x > 0$$

  which is called a scaling law, with

$$c_y = c_x^{\alpha} \qquad \text{and} \qquad c_z = \frac{1}{c_x^{\beta}}$$

  and the scaling function? $F()$ is arbitrary

- Equivalent expressions: $D(x,y) = x^{-\beta} F_2(x/y^{1/\alpha}) = y^{-\beta/\alpha} F_3(x/y^{1/\alpha})$, etc.
- We will distinguish scaling laws from power laws
  For univariate functions both are the same, with $F =$ constant

Mean earthquake energy...?

$$E[E] = \langle E \rangle = \int_{min}^{\infty} E D(E) dE \propto \int_{min}^{\infty} \frac{dE}{E^{\beta-1}} = \infty$$

... is infinite! (because $1 < \beta \leq 2$)

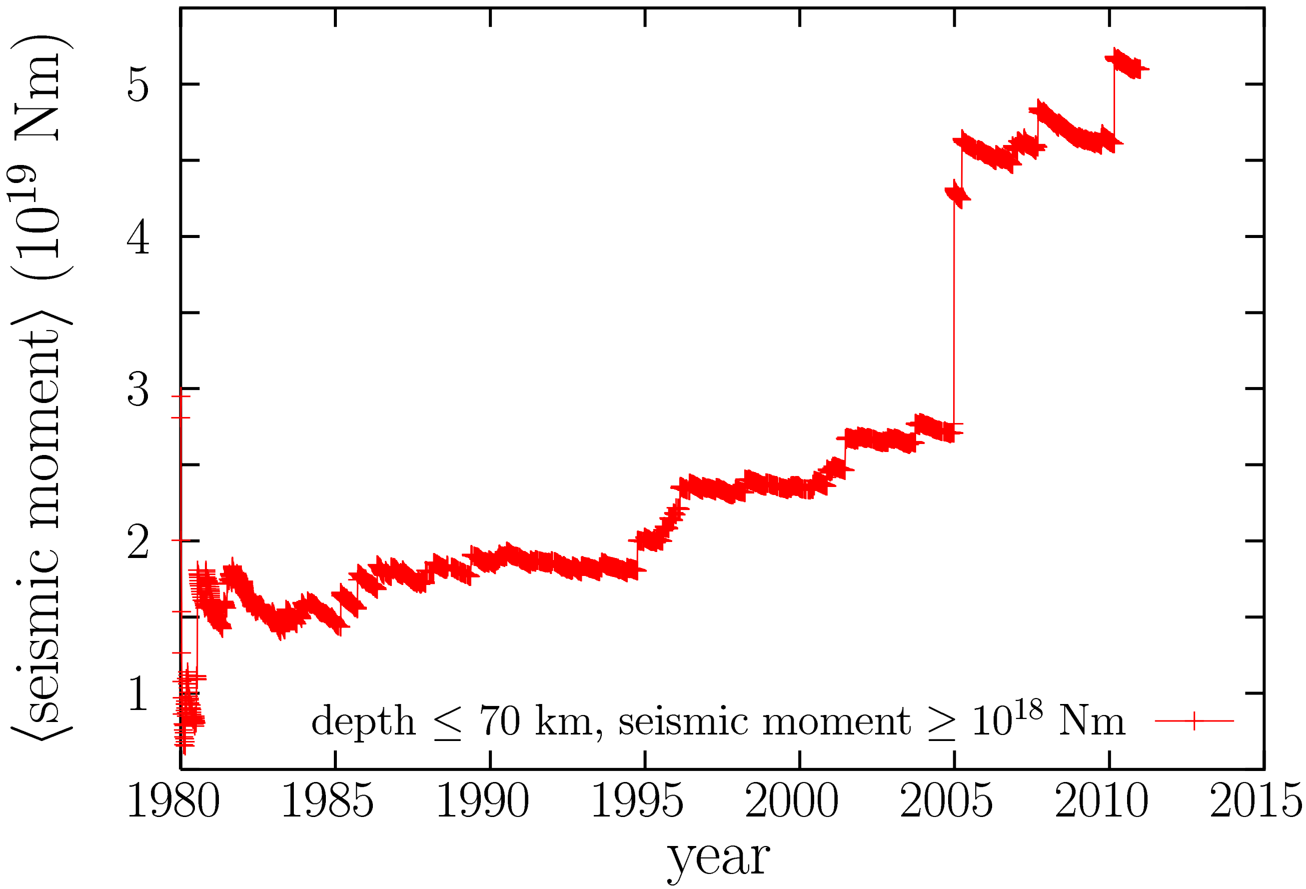
- Higher-order moment are also infinite.
- Which is the problem? Is mathematical?

  This process has a mean waiting time between events which is infinite:

  $$t_{i+1} = t_i + (1-u_i)^{1/(\beta-1)} \text{ with } u_i \text{ uniform random in } [0,1)$$

  Is physical then? The Earth contains a finite amount of energy!
- What does $\langle E \rangle = \infty$ mean in practice?

- Consider the average up to the $N-$th event, $\bar{E} = (E_1 + E_2 + \cdots + E_N)/N$



The rare big events are crucial for energy dissipation $\Rightarrow$ Bad news!!!

Discrete analog: the St. Petersburg paradox

N. Bernoulli 1713 & D. Bernoulli 1738

- Consider a game of chance in which a player tosses a (fair) coin until a tail appears for the 1st time. Each toss doubles the payoff

Outcome	Probability	Payoff
tail	$p_1 = 1/2$	1 \$
heads,tail	$p_2 = 1/4$	2 \$
heads,heads,tail	$p_3 = 1/8$	4 \$
...		
heads ... heads,tail	$p_k = 1/2^k$	2^{k-1} \$ for k tosses, in general

- You are a casino: which would be the fair price to pay to enter the game?

$$\langle \text{payoff} \rangle = \sum_{k=1}^{\infty} p_k \times \text{payoff}(k) = \sum_{k=1}^{\infty} \frac{1}{2^k} \times 2^{k-1}\$ = \frac{1}{2} \times 1\$ + \frac{1}{4} \times 2\$ + \ldots = \infty$$

- Note that the duration k of the game is geometrically (exponentially) distributed

$$p_k = \frac{1}{2^k} = e^{-k \ln 2} = 10^{-k \log 2} \qquad \Rightarrow \qquad \langle K \rangle = \sum_{k=1}^{\infty} p_k k = \frac{1}{1/2} = 2$$

so, the duration of the game is analogous to magnitude, with $b = \log 2 \neq 1$

- But the payoff $= 2^{k-1} \propto 10^{k \log 2} = 10^{ck}$ is analogous to energy, with $c = \log 2$
- Then, the payoff follows a (sort of) discrete power-law distribution with

$$\beta = 1 + \frac{b}{c} = 1 + \frac{\log 2}{\log 2} = 2$$

This is in the "boundary" of having a finite mean

Laplace transform

- Consider $D_x(x)$ defined for $x \geq 0$, then

$$\tilde{D}_x(z) = \int_0^\infty e^{-zx} D_x(x) dx = \langle e^{-zX} \rangle$$

if $D_x(x)$ is a probability density, normalization implies $\tilde{D}_x(z=0) = 1$

- Assuming that $\tilde{D}_x(z)$ exists and that all moments $\langle X^n \rangle$ are finite, and using $e^{-zx} = \sum_{n=0}^{\infty} (-1)^n z^n x^n / n!$

$$\tilde{D}_x(z) = 1 - \langle X \rangle z + \frac{1}{2} \langle X^2 \rangle z^2 - \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{\langle X^n \rangle z^n}{n!}$$

so, the Laplace transform of $D_x(x)$ is a sort of moment generating function

Sum and rescaling of independent random variables

skip!

- Define $S = X + Y$, then $F_S(s) = P\{\text{sum} < s\} = P\{Y < s - X\} \Rightarrow$

$$F_S(s) = \int_0^s dx \int_0^{s-x} dy D_X(x) D_Y(y) = \int_0^s dx D_X(x) F_Y(s-x)$$

Differentiating with the Leibniz rule, $D_S(s) = dF_S(s)/ds =$

$$= \int_0^s dx D_X(x) \frac{dF_Y(s-x)}{ds} + D_X(x) F_Y(s-x)\Big|_{x=s} = \int_0^s dx D_X(x) D_Y(s-x)$$

Calculating the Laplace transform, with $\theta(x)$ the step function,

$$\tilde{D}_S(z) = \int_0^\infty ds\, e^{-zs} D_S(s) = \int_0^\infty ds\, e^{-zs} \int_{-\infty}^{\infty} dx D_X(x)\theta(x) D_Y(s-x)\theta(s-x)$$

$$\Rightarrow \tilde{D}_S(z) = \int_0^\infty dy\, e^{-zy} D_Y(y) \int_0^\infty dx\, e^{-zx} D_X(x) = \tilde{D}_X(z)\tilde{D}_Y(z) \quad \text{using } s - x = y$$

The sum is a convolution of D_X and D_Y, which turns a product of $\tilde{D}_X$ and $\tilde{D}_Y$

Sum and rescaling of independent random variables

- Define $S = X + Y$, then, the Laplace transform of the distribution of S

$$\tilde{D}_S(z) = \langle e^{-zS} \rangle = \int_0^\infty ds\, e^{-zs} D_S(s) = \int_0^\infty \int_0^\infty dx dy D_X(x) D_Y(y)\, e^{-z(x+y)}$$

where we have used independence $[D_{X,Y}(x,y) = D_X(x) D_Y(y)]$, then

$$\tilde{D}_S(z) = \int_0^\infty dx D_X(x) e^{-zx} \int_0^\infty dy D_Y(y) e^{-zy} = \tilde{D}_X(x) \tilde{D}_Y(y)$$

So, the Laplace transform of the sum is the product of $\tilde{D}_X(z)$ and $\tilde{D}_Y(z)$
It is not necessary to know that $D_S(s)$ is the convolution of $D_X(x)$ and $D_Y(y)$

- In general, if $S = X_1 + X_2 + \cdots + X_N$, then

$$\tilde{D}_s(z) = [\tilde{D}_x(z)]^N$$

when all X_i are independent and identically distributed

- Rescaling by a constant, $R = S/C$

$$\tilde{D}_r(z) = \int_0^\infty dr D_r(r) e^{-zr} = \int_0^\infty ds D_s(s) e^{-z(s/C)} = \tilde{D}_s(z/C)$$

- Defining the rescaled mean, or "non-conserved" average

$$R = \frac{X_1 + X_2 + \cdots + X_N}{N^{1/\alpha}} \qquad \Rightarrow \qquad \tilde{D}_r(z) = [\tilde{D}_x(z/N^{1/\alpha})]^N$$

- Introducing a cumulant generating function

$$G_x(z) = \ln \tilde{D}_x(z) \qquad \Rightarrow \qquad G_r(z) = N G_x(z/N^{1/\alpha})$$

Sum of X's turns into product of m.g.f. and into a sum of cumulant g.f.
(if independence holds)

Note: g.f. $=$ generating function, m.g.f. $=$ moment g.f.

- If the moments are finite (and the generating function exists)

$$\tilde{D}_x(z) = 1 - \langle X \rangle z + \frac{1}{2}\langle X^2 \rangle z^2 - \ldots$$

Considering $\ln(1-y) = -y - y^2/2 - y^3/3 - \ldots$, with $-1 \leq y < 1$, then, the cumulant generating function

$$G_x(z) = \ln \tilde{D}_x(z) = \ln \left[1 - \left(\langle X \rangle z - \frac{1}{2}\langle X^2 \rangle z^2 + \ldots\right)\right] =$$

$$-\left(\langle X \rangle z - \frac{1}{2}\langle X^2 \rangle z^2 + \ldots\right) - \frac{1}{2}\left(\langle X \rangle z - \ldots\right)^2 + \ldots = -\langle X \rangle z + \frac{\langle X^2 \rangle - \langle X \rangle^2}{2} z^2 - \ldots$$

From the coefficients we can obtain the cumulants: $\langle X \rangle$, σ^2, etc.

Distributions stable under "averaging"

- Again a "silly" question: let us look at the fixed points of this transformation

$$G^*(z) = N G^*(z/N^{1/\alpha})$$

This is the scale invariance condition, whose only solution for all N is

$$G^*(z) \propto z^{\alpha}$$

- In the case of the arithmetic mean, $\alpha = 1$, then $D_x^*(x) = \delta(x - \mu)$, indeed

$$\tilde{D}_x^*(z) = \int_0^\infty dx\, e^{-zx} \delta(x - \mu) = e^{-\mu z} \qquad \Rightarrow \qquad G_x^*(z) = \ln D_x^*(z) = -\mu z$$

where $\delta(x - \mu)$ is a Dirac delta "function", which has mean μ and zero variance

Domain of attraction of the Dirac delta distribution

- Considering the expansion of $G_x(z)$ into cumulants (if they exist and are finite)

$$G_x(z) = -\langle X\rangle z + \frac{\sigma^2}{2} z^2 - \ldots$$

Applying the scale transformation we get the distribution of the mean

$$G_{\bar{x}}(z) = G_r(z) = N G_x(z/N) = N\left[-\langle X\rangle \frac{z}{N} + \frac{\sigma^2}{2}\left(\frac{z}{N}\right)^2 - \ldots\right] \to -\langle X\rangle z$$

The distribution of the mean tends to a delta centered at $\langle X\rangle$ when $N \to \infty$
So, the fixed point is attractive if $G_x(z)$ exists and all moments are finite
We will see that the domain of attraction is even bigger

- This constitutes a version of the law of large numbers (weak version) Feller 1971

  It is somehow analogous to the central limit theorem also
  Note that the Gaussian (normal) distribution also tends to a delta
  (because we do not have zero mean)
  If we had subtracted the mean the "central limit" would have been Gaussian

Stability and domain of attraction for "non-conserved" averaging

- Coming back to the general rescaled mean, $G^*(z) \propto z^\alpha$, consider $\alpha = 1/2$

$$G^*_x(z) = -2a\sqrt{z} \qquad \Rightarrow \qquad D^*_x(x) = e^{-a^2/x} \frac{a}{\sqrt{\pi}\, x^{3/2}}$$

Abramowitz & Stegun, 29.3.82; Bouchaud & Georges, *Phys Rep* 1990

As $(x_1 + x_2 + \cdots + x_N)/N^2$ converges, the mean diverges linearly with N

- Do it yourself! Simulate N random values of X. How?
  - Consider the transformation $X = 1/Y^2$
  - Y follows a half-normal (half-Gaussian) distribution
  - Use standard algorithm (like Box-Muller transformation) to simulate Y
- Alternative: simulate a power-law with exponent $3/2$ $\Rightarrow$ What happens?

- Power-law distributions belong to the domain of attraction of $G^*(z) \propto z^\alpha$
  Consider $D_x(x) = B/x^{1+\rho}$ for $x \geq c > 0$ (and 0 otherwise), then $B = \rho c^\rho$ and

$$\tilde{D}_x(z) = B \int_c^\infty e^{-zx} x^{-\rho-1} dx = B z^\rho \Gamma(-\rho, cz)$$

with $\Gamma(\gamma, z) = \int_z^\infty u^{\gamma-1} e^{-u} du$ the incomplete gamma function, with expansion

$$\Gamma(\gamma, z) = \Gamma(\gamma) - z^\gamma \sum_{n=0}^\infty \frac{(-z)^n}{(\gamma+n)n!} \qquad \gamma \neq 0, -1, -2, -3 \ldots \qquad \text{Abramowitz \& Stegun 6.5.29}$$

with $\Gamma(\gamma) = \Gamma(\gamma, 0)$ for $\gamma > 0$ and $\Gamma(\gamma) = \Gamma(\gamma+1)/\gamma$ for $\gamma < 0$ (non-integer)

$$\Rightarrow z^\rho \Gamma(-\rho, z) = z^\rho \Gamma(-\rho) - \sum_{n=0}^\infty \frac{(-z)^n}{(n-\rho)n!} = \frac{1}{\rho}\left[\rho\Gamma(-\rho)z^\rho + \left(1 + \frac{\rho z}{1-\rho} + \ldots\right)\right]$$

$\rho \neq 0, 1, 2, \ldots$ We are interested in $G_x(z) = \ln \tilde{D}_x(z) = \ln B z^\rho \Gamma(-\rho, cz)$, so

$$\ln z^\rho \Gamma(-\rho, z) = -\ln \rho + \ln [\,] = -\ln \rho + \rho\Gamma(-\rho)z^\rho + \frac{\rho z}{1-\rho} + \ldots$$

$$\Rightarrow G_x(z) = \ln \frac{B}{c^\rho} + \ln c^\rho z^\rho \Gamma(-\rho, cz) = \rho\Gamma(-\rho) c^\rho z^\rho + \frac{\rho c z}{1-\rho} + \ldots$$

using again the expansion of the logarithm. Applying the transformation

$$G_r(z) = N G_x(z/N^{1/\alpha}) = N \left[\rho\Gamma(-\rho) c^\rho \left(\frac{z}{N^{1/\alpha}} \right)^\rho + \frac{\rho c}{1-\rho} \left(\frac{z}{N^{1/\alpha}} \right) + \ldots \right]$$

If $0 < \rho < 1$ and $\alpha = \rho$ then $G_r(z) \to \rho\Gamma(-\rho) c^\rho z^\rho$

If $\rho > 1$ and $\alpha = 1$ then $G_r(z) = G_{\bar{x}}(z) \to -\frac{\rho c}{\rho - 1} z$

where the coefficient is the mean of the power-law distribution in this case
(in any other case $G_r(z) \to 0$ or ∞)

- The domain of attraction includes distributions that are asymptotically power laws

$$D_x(x) \sim \frac{B}{x^{1+\rho}} \qquad \text{for } x \to \infty$$

with $B \neq \rho c^{\rho}$. If ρ is not a positive integer

$$\tilde{D}_x(z) \sim B\Gamma(-\rho)z^{\rho} + \sum_{n=0}^{\infty} \frac{a_n(-z)^n}{n!}$$

Bleistein & Handelsman 4.6.23

which is, except for the multiplying constants, the same as before (with $a_0 = 1$)
So, again, there are 2 cases:

If $0 < \rho < 1$	and $\alpha = \rho$	then	$G_r(z) \to B\Gamma(-\rho)z^{\rho}$
If $\rho > 1$	and $\alpha = 1$	then	$G_r(z) = G_{\bar{x}}(z) \to a_1 z$

- Reciprocally, $G_x(z) \propto z^{\rho}$ corresponds to a distribution that is asymptotically power law if $0 < \rho < 1$

Summary

- Assuming independence:

  If the moments of $D_x(x)$ are finite and its generating function exist or

  If $D_x(x)$ is asymptotically a power law with exponent $\beta = 1 + \rho > 2$

  $\Rightarrow$ the arithmetic mean $\bar{x}$ follows a Dirac's delta distribution

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \quad \sim \quad \delta(\bar{x} - \langle x \rangle) \quad \text{when } N \to \infty$$

  $\Rightarrow$ Law of large numbers — Feller 1971

  $\Rightarrow$ Makes sense of the arithmetic mean!

- Assuming independence:
  If $D_x(x)$ is asymptotically a power law with exponent $\beta = 1 + \rho < 2$ (but > 1)

$$\frac{\bar{x}}{N^{1/\rho - 1}} = \frac{1}{N^{1/\rho}} \sum_{i=1}^{N} x_i \quad \text{follows a power-law tailed distribution}$$

  with exponent $1 + \rho$, when $N \to \infty$

⇒ The arithmetic mean diverges as $N^{1/\rho - 1}$ (as $1/\rho - 1 > 0$)

⇒ Case of the generalized central limit theorem   Bouchaud & Georges, *Phys Rep* 1990

- Therefore, the sum of earthquake energies "converges", if rescaled by

$$N^{1/\rho} = N^{1/(\beta - 1)} = N^{3/(2b)}$$

  Standard (conserved) averaging leads to divergence of $\bar{E}$ as $N^{3/(2b) - 1} \simeq \sqrt{N}$

- Concrete example, case $\beta = 1 + \rho = 3/2$

$$D_x(x) \sim \frac{A}{x^{3/2}} \quad \text{for "large" } x$$

The "correct" average to get convergence is

$$\frac{\bar{x}}{N^{1/\rho - 1}} = \frac{1}{N^{1/\rho}} \sum_{i=1}^{N} x_i = \frac{1}{N^2} \sum_{i=1}^{N} x_i = y$$

Then, for $N \to \infty$, the "non-conserved" average y follows

$$D_y(y) = e^{-a^2/y} \frac{a}{\sqrt{\pi}\, y^{3/2}}$$

with $a = -A\Gamma(-1/2)/2 = \sqrt{\pi} A$

What Should One Expect from a Theory of a Complex Phenomenon?

P. Bak, *How Nature Works*, 1996

- A theory must be abstract
  - ⋆ A theory of life does not need to predict elephants (if your theory predicts elephants it is not general enough)
  - ⋆ *Explain why there is variability, or what typical patterns may emerge*
  - ⋆ *If ... we concentrate on an accurate description of the details, we lose perspective*

- A theory must be statistical
  - ⋆ *Collecting anecdotal evidence can only be an intermediate goal.*
  - ⋆ *Anecdotal evidence carries weight only if enough of it can be gathered to form a statistical statement.*
  - ⋆ *Confrontation between theories and ... observations... takes place by comparing the statistical features of general patterns.*

- *The abstractness and the statistical, probabilistic nature of any such theory might appear revolting to geophysicists, biologists, and economists, expecting to aim for photographic characterization of real phenomena.*

- *Perhaps too much emphasis has been put on detailed prediction ...
in today's materialistic world.*

- *To predict the statistics of actual phenomena rather than the specific outcome is a quite legitimate and ordinary way of confronting theory with observations.*

I cannot imagine a theory of earthquakes that does not explain the GR law

Per Bak, 1996

Domino-like theory: Otsuka's model (1971)

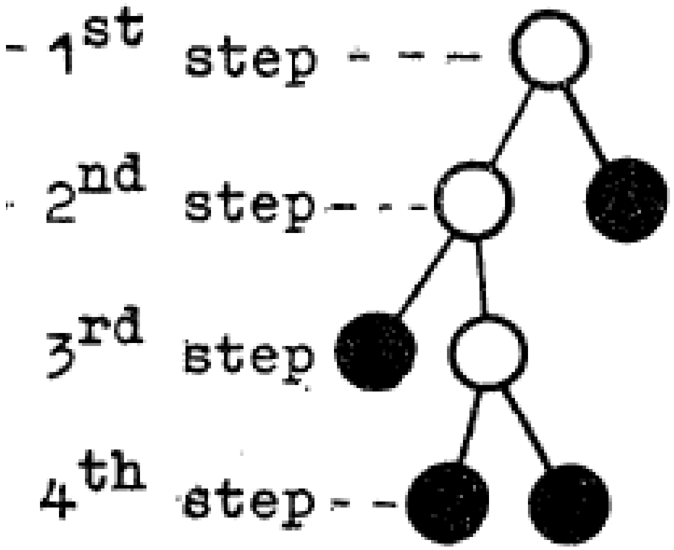
- Earthquake rupture = cascade process of topplings, but:

M. A. Francisco

(1) No domino effect: one toppling does not lead to another one and so on

(2) Pieces are not in a row, rather, in a network or tree, and disordered

⇒ When one piece topples, what happens next is random



appadvice.com

Otsuka, *Zisin* 1971

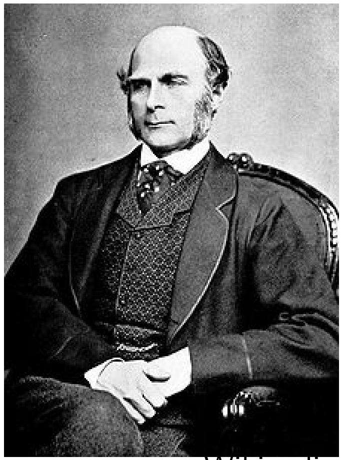
Seismic fault =
patches that may fail and trigger other patches to fail with some probability and so on

Kanamori & Mori, in Boschi et al. 2000

Besides gambling, many probabilists have been interested in reproduction

G. Grimmett and D. Stirzaker, 2001

Galton-Watson (Branching) Process (1873)


Wikipedia


www.wolframalpha.com

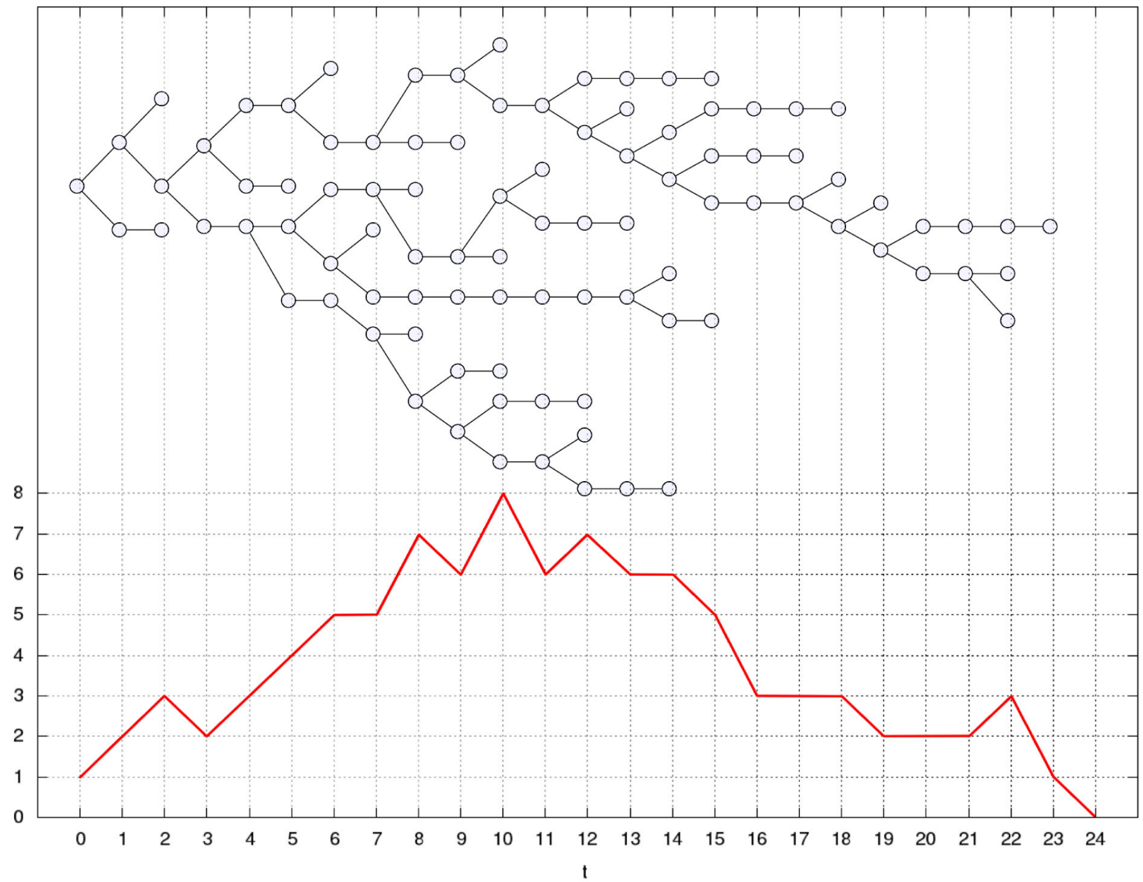
- Definition
  Start with 1 “element” (parent) which generates $K = 0, 1, \ldots$ elements (offsprings) with some probability $p_0, p_1, \ldots$ and so on...

  K's are independent identically distributed

  Galton was not interested in earthquakes
  Rather, he was worried by the extinction of prominent families:
  a rise in physical comfort and intellectual capacity is necessarily accompanied by diminution in “fertility”... If that conclusion be true, our population is chiefly maintained though the “proletariat,” and thus a large element of degradation is inseparably connected with those elements which tend to ameliorate the race

Extinction

- $N_t =$ total number of elements in generation t (with $N_0 = 1$)



Extinction $\Rightarrow$ $N_t = 0$ at some t

- Extinction = extinction in $t = 1$ or in $t = 2$ or $\ldots \in \lim_{t\to\infty}\{N_t = 0\}$

$$\Rightarrow P_{extinction} = \lim_{t\to\infty} P\{N_t = 0\}$$

- Probability generating function of a discrete random variable X

$$f_X(z) = \langle z^X \rangle = \sum_{x=0}^{\infty} P\{X = x\} z^x = P\{X = 0\} + P\{X = 1\} z + \ldots$$

$$\Rightarrow f_X(0) = P\{X = 0\}$$

This is valid for any random variable, also for N_t, so,

$$\Rightarrow P_{extinction} = \lim_{t\to\infty} f_{N_t}(0)$$

which is easier to calculate

- Main equation

$$N_{t+1} = \sum_{i=1}^{N_t} K_i(t)$$

- If N_t were a constant

$$f_{N_{t+1}}(z) = [f_K(z)]^{N_t}$$

Proof:

$$f_{N_{t+1}}(z) = \langle z^{N_{t+1}} \rangle = \langle z^{\sum_i K_i} \rangle = \langle z^{K_1} \dots z^{K_{N_t}} \rangle = \langle z^{K_1} \rangle \dots \langle z^{K_{N_t}} \rangle = [f_K(z)]^{N_t},$$

assuming independence.

- Let us repeat, $N_{t+1} = \sum_{i=1}^{N_t} K_i(t)$. If N_t is constant, $f_{N_{t+1}}(z) = [f_K(z)]^{N_t}$
- But N_t is random, so

$$f_{N_{t+1}}(z) = f_K^{t+1}(z)$$

with $f_K^{t+1}(z) = f_K(f_K(\ldots f_K(z)\ldots)) \; = \;$ composition $t+1$ times

Proof:

$$f_{N_{t+1}}(z) = \langle z^{N_{t+1}} \rangle = \Big\langle \langle z^{N_{t+1}} \rangle_{K_i} \Big\rangle_{N_t} = \langle [f_K(z)]^{N_t} \rangle_{N_t} = f_{N_t}(f_K(z)).$$

As $f_{N_1}(z) = f_K(z)$, then

$$\Rightarrow f_{N_2}(z) = f_{N_1}(f_K(z)) = f_K(f_K(z)) \equiv f_K^2(z)$$

and the result follows by induction

- In conclusion

$$P_{extinction} = \lim_{t \to \infty} P\{N_t = 0\} = \lim_{t \to \infty} f_{N_t}(0) = \lim_{t \to \infty} f_K^t(0)$$

- Let us calculate $f_{N_t}(z)$. Note that $N_1 = K \Rightarrow f_{N_1}(z) = f_K(z)$. Also

$$N_2 = \sum_{i=1}^{N_1} K_i, \qquad \text{and, in general} \qquad N_{t+1} = \sum_{i=1}^{N_t} K_i$$

If $M = \sum_{i=1}^{N} K_i$, with N constant, then

$$f_M(z) = \langle z^M \rangle = \langle z^{\sum_i K_i} \rangle = \langle z^{K_1} \ldots z^{K_N} \rangle = \langle z^{K_1} \rangle \ldots \langle z^{K_N} \rangle = [f_K(z)]^N,$$

assuming independence between the K_i's.
But if N is random, with $f_N(z) = \langle z^N \rangle$, then

$$f_M(z) = \langle z^M \rangle = \Big\langle \langle z^M \rangle_{K_i} \Big\rangle_N = \langle [f_K(z)]^N \rangle_N = f_N(f_K(z)).$$

$$\Rightarrow f_{N_2}(z) = f_{N_1}(f_K(z)) = f_K(f_K(z)) \equiv f_K^2(z)$$

In the same way

$$N_{t+1} = \sum_{i=1}^{N_t} K_i$$

As $f_M(z) = f_N(f_K(z))$, then, $f_{N_3}(z) = f_{N_2}(f_K(z)) = f_K^2(f_K(z)) \equiv f_K^3(z)$
In general, by induction

$$f_{N_t}(z) = f_{N_{t-1}}(f_K(z)) \equiv f_K^t(z) \qquad (t - \text{times composition})$$

Therefore

$$P_{extinction} = \lim_{t \to \infty} P\{N_t = 0\} = \lim_{t \to \infty} f_{N_t}(0) = \lim_{t \to \infty} f_K^t(0)$$

Expected size of population at t

- Property of $f_X(z) = \sum_{x=0}^{\infty} p_x z^x \quad \Rightarrow \quad f'_X(1) = \langle X \rangle$
- Therefore, $\langle N_t \rangle =$

Valid for any generating function, so, in the same way

$$f'_{N_t}(1) = \langle N_t \rangle$$

$$\left.\frac{df_{N_t}(z)}{dz}\right|_{z=1} = \left.\frac{df_K^t(z)}{dz}\right|_{z=1} = \left.\frac{df_K(f_K^{t-1}(z))}{dz}\right|_{z=1} = \left. f'_K(f_K^{t-1}(z))\frac{df_K^{t-1}(z)}{dz}\right|_{z=1}$$

by the chain rule, and by induction

$$\langle N_t \rangle = f'_K(f_K^{t-1}(z)) f'_K(f_K^{t-2}(z)) \cdots f'_K(f_K^2(z)) f'_K(f_K(z))\, f'_K(z)\Big|_{z=1}$$

using $f_K(1) = 1 \Rightarrow f_K^2(1) = 1$, etc., and $f'_K(1) = \langle K \rangle$ then $\langle N_t \rangle = \langle K \rangle^t$

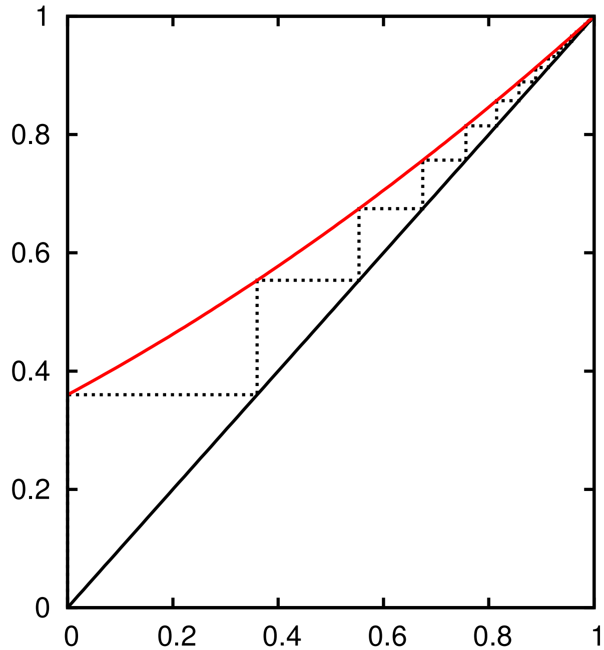
Extinction probability as a function of K

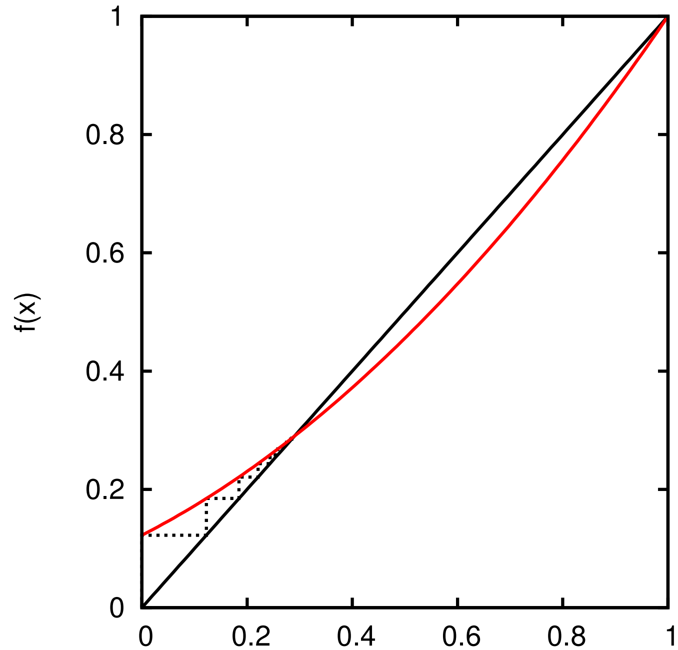
- Properties of $f_K(z)$ in $[0,1]$
  - ⋆ $f_K(0) = p_0$
  - ⋆ $f_K(1) =$
  - ⋆ $f'_K(1) =$
  - ⋆ $f'_K(z)$
  - ⋆ $f''_K(z)$

Extinction probability as a function of K

- Properties of $f_K(z)$
  - $\star$ $f_K(0) = p_0$
  - $\star$ $f_K(1) = 1$
  - $\star$ $f'_K(1) = \langle K \rangle$
  - $\star$ $f'_K(z) \geq 0$
  - $\star$ $f''_K(z) \geq 0$

  Valid for all probability generating functions





$$\langle K \rangle \leq 1 \Rightarrow P_{extinction} = \lim_{t \to \infty} f^t(0) = 1,$$ i.e., extinction for sure

$$\langle K \rangle > 1 \Rightarrow P_{extinction} = \lim_{t \to \infty} f^t(0) = z^* < 1,$$ i.e., non-sure extinction

Except for the "monarchic" case $p_1 = 1$, which has $\langle K \rangle = 1$ but $P_{extinction} = 0$

Phase transition in branching processes

- The fixed point condition for the probability of non-extinction $\rho = 1 - P_{extinction}$,

$$P_{extinction} = 1 - \rho = f_K(P_{extinction}) = f_K(1-\rho) = \sum_{k=0}^{\infty} p_k (1-\rho)^k$$

(because $P\{A\} + P\{no\,A\} = 1$). Expanding using the binomial theorem

$$1 - \rho = \sum_{k=0}^{\infty} p_k \left[1 - k\rho + \frac{1}{2}k(k-1)\rho^2 - \ldots \right] =$$

$$= \sum_{k=0}^{\infty} p_k - \left(\sum_{k=0}^{\infty} p_k k \right) \rho + \frac{1}{2} \left(\sum_{k=0}^{\infty} p_k k(k-1) \right) \rho^2 + \ldots =$$

$$= 1 - \langle K \rangle \rho + \frac{1}{2} \langle K(K-1) \rangle \rho^2 + \ldots$$

- For small ρ (large $P_{extinction}$), introducing $\phi = \langle K(K-1)\rangle$ (2nd factorial moment)

$$\frac{1}{2}\phi\rho^2 - (\langle K\rangle - 1)\rho \simeq 0$$

which has 2 solutions,

$$\rho = 0 \qquad \text{and} \qquad \rho \simeq 2\frac{\langle K\rangle - 1}{\phi}$$

We need to consider the solution closer to (but smaller than) 1, so

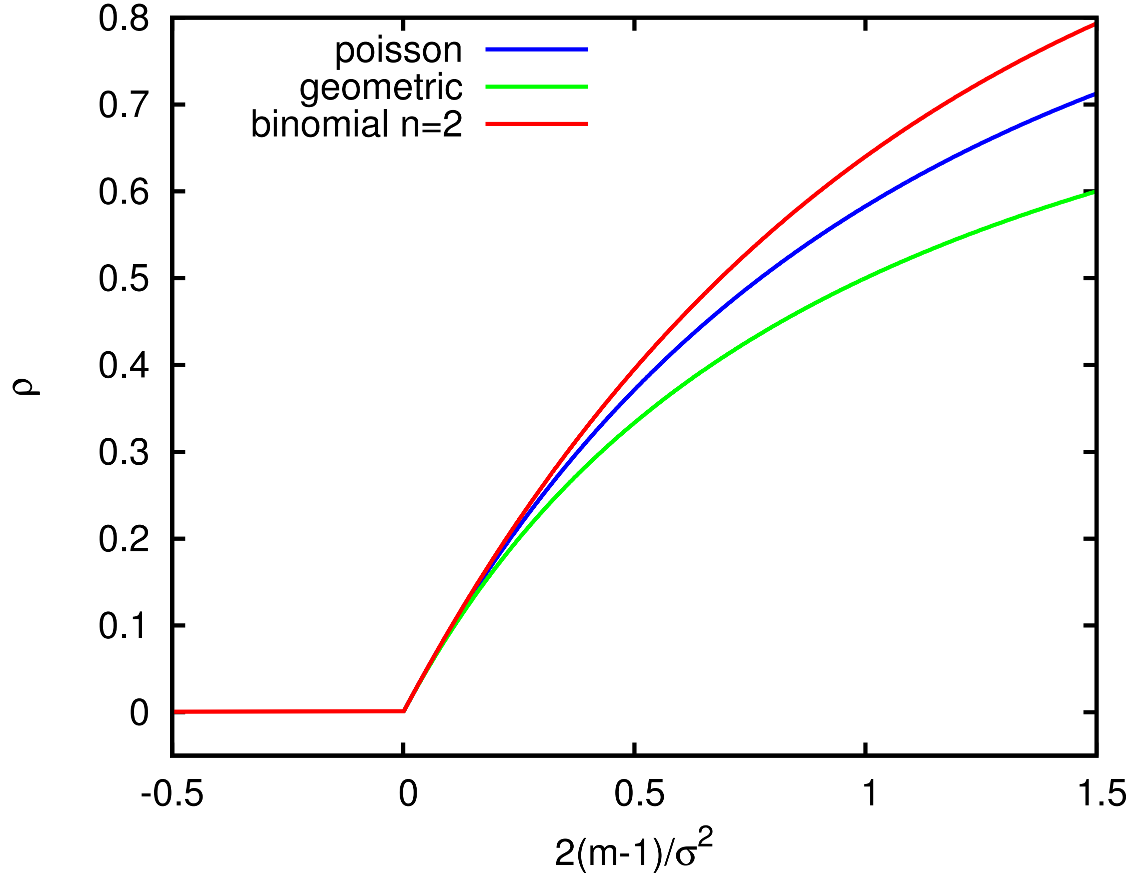
$$\rho = 0 \quad \text{for } \langle K\rangle \leq 1 \qquad \text{and} \qquad \rho \simeq 2\frac{\langle K\rangle - 1}{\sigma_c^2} \quad \text{for } \langle K\rangle \geq 1$$

where we have used $\phi = \sigma^2 + \langle K\rangle(\langle K\rangle - 1)$, if $\rho \simeq 0$ then $\langle K\rangle \simeq 1$ and $\phi \simeq \sigma_c^2$

- The transition is continuous, but sharp $\Rightarrow$ 2nd order phase transition

The case $\langle K\rangle = 1$ is critical, as it separates two very different behaviors

Universality: close to the critical point



$$m = \langle K \rangle$$

Continuous (or second order) phase transition

- Let m be a control parameter ($\langle K \rangle$ in branching or temperature, etc.)

  Let ρ be an order parameter (non-extinction probability, magnetization, etc.)

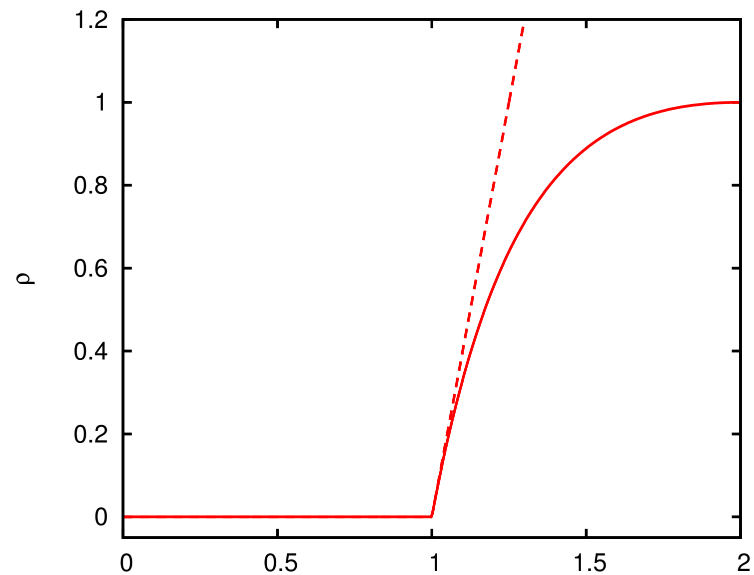
  Then

$$\rho \propto \begin{cases} 0 & \text{for } m \text{ below } m_c = \text{critical point} \\ (m - m_c)^\beta & \text{for } m \text{ above but close to } m_c \end{cases}$$

- Abrupt change in the derivative

  The derivative is discontinuous if $\beta \leq 1$

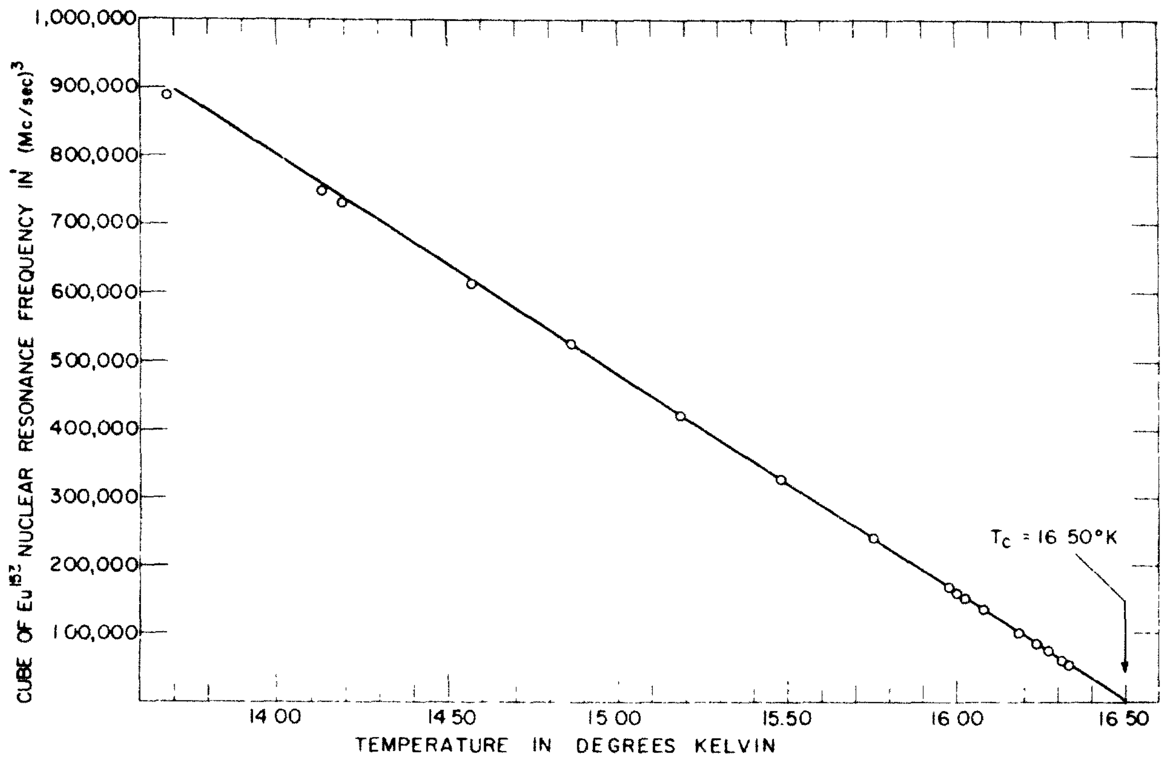
- For a branching process,
  $m_c = 1$ and $\beta = 1$



- For a magnetic system, m is the inverse of the temperature, ρ is magnetization

$\Rightarrow$ m_c is the inverse of Curie temperature and $\beta = 1/3$

Cube of magnet- ization



Magnetization dissappears sharply

Heller & Benedek

Example: binomial number of offsprings

- Each element has only a fixed number of trials n to generate other elements

$$p_k = P\{K = k\} = \left(\begin{array}{c} n \\ k \end{array} \right) p^k q^{n-k}, \text{ for } k = 0, 1, \ldots n.$$

with p the probability of being successful in each trial, and $q = 1 - p$

- The probability generating function

$$f_K(z) = \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) q^{n-k} p^k z^k = (q + pz)^n.$$

using the binomial theorem. We will consider $n = 2$

- $P_{extinction}$ will come from the smallest solution in $[0, 1]$ of

$$z^* = (q + pz^*)^2 \Rightarrow z^* = \frac{1 - 2pq \pm \sqrt{(1 - 2pq)^2 - 4p^2q^2}}{2p^2}$$

but for the square root we can write $\sqrt{1 - 4p(1 - p)} = \sqrt{(1 - 2p)^2} = (1 - 2p)$

$$\Rightarrow z = \frac{1 - 2p + 2p^2 \pm (1 - 2p)}{2p^2} = \left\{ \begin{array}{ll} (1 - 2p + p^2)/p^2 & = (q/p)^2 \\ p^2/p^2 & = 1 \end{array} \right.$$

The smallest root depends on whether p is below or above $1/2$

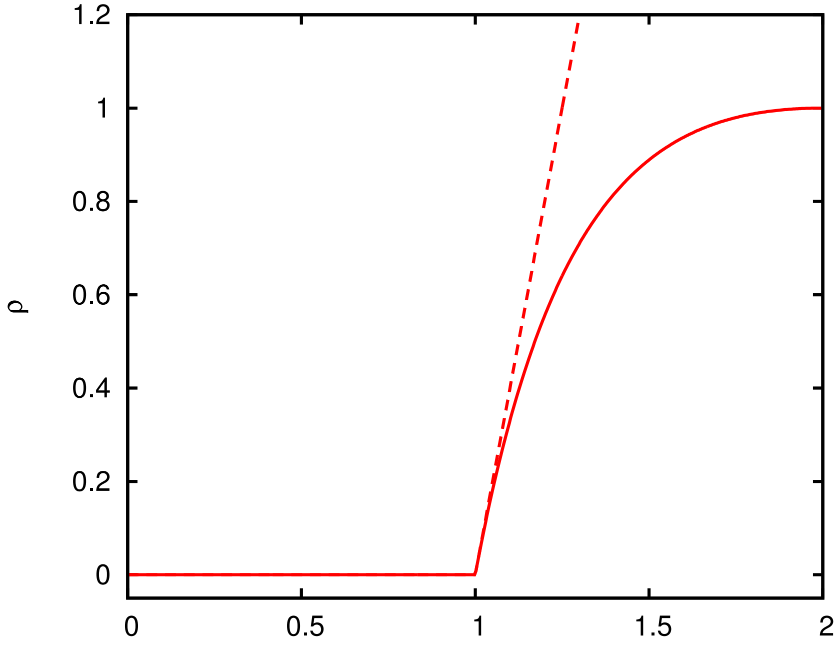
$$P_{extinction} = \begin{cases} 1 & \text{for } p \leq 1/2 \\ (q/p)^2 & \text{for } p \geq 1/2 \end{cases}$$

As $\langle K \rangle = np = 2p$ the critical case $\langle K \rangle = 1$ corresponds to $p = p_c = 1/2$
(in agreement with the behavior of $P_{extinction}$)

- In terms of the non-extinction probability $\rho = 1 - P_{extinction}$

$$\rho = 0 \;\text{ for } p$$

$$\rho = 1 - \left(\frac{q}{p}\right)^2 = \frac{2p-1}{p^2} =$$



using $\langle K \rangle = 2p$

- Expanding around $\langle K \rangle - 1 \simeq 0$

$$\rho \simeq 4(\langle K \rangle - 1) \simeq 2\frac{\langle K \rangle - 1}{\sigma_c^2} \text{ for } p \geq \frac{1}{2}$$

with $\sigma^2 = 2pq$ and $\sigma_c^2 = 1/2$
Then, $p = p_c = 1/2$ or $\langle K \rangle = 1$
is the critical point

Total size of the population

- The size of the population, summing across generations is

$$S = \sum_{t=0}^{\infty} N_t,$$

  - ⋆ total number of individuals that have ever existed, or
  - ⋆ total number of domino pieces toppling,
  - ⋆ "size" of an earthquake, etc...

- Its mean value, for $\langle K \rangle < 1$, using the geometric series, and $\langle N_t \rangle = \langle K \rangle^t$ (new!)

$$\langle S \rangle = \langle N_0 \rangle + \langle N_1 \rangle + \langle N_2 \rangle + \ldots = 1 + \langle K \rangle + \langle K \rangle^2 + \ldots = \frac{1}{1 - \langle K \rangle}$$

Note that when $\langle K \rangle \to 1$, the probability of extinction is 1, but $\langle S \rangle \to \infty$ (!)

Total size of the population: binomial case

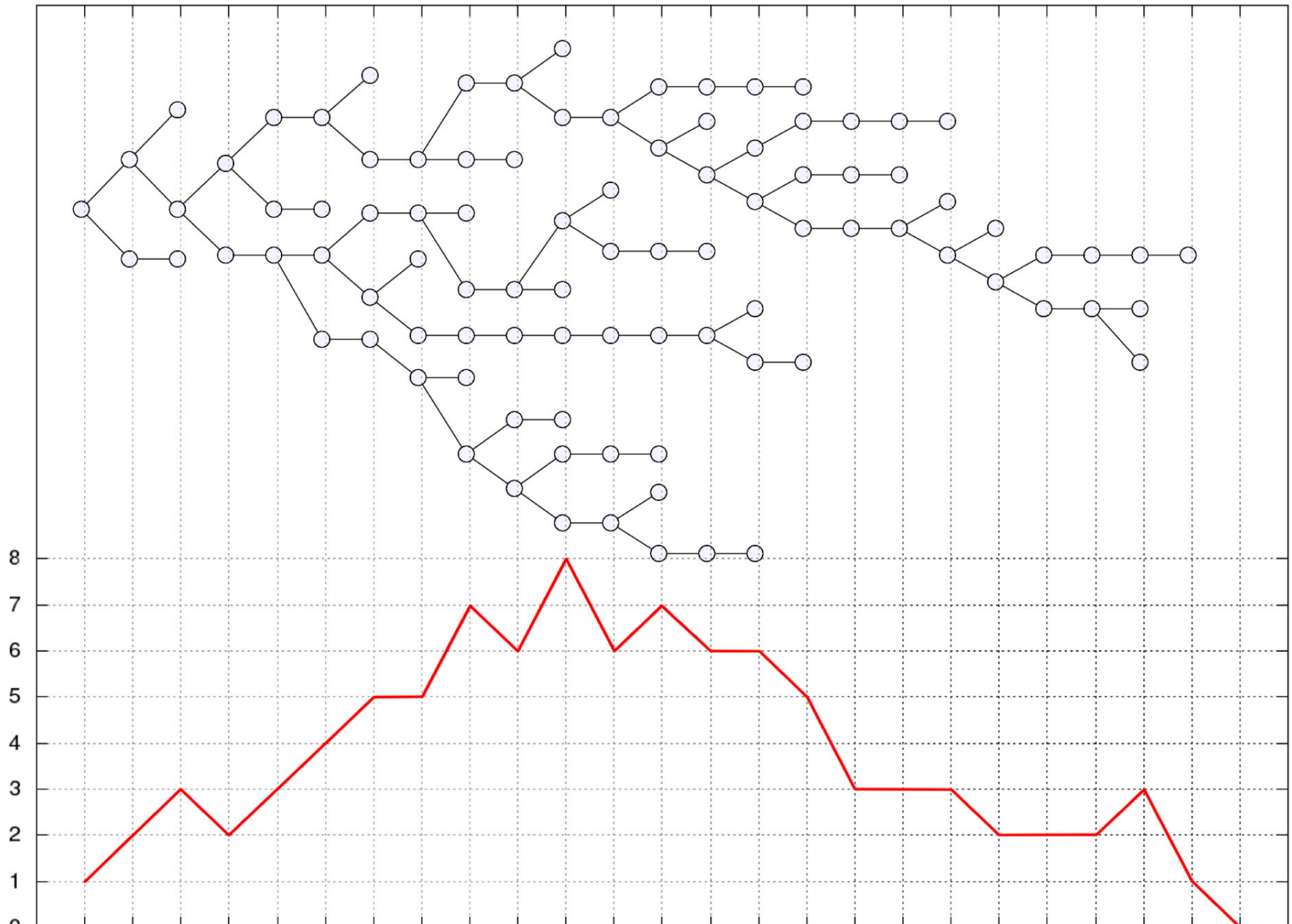
- Each element has only a fixed number of trials n to generate other elements

$$p_k = P\{K = k\} = \binom{n}{k} p^k q^{n-k}, \text{ for } k = 0, 1, \ldots n.$$

with p the probability of being successful in each trial, and $q = 1 - p$

- Remember $\langle K \rangle = np$, so the critical point is at $p_c = 1/n$

- Representation of a branching process as a tree (connected graph with no loops).
  - ⋆ Each element is associated to a node
  - ⋆ Branches linking nodes indicate an offspring relationship between two nodes


8
7
6
5
4
3
2
1
0

- Representation of a branching process as a tree (connected graph with no loops).
  - ⋆ Each element is associated to a node
  - ⋆ Branches linking nodes indicate an offspring relationship between two nodes

- All nodes have just one incoming branch, except the one in the zero generation
  - ⋆ the number of branches is the number of nodes minus 1, i.e., $s-1$
  - ⋆ the number of possible branches arising from s nodes is ns (in a n-tree)
  - ⋆ the number of missing branches (non-successful trials) is $ns-(s-1)$

  A particular tree of size s comes with a probability

$$p^{s-1}(1-p)^{(n-1)s+1} \text{ with } s=1,2,\ldots$$

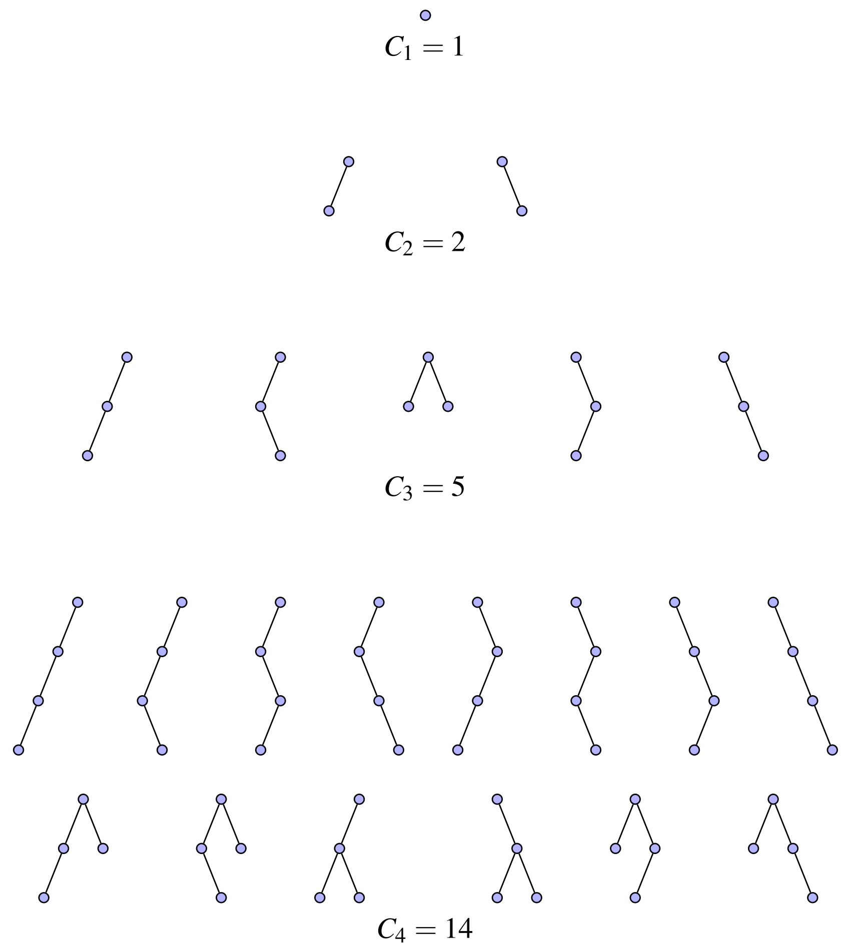
- For $n = 2$, the probability of having an undefined tree of size $s = 1, 2 \ldots$ comes from the Catalan numbers! ...

$$P\{S = s\} = C_s p^{s-1}(1-p)^{s+1} =$$
$$= \frac{1}{s+1}\begin{pmatrix} 2s \\ s \end{pmatrix} p^{s-1}(1-p)^{s+1}$$

with $C_s = \frac{1}{s+1}\begin{pmatrix} 2s \\ s \end{pmatrix}$ the number of different trees of size s, called Catalan numbers

The trees are the internal part of rooted binary trees

Can you draw them?



Calculation of the Catalan numbers

- Let us decompose a tree of size s into its root (zeroth generation) and the rest
  This can be done as
  - ⋆ A subtree of size $s-1$ in the 1st branch and another of size 0 in the 2nd
  - ⋆ A subtree of size $s-2$ in the 1st branch and another of size 1 in the 2nd
  - ⋆ ...
  - ⋆ A subtree of size 0 in the 1st branch and another of size $s-1$ in the 2nd

  So, the total number of trees of size s is

$$C_s = C_0 C_{s-1} + C_1 C_{s-2} + \cdots + C_{s-2} C_1 + C_{s-1} C_0 \qquad \text{with } C_0 = 1$$

- We define a generating function for the Catalan numbers

$$h(x) = C_0 + C_1 x + C_2 x^2 + \ldots = \sum_{s=0}^{\infty} C_s x^s$$

The properties of the Catalan numbers will allow the calculation of $h(x)$

$$[h(x)]^2 = \sum_{i,j=0}^{\infty} C_i C_j x^{i+j} = \sum_{s=0}^{\infty} \underbrace{\left[\sum_{i+j=s} C_i C_j\right]}_{C_{s+1}} x^s = \frac{1}{x}\sum_{s=0}^{\infty} C_{s+1} x^{s+1} = \frac{h(x) - C_0}{x}$$

so

$$h(x) = \frac{1 \pm \sqrt{1-4x}}{2x}$$

but this tell us nothing yet. Using the Taylor expansion of $\sqrt{1-x}$

$$\sqrt{1-x} = 1 - \frac{x}{2} - \frac{1}{4}\left(\frac{x^2}{2!}\right) - \frac{3}{8}\left(\frac{x^3}{3!}\right) - \ldots = 1 - \frac{x}{2} - \sum_{s=1}^{\infty} \frac{(2s-1)!!}{2^{s+1}(s+1)!} x^{s+1}$$

then

$$\sqrt{1-4x} = 1 - 2x - \sum_{s=1}^{\infty} \frac{(2s-1)!!2^{s+1}}{(s+1)!} x^{s+1}$$

and so, taking the minus sign (otherwise $h(x)$ is not a g.f.)

$$h(x) = \frac{1-\sqrt{1-4x}}{2x} = 1 + \frac{1}{2x}\sum_{s=1}^{\infty} \frac{(2s-1)!!2^{s+1}}{(s+1)!} x^{s+1} = 1 + \sum_{s=1}^{\infty} \frac{(2s-1)!!2^{s}}{(s+1)!} x^{s}$$

then the Catalan numbers are, and using $(2s)! = (2s)!!(2s-1)!! = s!2^s(2s-1)!!$

$$C_s = \frac{(2s-1)!!2^s}{(s+1)!} = \frac{(2s)!}{s!(s+1)!} = \frac{1}{s+1}\begin{pmatrix} 2s \\ s \end{pmatrix}$$

the latter being valid for $s = 0, 1, 2 \ldots$

Coming back to the Taylor expansion of $\sqrt{1-4x}$

$$\sqrt{1-4x} = 1 - 2x \sum_{s=0}^{\infty} C_s x^s$$

Parenthesis: many uses of the Catalan numbers

Davis, *geometer* 2010

- Number of balanced configurations with n pairs of parenthesis

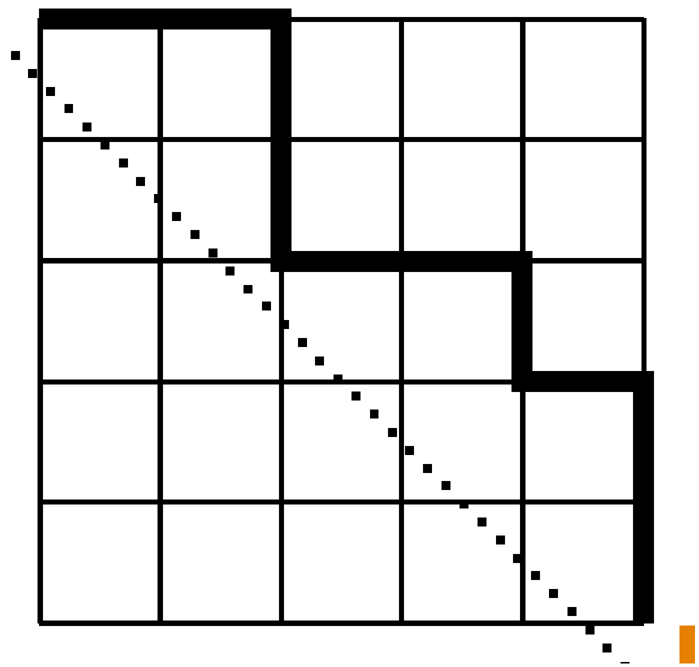
$n=0$:	`*`	1 way
$n=1$:	`()`	1 way
$n=2$:	`()(), (())`	2 ways
$n=3$:	`()()(), ()(()), (())(), (()()), ((()))`	5 ways

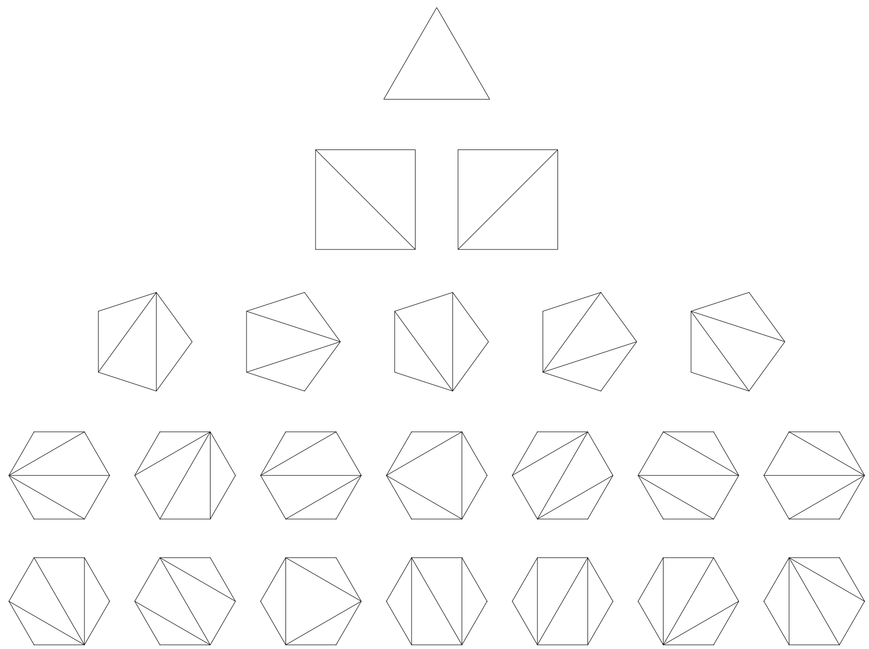
$n=4$:	`()()()(), ()()(()), ()(())(), ()(()()), ()((())),`
	`(())()(), (())(()), (()())(), ((()))(), (()()()),`
	`(()(())), ((())()), ((()())), (((())))`
		14 ways

- Number of mountains profiles with n upstrokes and n downstrokes

<table>
<tr><td>$n=0$:</td><td><pre>*</pre></td><td>1 way</td></tr>
<tr><td>$n=1$:</td><td><pre>/\</pre></td><td>1 way</td></tr>
<tr><td>$n=2$:</td><td><pre>   /\
/\/\, / \</pre></td><td>2 ways</td></tr>
<tr><td>$n=3$:</td><td><pre>   /\
           /\   /\   /\/\   / \
/\/\/\, /\/ \, / \/\, /   \, /   \</pre></td><td>5 ways</td></tr>
</table>

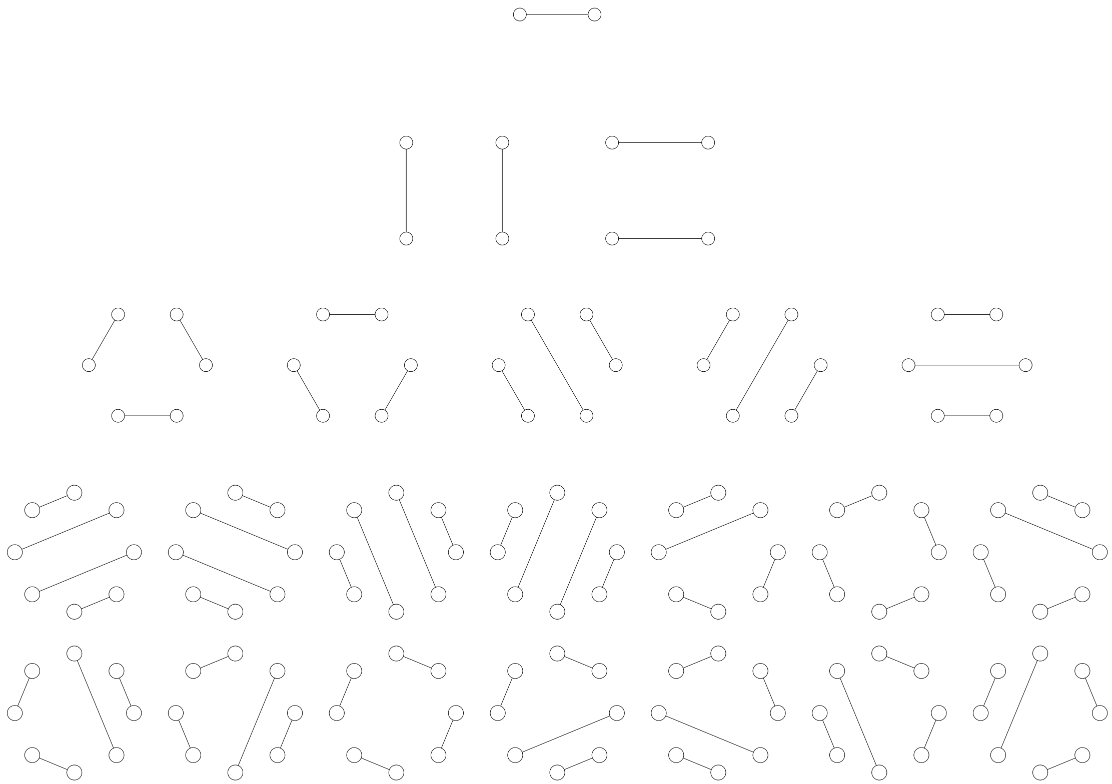
- Number of paths above (or on) the diagonal in a $n \times n$ lattice





- Number of triangulations of polygons with $n+2$ sides

- Non-crossing hand-shaking configurations of $2n$ people in a round table



- Many more!

Stanley *Enumerative Combinatorics 1999*

Normalization of the size distribution

- $P\{S=s\}$ is normalized for $p \leq 1/2$ but not for $p > 1/2$

$$\sum_{s=1}^{\infty} P\{S=s\} = \frac{q}{p} \sum_{s=1}^{\infty} C_s p^s q^s = \frac{q}{p}[h(pq)-1]$$

with $q = 1-p$ and introducing $h(x) = \sum_{s=0}^{\infty} C_s x^s$. As $h(x) = (1-\sqrt{1-4x})/(2x)$

$$h(pq) = \frac{1-\sqrt{1-4pq}}{2pq} = \frac{1-\sqrt{(1-2p)^2}}{2pq} = \frac{1-|1-2p|}{2pq} = \begin{cases} \frac{2p}{2pq} = \frac{1}{q} & \text{if } 1 \geq 2p \\ \frac{2(1-p)}{2pq} = \frac{1}{p} & \text{if } 1 \leq 2p \end{cases}$$

Therefore

$$\sum_{s=1}^{\infty} P\{S=s\} = \frac{q}{p}[h(pq)-1] = \begin{cases} \frac{q}{p}(\frac{1}{q}-1) = 1 & \text{if } p \leq 1/2 \\ \frac{q}{p}(\frac{1}{p}-1) = (q/p)^2 & \text{if } p \geq 1/2 \end{cases}$$

which turns out into

$$\sum_{s=1}^{\infty} P\{S=s\} = P_{extinction}$$

But how does $P\{S=s\}$ look like?

And what this has to do with power laws?

- Summarizing, the size distribution

$$P\{S = s\} = C_s p^{s-1}(1-p)^{s+1} = \frac{1}{s+1}\left(\begin{array}{c} 2s \\ s \end{array}\right) p^{s-1}(1-p)^{s+1}$$

for a branching process with binomial distribution and $n = 2$

- But what this has to do with power laws??

Asymptotic total size of the population

- Using Stirling's approximation, valid for $s \to \infty$ Christensen & Moloney 2005; A.C. & Font-Clos 2013

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

the binomial coefficient turns out to be

$$\binom{2s}{s} = \frac{(2s)!}{s!s!} \sim \frac{4\pi s}{2\pi s}\frac{(2s)^{2s}}{s^{2s}} \sim \frac{4^s}{\sqrt{\pi s}}$$

and the Catalan number, replacing $s+1 \sim s$

$$C_s = \frac{1}{s+1}\binom{2s}{s} \sim \frac{4^s}{\sqrt{\pi}\, s^{3/2}}$$

essentially, an exponential increasing function of s

- Introducing the factor $p^{s-1}q^{s+1}$ we get $P\{S=s\}$

$$P\{S=s\} \sim \frac{q}{\sqrt{\pi}p}\frac{(4pq)^s}{s^{3/2}}$$

How does this function looks like for large s?

⋆ If $p(1-p) < 1/4 \Rightarrow p \neq 1/2 \Rightarrow$ decreasing exponential
⋆ If $p(1-p) = 1/4 \Rightarrow p = 1/2 \Rightarrow$ exponential dissapears $\Rightarrow$ power law!

It becomes more transparent writting

$$(4pq)^s = e^{s\ln[4p(1-p)]} = e^{-s/\xi(p)}$$

with the characteristic size defined as

$$\xi(p) = \frac{-1}{\ln[4p(1-p)]} = \left(\ln\frac{1}{4p(1-p)}\right)^{-1}$$

and then

$$P(S = s) \sim \frac{q}{\sqrt{\pi} p} \frac{e^{-s/\xi(p)}}{s^{3/2}}$$

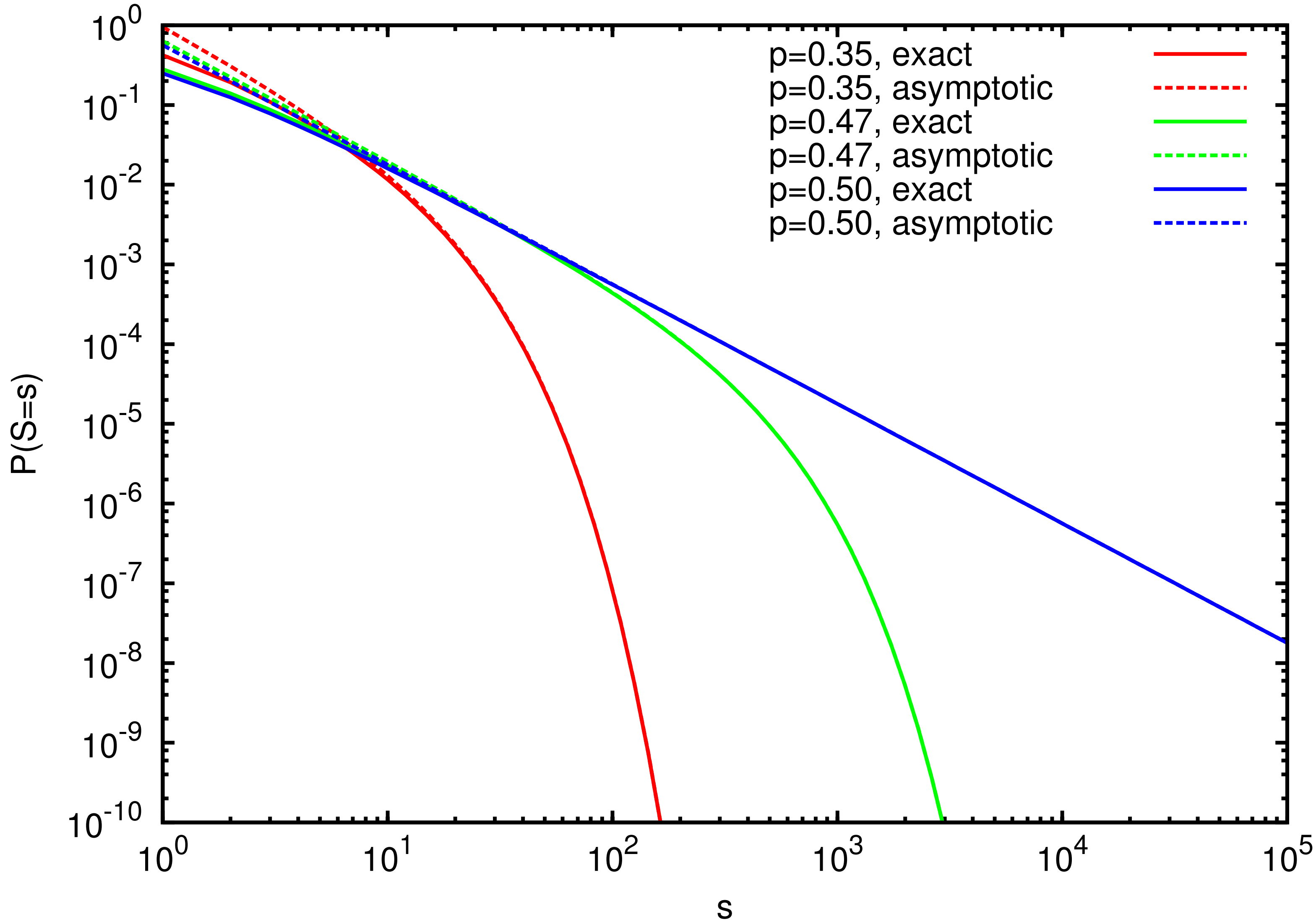
Case $p \neq 1/2$

- ⋆ For s large but $s \ll \xi(p) \Rightarrow$ power law with exponent $3/2$
- ⋆ For s large with $s \gg \xi(p) \Rightarrow$ exponential decay

Case $p = 1/2$

- ⋆ Then, $\xi \to \infty$ and for large s we obtain a power law

The critical exponent for the size distribution is 3/2



p=0.35, exact
p=0.35, asymptotic
p=0.47, exact
p=0.47, asymptotic
p=0.50, exact
p=0.50, asymptotic
P(S=s)
s

Divergence of the characteristic size

- Another critical exponent arises for the divergence of $\xi(p)$ at the critical point.

  Introducing the deviation with respect to the critical point, $\Delta \equiv p - p_c = p - 1/2$

$$p(1-p) = \left(\frac{1}{2} + \Delta\right)\left(\frac{1}{2} - \Delta\right) = \frac{1}{4} - \Delta^2$$

  So, close to the critical point (for small Δ)

$$\frac{1}{4p(1-p)} = \frac{1}{1-4\Delta^2} \simeq 1 + 4\Delta^2 + \ldots$$

  (using the formula of the geometric series), then

$$\ln \frac{1}{4p(1-p)} \simeq \ln(1+4\Delta^2) \simeq 4\Delta^2 + \ldots$$

(using the Taylor expansion of the logarithm at point 1), therefore

$$\xi(p) = \left(\ln \frac{1}{4p(1-p)} \right)^{-1} \simeq \frac{1}{4\Delta^2} + \dots$$

So, $\xi(p)$ diverges at the critical point as a power law, with an exponent $= 2$

Then, for s large and Δ small

$$P(S = s) \sim \frac{1}{\sqrt{\pi}} \frac{e^{-4(p-p_c)^2 s}}{s^{3/2}}$$

Expected value of the size

- We already know that for $\langle K \rangle < 1$ (i.e., $p < 1/2$, i.e., $\Delta < 0$)

$$\langle S \rangle = \langle N_0 \rangle + \langle N_1 \rangle + \langle N_2 \rangle + \ldots = 1 + \langle K \rangle + \langle K \rangle^2 + \ldots = \frac{1}{1 - \langle K \rangle} = -\frac{1}{2\Delta}$$

substituting $\langle K \rangle = 2p$ and $\Delta = p - 1/2 =$ deviation with respect criticality

This defines another critical exponent

- As $\xi(p) \simeq \Delta^{-2}/4$ close but below $p_c = 1/2$ then

$$\xi(p) \simeq \langle S \rangle^2$$

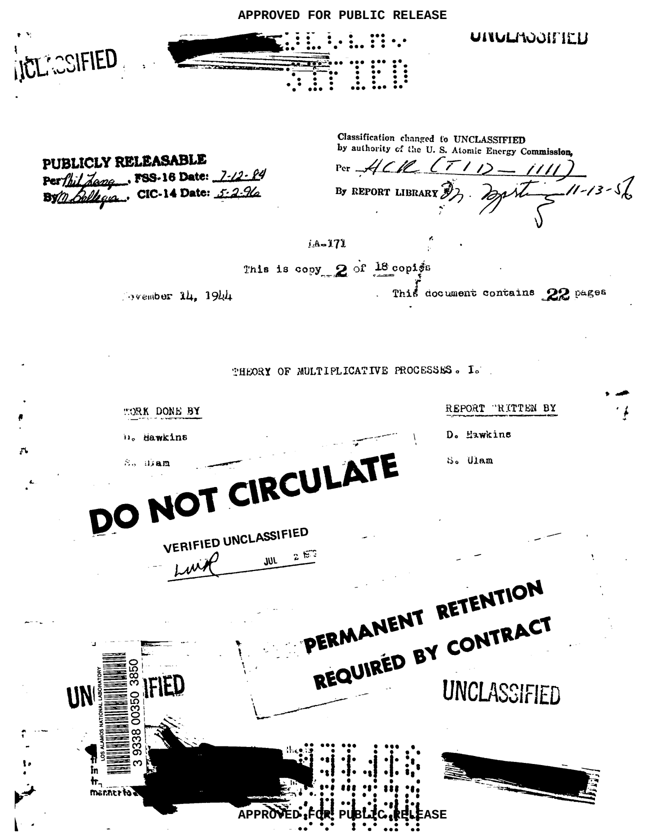
So, if the mean increases by 2, the extreme values given by ξ increase by 4

Total size of the population: general case

- Let $g(z) = f_S(z)$ be the generating function of $S = \sum_{t=0}^{\infty} N_t$. Then

$$g(z) = z f_K(g(z))$$

  with $f_K(z)$ the p.g.f. of the number of offsprings per element
- 1st demonstrated by Hawkins and Ulam in 1944 for nuclear chain reactions (as a part of the Manhattan project)
  - ⋆ A neutron may produce a fission reaction
  - ⋆ Each reaction releases neutrons
  - ⋆ Each neutron may trigger more reactions, and so on.



APPROVED FOR PUBLIC RELEASE

UNCLASSIFIED [redacted]

PUBLICLY RELEASABLE
Per Phil Long, FSS-16 Date: 7-12-94
By M. Ballegos, CIC-14 Date: 5-2-96

Classification changed to UNCLASSIFIED
by authority of the U. S. Atomic Energy Commission,
Per ACR (TID-1111)
By REPORT LIBRARY 11-13-56

LA-171

This is copy 2 of 18 copies

ovember 14, 1944

This document contains 22 pages

THEORY OF MULTIPLICATIVE PROCESSES. I.

WORK DONE BY	REPORT WRITTEN BY
D. Hawkins	D. Hawkins
S. Ulam	S. Ulam

DO NOT CIRCULATE

VERIFIED UNCLASSIFIED

PERMANENT RETENTION REQUIRED BY CONTRACT

UNCLASSIFIED

[redacted]

APPROVED FOR PUBLIC RELEASE

- Demonstration Hawkins & Ulam 1944; A.C. & Font-Clos, arXiv 2012

Consider the size from generation 1 to ∞ (excluding the $0-$th generation)

$$S_{\tilde{0}} = S - 1 = \sum_{t=1}^{\infty} N_t$$

with $q_s = P(S_{\tilde{0}} = s)$ and a generating function $\tilde{g}(z) = \sum_{\forall s} q_s z^s$

A size s in generations from 1 to ∞ can be decomposed into

⋆ a size k in the first generation, with probability p_k, and

⋆ a size $s-k$ in the remaining generations (from 2 to ∞) but starting with k elements; this has a probability $q_{s-k}^{(k)}$ (note that $q_s = q_s^{(1)}$)

Using the law of total probability,

$$q_s = \sum_{k=1}^{s} p_k q_{s-k}^{(k)},$$

except for $s = 0$, where $q_0 = p_0$

If we multiply by z^s and sum for all s we will obtain the g.f. of $S_{\tilde{0}}$

$$\tilde{g}(z) = p_0 + \sum_{s=1}^{\infty} \sum_{k=1}^{s} p_k q_{s-k}^{(k)} z^s = p_0 + \sum_{k=1}^{\infty} p_k \left[\sum_{s=k}^{\infty} q_{s-k}^{(k)} z^{s-k} \right] z^k$$

The term in [ ] is the g.f. of the size from 1 to ∞ generations
but starting with k elements ($N_1 = k$)
As these k parents are independent of each other
$\Rightarrow$ size will be the sum of k independent random variables each with g.f. $\tilde{g}(z)$
This yields $[\tilde{g}(z)]^k$ as the corresponding generating function,

$$[\tilde{g}(z)]^k = \sum_{s-k=0}^{\infty} q_{s-k}^{(k)} z^{s-k},$$

Substituting into the equation above

$$\tilde{g}(z) = p_0 + \sum_{k=1}^{\infty} p_k [\tilde{g}(z)]^k z^k = f_K(z\tilde{g}(z))$$

As $S = 1 + S_{\tilde{0}}$ we need to add an independent variable with g.f. $= z$
(as N_0 takes the value 1 with probability 1) then, the g.f. of the size from generation 0 to ∞ is the product $z\tilde{g}$, so

$$g(z) = z\tilde{g}(z) = zf_K(z\tilde{g}(z)) = zf_K(g(z))$$

- Binomial case. Substituting $f_K(z) = (q + pz)^2$ then

$$g(z) = zf_K(g(z)) = z(q + pg(z))^2 \qquad \Rightarrow \qquad g(z) = \frac{1 - 2pqz \pm \sqrt{1 - 4pqz}}{2p^2 z}$$

Using the Taylor expansion for the square root

$$\sqrt{1 - 4pqz} = 1 - 2pqz - \sum_{s=1}^{\infty} \frac{(2s-1)!!2^{s+1}}{(s+1)!}(pqz)^{s+1}$$

where we do not need to compute the Catalan numbers C_s, so, taking "'-"

$$g(z) = \frac{q}{p} \sum_{s=1}^{\infty} C_s (pqz)^s$$

From the coefficients we recover the probability distribution we knew

$$P\{S = s\} = C_s p^{s-1} q^{s+1}$$

- Geometric case. "Success" probability p and $q = 1 - p$ and values $k = 0, 1, 2 \ldots \infty$

$$p_k = P\{K = k\} = q^k p \qquad \Rightarrow \qquad f_K(z) = \sum_{k=0}^{\infty} p_k z^k = p \sum_{k=0}^{\infty} q^k z^k = \frac{p}{1 - qz}$$

using the geometric series. The generating function for the size is

$$g(z) = z f_K(g(z)) = \frac{pz}{1 - qg(z)} \quad \Rightarrow \quad g(z) = \frac{1 - \sqrt{1 - 4pqz}}{2q} = pz + \sum_{s=2}^{\infty} C_{s-1} q^{s-1} p^s z^s$$

where we have used the following, with C_i the $i-$th Catalan number

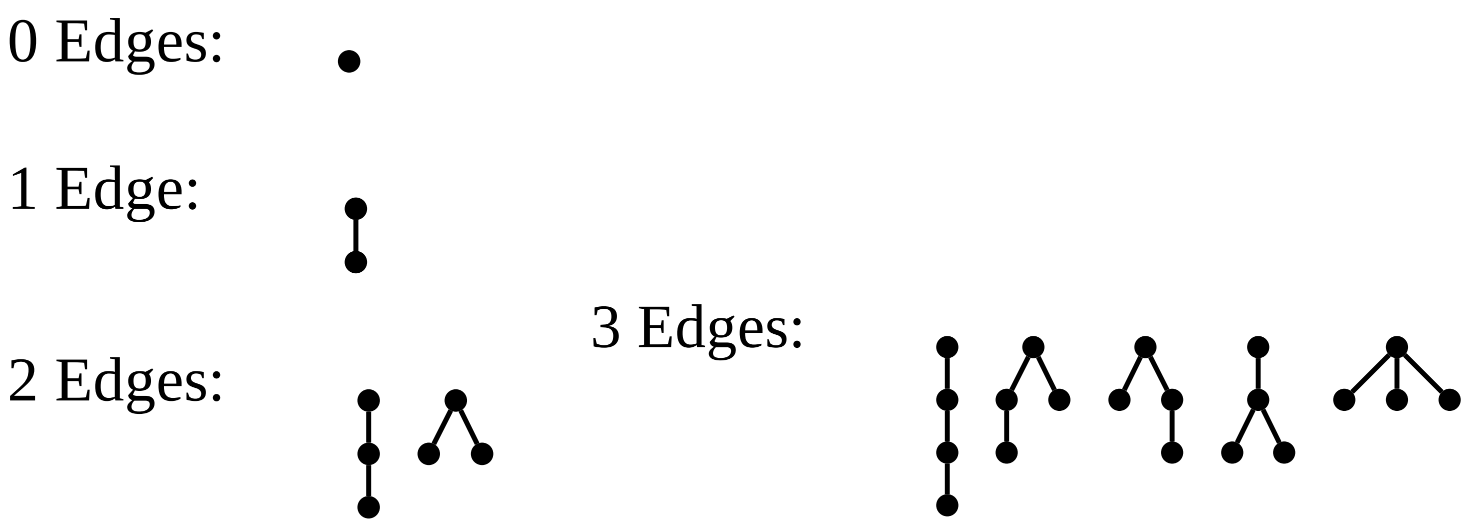
$$\sqrt{1 - 4x} = 1 - 2x - \sum_{i=1}^{\infty} 2C_i x^{i+1}$$

Therefore, the size distribution (without binary trees!)

$$P\{S = s\} = C_{s-1} q^{s-1} p^{s} \qquad \sim \frac{1}{4\sqrt{\pi} q} \frac{(4pq)^{s}}{s^{3/2}} \text{ for } s \to \infty$$

so we again obtain a critical exponent $= 3/2$ (and also the others)

- C_s also counts number of (non-necessarily-binary) trees with s edges



Davis 2010

- Normalization of the size distribution in the geometric case

$$\sum_{s=1}^{\infty} P\{S = s\} = C_{s-1} q^{s-1} p^{s} = \begin{cases} 1 & \text{if } q \leq 1/2 \\ p/q & \text{if } q \geq 1/2 \end{cases}$$

which corresponds to the probability of extinction in the geometric case
Note that $\langle K \rangle = q/p$, so $p_C = q_C = 1/2$

- Another offspring distribution
  - 0 offsprings with probability $q = 1 - p$
  - 2 offsprings with probability p

  Then $f_K(z) = q + pz^2$. The generating function for the size is

$$g(z) = z f_K(g(z)) = z(q + pg(z)^2) \quad \Rightarrow \quad g(z) = \frac{1 \pm \sqrt{1 - 4pqz^2}}{2pz} = \sum_{i=0}^{\infty} C_i p^i q^{i+1} z^{2i+1}$$

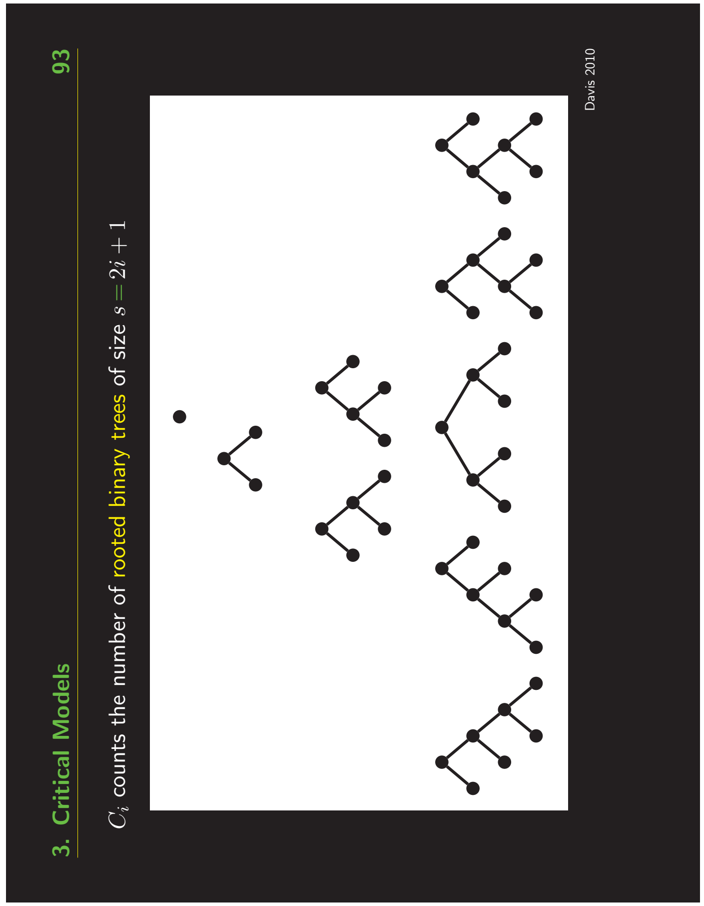
Therefore

$$P\{S = s\} = C_{\frac{s-1}{2}} \, p^{\frac{s-1}{2}} q^{\frac{s+1}{2}} \qquad \text{for } s = 1, 3, 5 \ldots$$

So, C_i counts the number of rooted binary trees of size $s = 2i + 1$
Asymptotically we do not scape from the exponent 3/2

$$P\{S = s\} \sim \sqrt{\frac{2q}{\pi p}} \frac{(4pq)^{s/2}}{s^{3/2}} \; \text{ for } s \to \infty$$

C_i counts the number of rooted binary trees of size $s = 2i + 1$



Finite size effects in branching processes

Garcia-Millan, Font-Clos & A. C. *Phys Rev E* 2015

- Let us consider a limitation in the number of generations: $t = 0, 1, \ldots L$ (this plays the role of boundaries)

  The probability of extinction, with $f(z) \equiv f_K(z)$, will be

$$P_{ext}(L) = f^L(0) \qquad < \qquad P_\infty = \lim_{t \to \infty} f^t(0)$$

- Consider a very large number of generations, n

  $\Rightarrow$ $f^n(0)$ will be close to $f^\infty(0) = P_\infty$

  Let us Taylor expand $f(f^n(0))$ around the fixed point P_∞

$$f^{n+1}(0) = f(f^n(0)) = P_\infty + f'(P_\infty)\,(f^n(0) - P_\infty) + \ldots$$

- Taking up to 2nd-order terms and arranging, the inverse of the distance is[1]

$$c_{n+1} \equiv \frac{1}{P_\infty - f^{n+1}(0)} = \frac{c_n}{M} + \frac{C}{M^2}$$

with $M = f'(P_\infty)$ and $C = f''(P_\infty)/2$. Iterating

$$c_{n+\ell} = \frac{c_n}{M^\ell} + \frac{C(1-M^\ell)}{M^{\ell+1}(1-M)}$$

In the subcritical case, $P_\infty = 1$, then $M = \langle K \rangle$ and $2C = \sigma^2 + \langle K \rangle(\langle K \rangle - 1)$, so

$$c_{n+\ell} = \frac{c_n}{\langle K \rangle^\ell} + \frac{\sigma^2(1-\langle K \rangle^\ell)}{2\langle K \rangle^{\ell+1}(1-\langle K \rangle)} - \frac{1-\langle K \rangle^\ell}{2\langle K \rangle^\ell}$$

[1] do not confuse distance to the fixed point with distance to the critical point

- Let us introduce a rescaled distance to the critical point $y = \ell(\langle K \rangle - 1)$, so

$$c_{n+\ell} = \frac{\sigma^2(1 - \langle K \rangle^\ell)}{2\langle K \rangle^{\ell+1}(1 - \langle K \rangle)} + \ldots = -\frac{\sigma_c^2(1 - e^y)\ell}{2e^y y}$$

with $\langle K \rangle = 1 + y/\ell$ and $\langle K \rangle^\ell = e^y$ and with ℓ large (then $\langle K \rangle$ is close to 1)

For $L = \ell + n \gg n$, we have that the probability of non-extinction will be

$$\rho(L) = 1 - P_{ext}(L) = 1 - f^L(0) = \frac{1}{c_L} = \frac{2e^y y}{\sigma_c^2(e^y - 1)L},$$

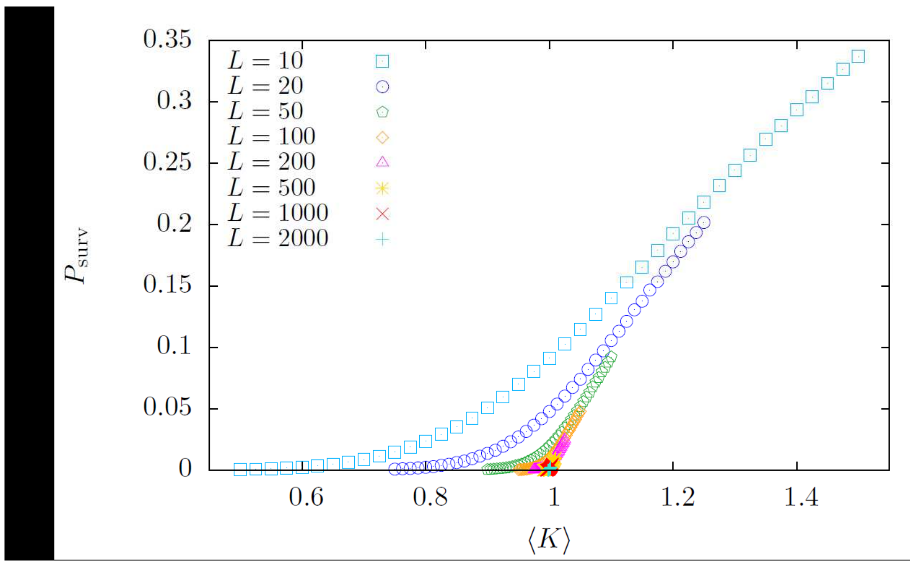
with $L \simeq \ell$. So, a scaling law is fulfilled, with scaling function $\mathcal{G}(y)$

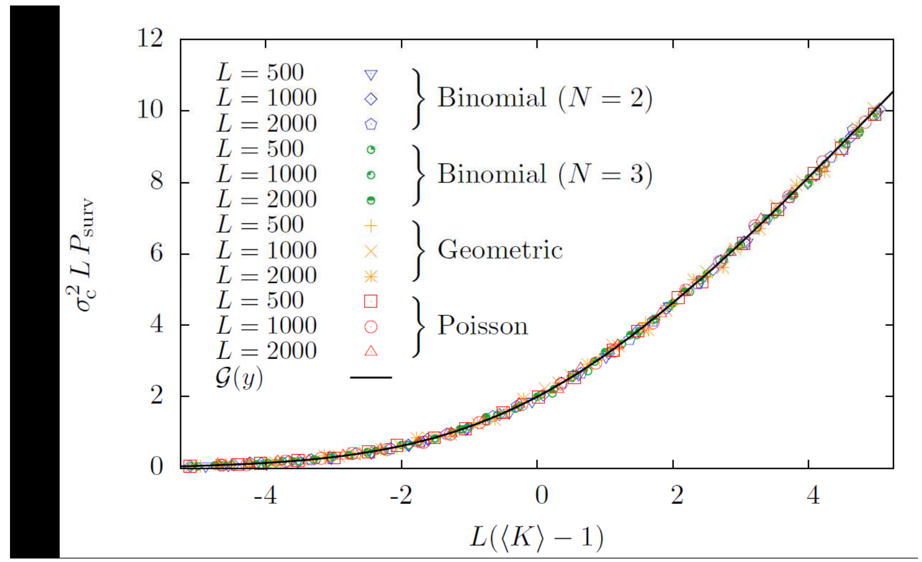
$$\rho(L) = \frac{1}{L\sigma_c^2}\mathcal{G}(L(\langle K \rangle - 1)) \quad \text{with} \quad \mathcal{G}(y) = \frac{2ye^y}{e^y - 1}$$

valid also for the supercritical case. This is called finite-size scaling

- Let us repeat

$$\rho(L) = \frac{1}{L\sigma_c^2}\mathcal{G}(L(\langle K\rangle - 1)) \quad \text{with} \quad \mathcal{G}(y) = \frac{2ye^y}{e^y - 1}$$



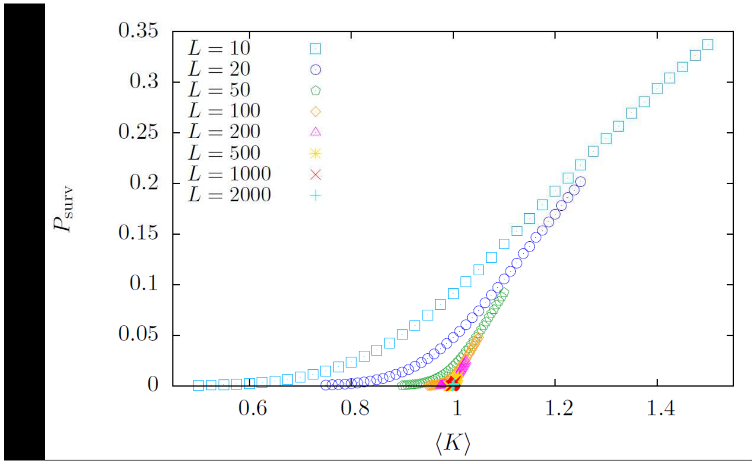


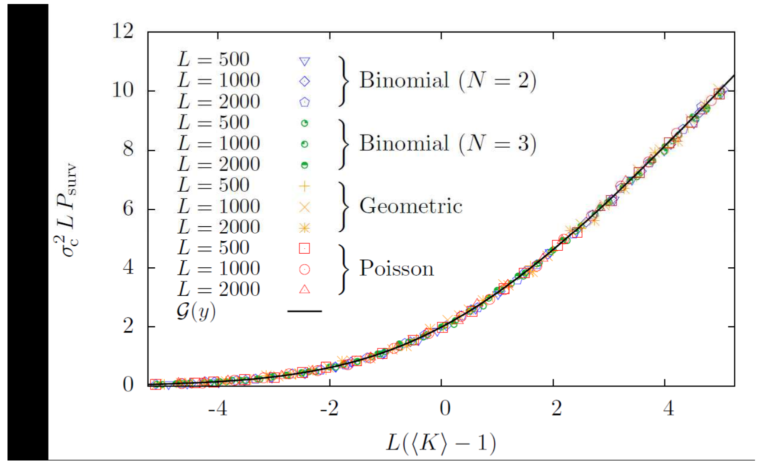
Phase transitions only exist in the infinite-system limit (thermodynamic limit)

$$\mathcal{G}(y) = \frac{2ye^y}{e^y - 1} \quad \rightarrow \quad \begin{cases} 0 & \text{when } y \rightarrow -\infty \\ 2 & \text{when } y \rightarrow 0 \\ 2y & \text{when } y \rightarrow \infty \end{cases}$$

So, for $L \rightarrow \infty$

$$\rho(L) = \frac{1}{L\sigma_c^2}\mathcal{G}(L(\langle K \rangle - 1)) \quad \rightarrow \quad \begin{cases} 0 & \text{for } \langle K \rangle < 1 \\ 2\sigma_c^{-2}L^{-1} & \text{for } \langle K \rangle = 1 \\ 2(\langle K \rangle - 1)/\sigma_c^2 & \text{for } \langle K \rangle > 1 \end{cases}$$





Simulation of a branching process

- Initialize $t = 0$ and $N_0 = 1$ (one single ancestor)

- Loop for t
  - ⋆ Simulate N_t values of K_i
  - ⋆ Compute $N_{t+1} = \sum_{i=1}^{N_t} K_i$
  - ⋆ If $N_{t+1} = 0 \Rightarrow$ stop
  - ⋆ $t = t + 1$

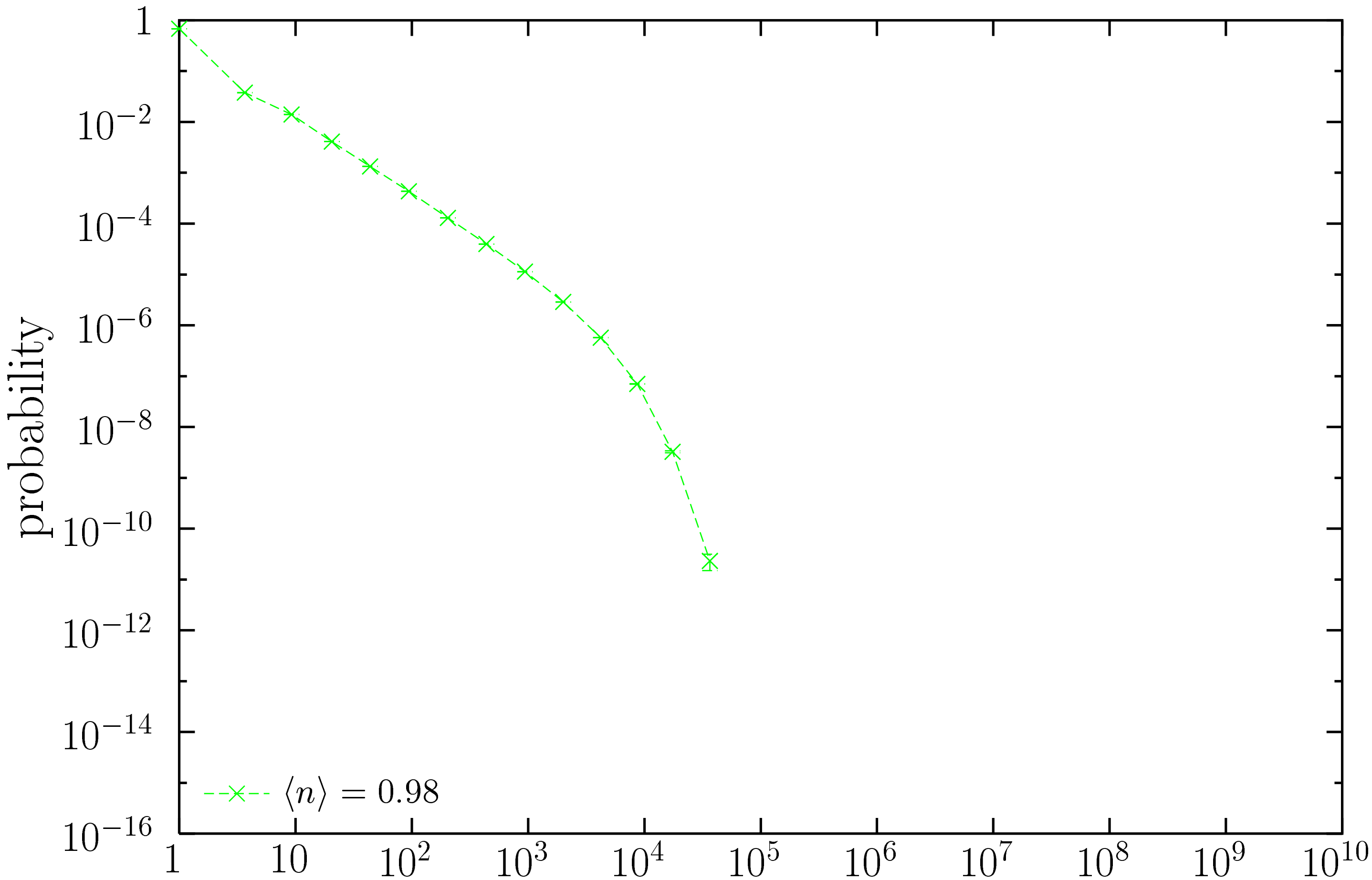
- For the twins-or-nothing example

$$K = \begin{cases} 2 & \text{if } u \leq p \\ 0 & \text{otherwise} \end{cases}$$

with u a uniform random number between 0 and 1

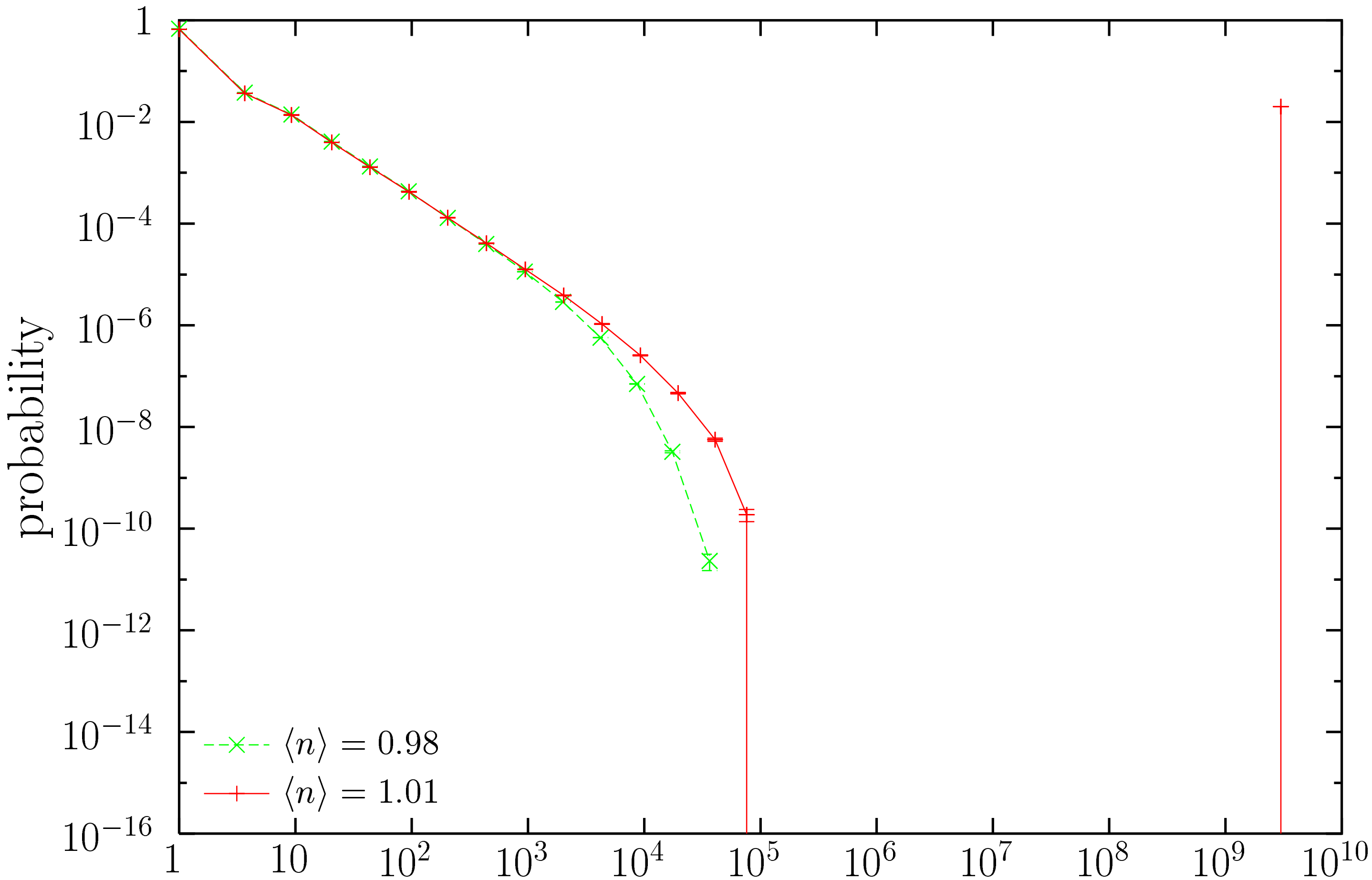
As the mean of the number of offsprings is $\langle K \rangle = 2p$, then, $p_c = 1/2$

Plot of $P\{S = s\}$ (with $\langle n \rangle = \langle K \rangle$)



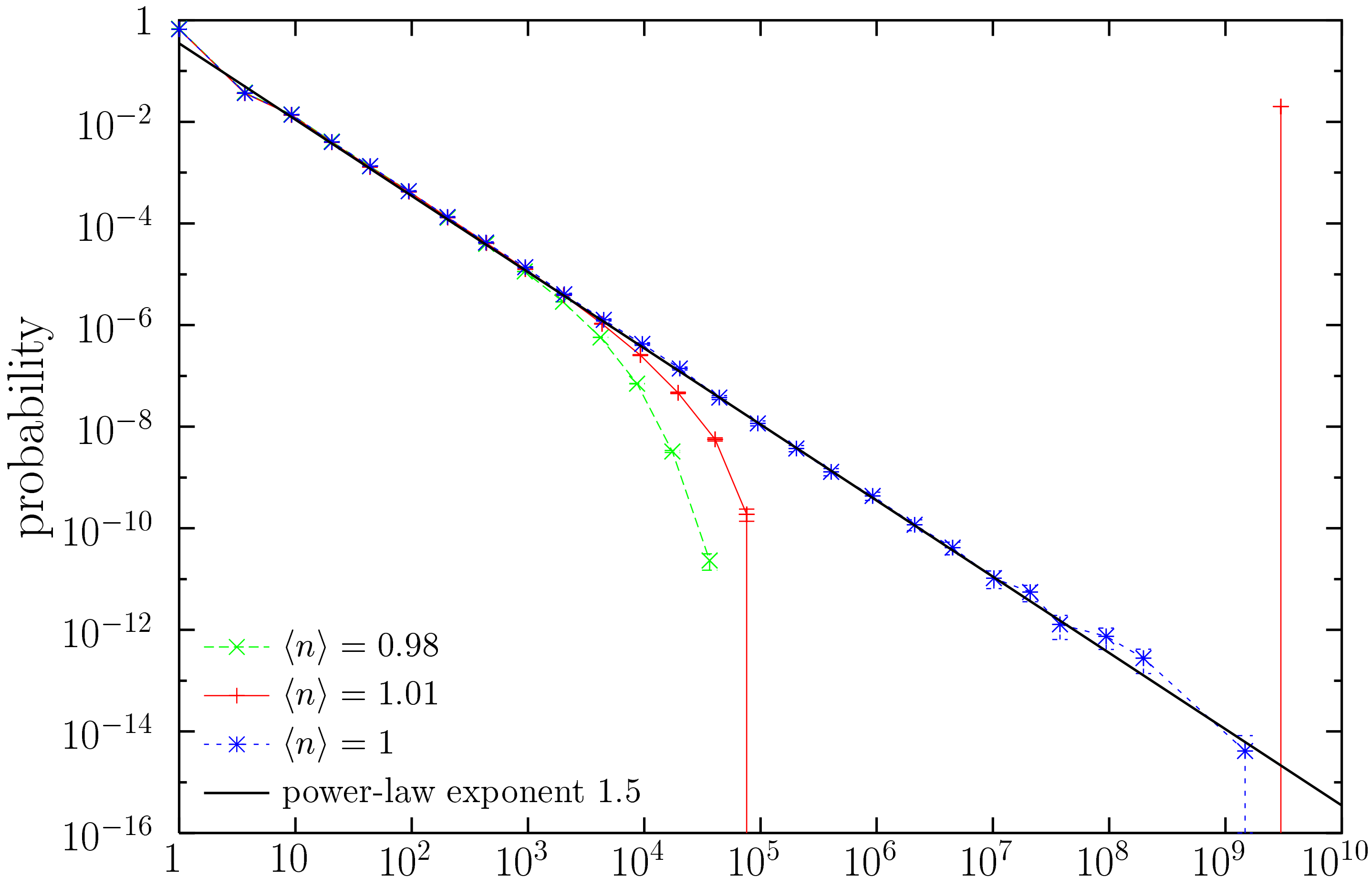
As the mean of the number of offsprings is $\langle K \rangle = 2p$, then, $p_c = 1/2$

Plot of $P\{S = s\}$ (with $\langle n \rangle = \langle K \rangle$)



As the mean of the number of offsprings is $\langle K \rangle = 2p$, then, $p_c = 1/2$

Plot of $P\{S = s\}$ (with $\langle n \rangle = \langle K \rangle$)



Earthquakes and branching processes

- Gutenberg-Richter power law holds only for $\langle K \rangle = 1$
  Critical branching process $\Rightarrow$ Fine tuning of mean number of offsprings
  $\Rightarrow$ Very difficult to get in practice!

- Agreement only qualitative, not quantitative

$$1 + \frac{2b}{3} \simeq 1.67 \qquad \neq \qquad \frac{3}{2}$$

- Model too simple, still

- Kagan: Gutenberg-Richter exponent should be 3/2 (i.e., $b-$value $= 0.75$)

  Instrumental artifacts makes the exponent increase Kagan, *Tectonophys* 2010

Consequences for predictability

- Consider $\langle N_{t+1}|N_t\rangle$ with N_t known, then

$$\langle N_{t+1}|N_t\rangle = \langle K\rangle N_t$$

  using $N_{t+1} = \sum_{i=1}^{N_t} K_i$

  For critical branching processes $\langle K\rangle = 1$ and then

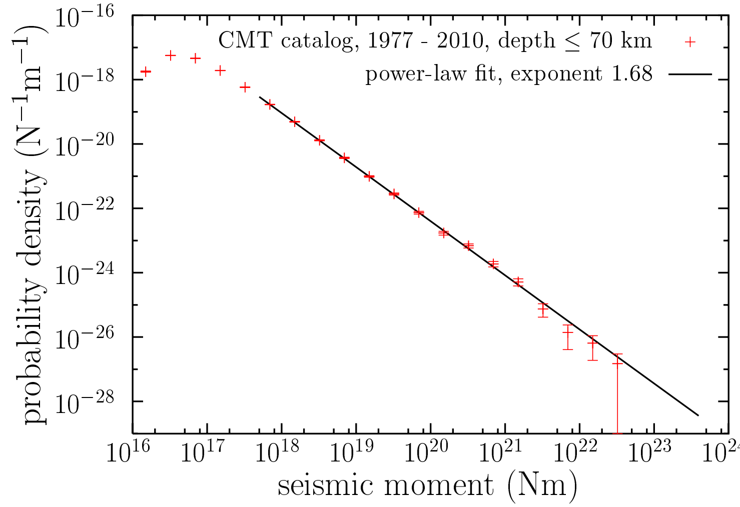
$$\langle N_{t+1}|N_t\rangle = N_t$$

  Note that it is not only that the outcome of the next step is random

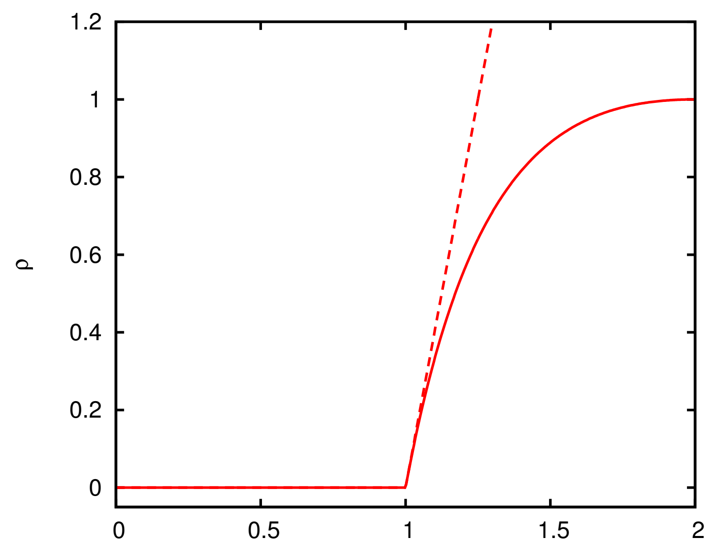
  It is much worst: the earthquake is in the limit of attenuation and intensification

- But what makes earthquakes critical?

Summary



- The size (energy) of earthquakes (and other natural hazards) follows a power-law distribution
- A power law signals the absence of a characteristic scale
- (Decreasing) power-law densities, with $\beta \leq 2$ have an infinite mean value
- Galton-Watson branching process can be a model of earthquakes
  - ⋆ Continuous phase transition at $\langle K \rangle = 1$
  - ⋆ Size distribution is only power law at the critical point



Self-Organized Branching Process

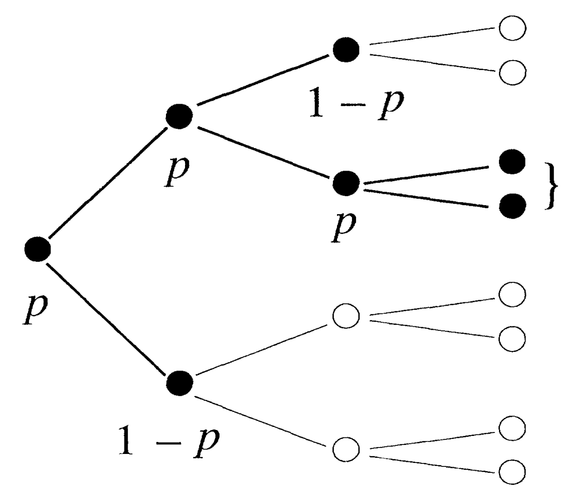
Zapperi et al. *Phys Rev Lett* 1995

- Consider: 0 offsprings with prob $1-p$
  2 offsprings with prob p

  Limit the maximum number of generations
  $\Rightarrow$ analogous to introduce a boundary at $t=L$

  Change p from one realization T to the next as

$$p(T+1)=p(T)+\frac{1-N_L(T)}{M}$$

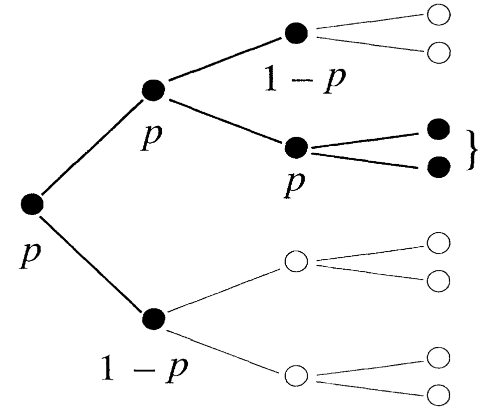


where N_L is the population in the last generation ($=2$ in Fig.)

and M is a big number (explained later)

Note that there are 2 times scales

⋆ $t=$ fast time scale, counts generations ,from $t=0$ to L
⋆ $T=$ slow time scale, counts realizations



$$p(T+1) = p(T) + \frac{1 - N_L(T)}{M}$$

- Dynamics

  ⋆ If p is low $\Rightarrow$ small size $\Rightarrow$ $N_L = 0$ $\Rightarrow$ p increases
  ⋆ If p is high $\Rightarrow$ large size $\Rightarrow$ $N_L > 1$ $\Rightarrow$ p decreases

- Indeed, we know that $\langle N_L \rangle = \langle K \rangle^L = (2p)^L$

  So, we can write, $N_L = (2p)^L + \eta$, with $\langle \eta \rangle = 0$

  Considering the deterministic equation (removing η)

$$p(T+1) = F(p(T)) = p(T) + \frac{1 - (2p(T))^L}{M}$$

  Therefore, the deterministic equation has a fixed point $p^* = 1/2 = p_c$

- Moreover, if M is big enough then $|F'(p^*)| < 1$ and the fixed point is attractive,so

$$p(T) \to p^* = p_c$$

As the noise is small, it only adds small perturbations to p^*

Then, p tends, or self-organizes, to its critical value, on average

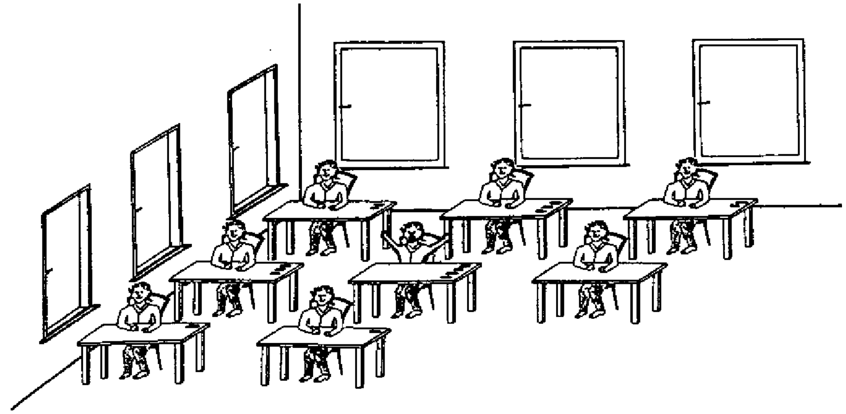
- Note:
Self-organization is the spontaneous emergence of structures or global order (here we do not have any structure yet, but wait...)

Examples:
convection patterns in fluids, chemical oscillations, self-regulations of markets

- Nevertheless, the global condition (on p) is very difficult to justify, in practice

4. Self-organization towards criticality

Cellular automaton Manna model[2]



Bak 1996, after Grassberger

- Let us consider a lattice in d dimensions
  - ⋆ Each site can store only 1 particle (or 0)
  - ⋆ If extra particles arrive at a site:
    - ⇒ 2 of them are transferred to 2 randomly chosen sites among its neighbors (this may generate an avalanche)
  - ⋆ Particles leave the system through the (open) boundary
  - ⋆ If there is no activity (all sites with 1 particle or less):
    - ⇒ Add 1 particle to a random site

In a formula, with $nn(j)$ denoting 2 random neighbors of j

$$\text{if } z_j \geq 2 \quad \Rightarrow \quad \begin{cases} z_j & \rightarrow & z_j & - & 2 \\ z_{nn(j)} & \rightarrow & z_{nn(j)} & + & 1 \end{cases}$$

$$\text{if } z_k < 2 \ \forall k \quad \Rightarrow \quad z_n \rightarrow z_n + 1 \ \text{ with } n = rand$$

[2] Cellular automaton = dynamical system with discrete time, space, and variable (field)

- The Manna model defines a complex system:

  System composed of many interacting parts, such that the collective behavior of those parts together is more than the sum of their individual behaviors

  Other examples, more complex: the cell, the brain, ecosystems, the economy, the Earth's crust... Newman, *Am J Phys* 2011

- Let us go back to the Manna model in the limit of infinite dimensions, $d \to \infty$

  Then, the propagation of the activity will show no loops $\simeq$ mean field (a neighbor will not be selected twice to get a grain $\Rightarrow$ no overlap)

  So, there will be no spatial correlations, and all sites are equivalent (the boundary conditions need to be readjusted)

  Each site will become active $(z \geq 2)$ with the same probability

$$p = \text{fraction of sites with one particle} = P[z = 1]$$

- Then, the activity propagates through the system as a branching process
- The offspring distribution will be binomial, with $n = 2$ and parameter p

  But note that there is no pre-existing tree
- The total number of particles will evolve as

$$\text{mass}(T+1) = \text{mass}(T) + 1 - \text{out}(T)$$

(one particle added before the avalanche, "out" particles lost at the boundaries)

Dividing by the total number of sites M, with $p = P[z = 1] = \text{mass}/M$

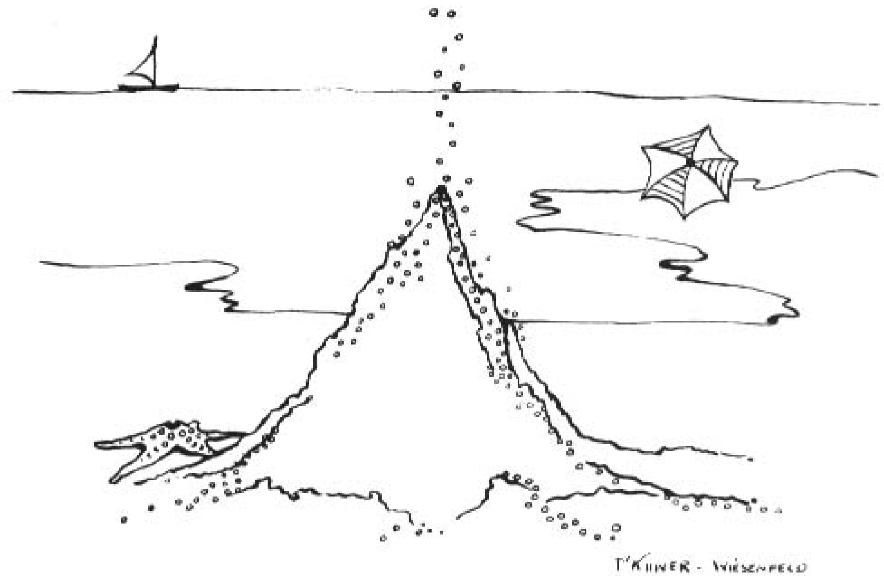
$$p(T+1) = p(T) + \frac{1 - \text{out}(T)}{M}$$

which corresponds to the self-organized branching process Christensen & Moloney 2005

The evolution and adjustment of p is implemented in a natural way

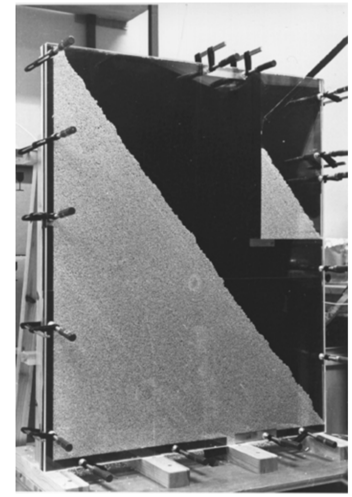
Self-Organized Criticality (SOC) Bak et al. *Phys Rev Lett* 1987

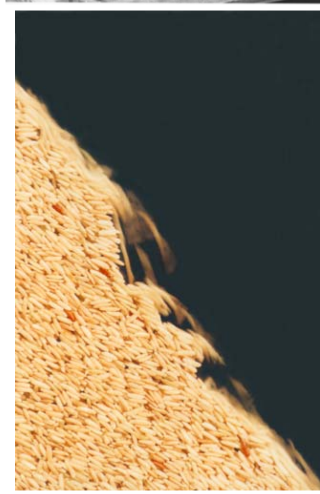
- The dynamics arises from the sandpile metaphor
  - ⋆ If there are few grains (flat pile)
    $\Rightarrow$ small avalanches, pile grows
  - ⋆ If there are many grains (steep pile)
    $\Rightarrow$ large avalanches, pile decreases
    (through boundary dissipation)



This mechanism makes the slope of the pile fluctuate around the critical state

$\Rightarrow$ Bak-Tang-Wiesenfeld (BTW) model





- BTW model: one-dimensional lattice, $d = 1$, with $j = 1 \ldots L$

  Modification: no random selection of neighbors

$$
\begin{array}{lll}
\text{if } z_j \geq 2 & \Rightarrow & \left\{ \begin{array}{lllll} z_j & \to & z_j & - & 2 \quad \text{for } j \neq L \\ z_{j\pm 1} & \to & z_{j\pm 1} & + & 1 \end{array} \right. \\
\text{if } z_k < 2 \ \forall k & \Rightarrow & z_n \to z_n + 1 \quad \text{with } n = rand
\end{array}
$$

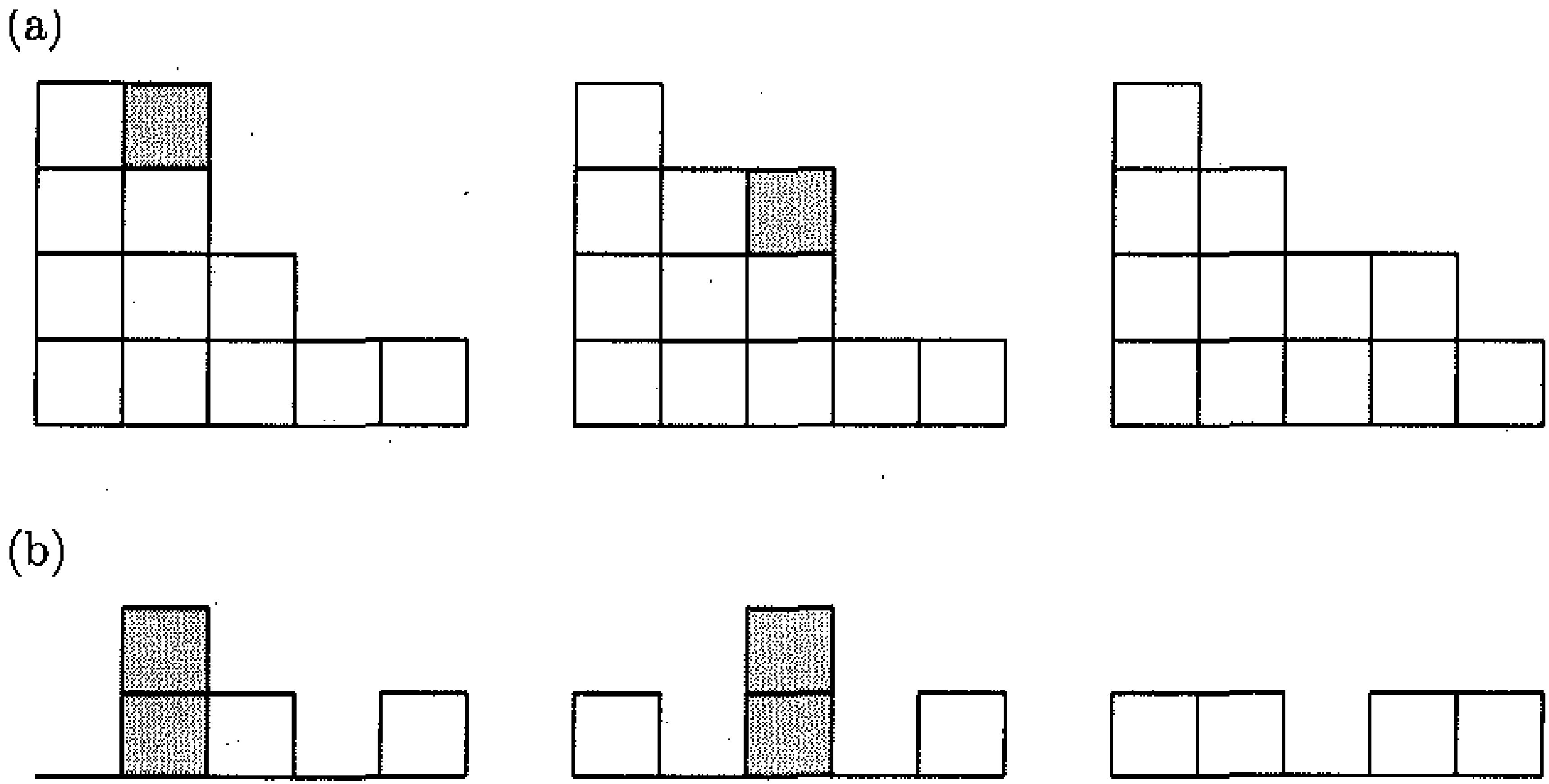
The "particles" are in fact elements of slope in a $2-d$ sandpile

$$\text{height at } j = h_j = h_{j+1} + z_j \quad \Rightarrow \quad z_j = h_j - h_{j+1}$$

with $h_{L+1} = 0 \Rightarrow z_L = h_L \Rightarrow z_L \to z_L - 1$ (conserved BC)

$$
\begin{array}{lll}
\text{if } h_j - h_{j+1} \geq 2 & \Rightarrow & \left\{ \begin{array}{lllll} h_j & \to & h_j & - & 1 \\ h_{j+1} & \to & h_{j+1} & + & 1 \end{array} \right. \\
\text{if } h_k - h_{k+1} < 2 \ \forall k & \Rightarrow & h_m \to h_m + 1 \text{ for } m \leq n \text{ with } n = rand
\end{array}
$$

- Height h picture (grains) versus slope z picture (repelling particles)



Christensen & Moloney 2005

Relation with interface depinning

Paczuski & Boettcher, *Phys Rev Lett* 1996

- Define H_j as the total number of topplins in a sandpile

  When:

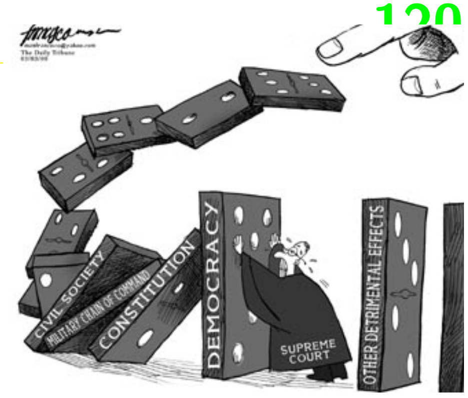
  ⋆ the initial condition is empty ($h_j = 0$ for all j and for $T = 0$) and
  ⋆ the addition takes place at $j = 1$

  then, H defines an advancing interface, whose gradient gives the pile height

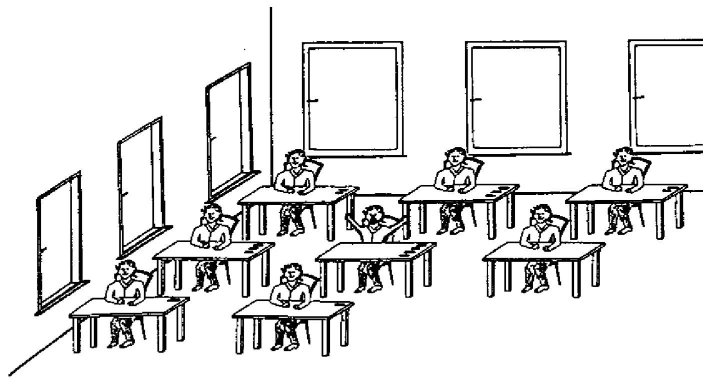
$$h_j = H_{j-1} - H_j$$

  with H_0 giving the total number of grains added

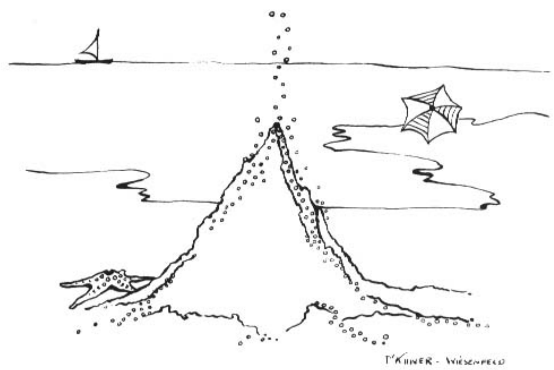
Retrospective of models



- Domino (Otsuka) model of fracture
- Galton-Watson branching process
- Self-organized branching model



- Cellular automaton Manna (bureaucrats) model
- Bak et al. sandpile model



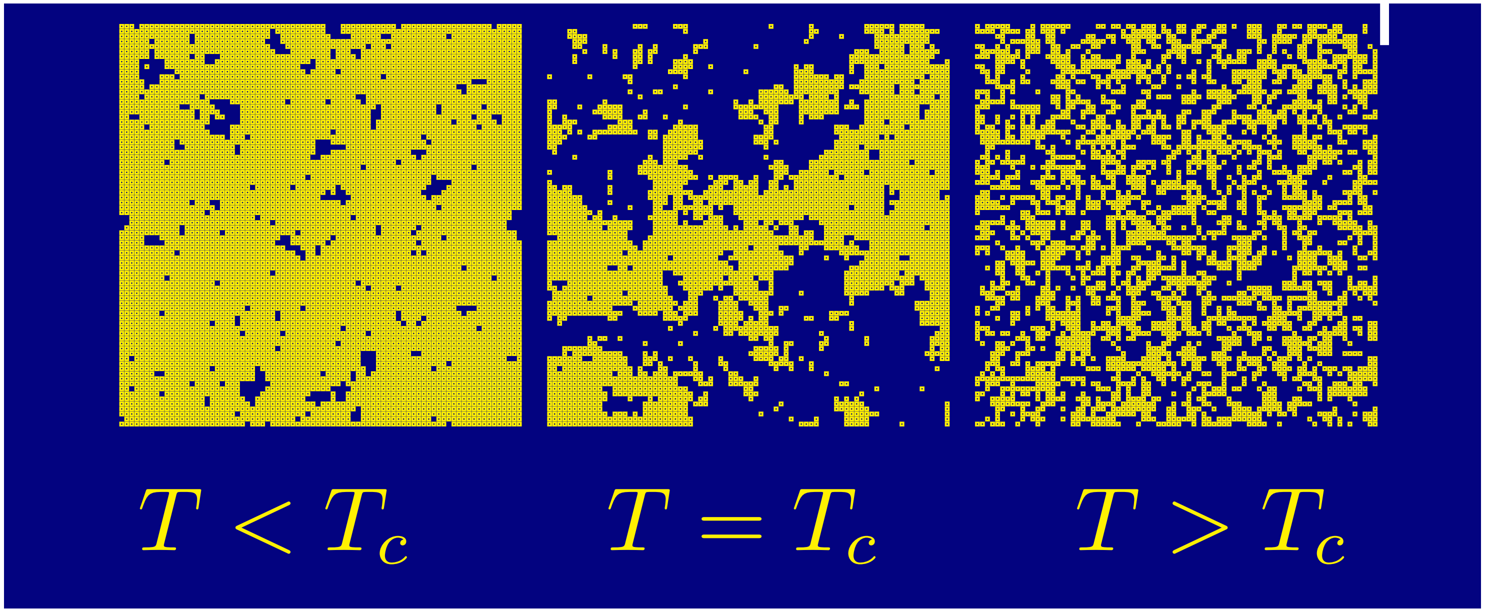
- Interface depinning model
- These models serve as metaphors for earthquakes

- **Inspiration: Critical Points of Thermodynamic Phase Transitions**

  Magnetic material: atom $=$ spin with 2 states

  There exists a critical temperature T_c

  ⋆ Above T_c: no magnetization, small clusters
  ⋆ Below T_c: magnetization, one very large cluster
  ⋆ At the precise value $T = T_c$ $\Rightarrow$ clusters of all sizes $\Rightarrow$ power law!



Christensen & Moloney, *Complexity and Criticality* 2005

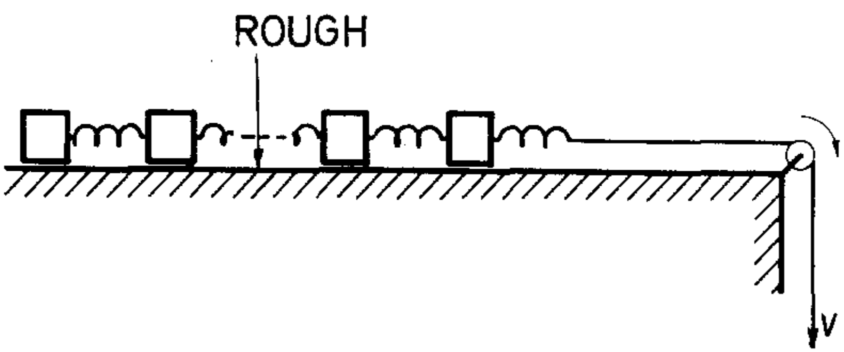
Burridge-Knopoff spring-block models

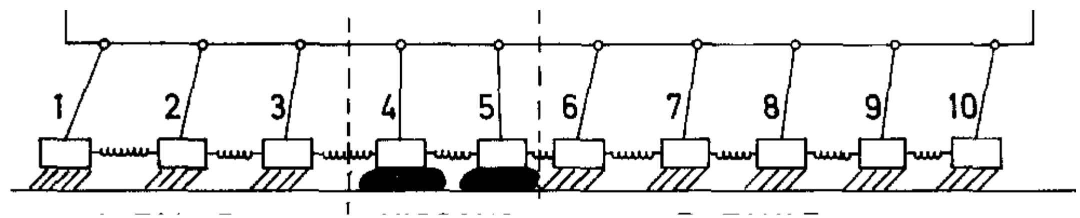
Bull Seism Soc Am 1967

- Earthquakes take place in "pre-existing" faults

⇒ Alternative: modeling friction in a fault

⋆ Experiment: spring-block system pulled from one end





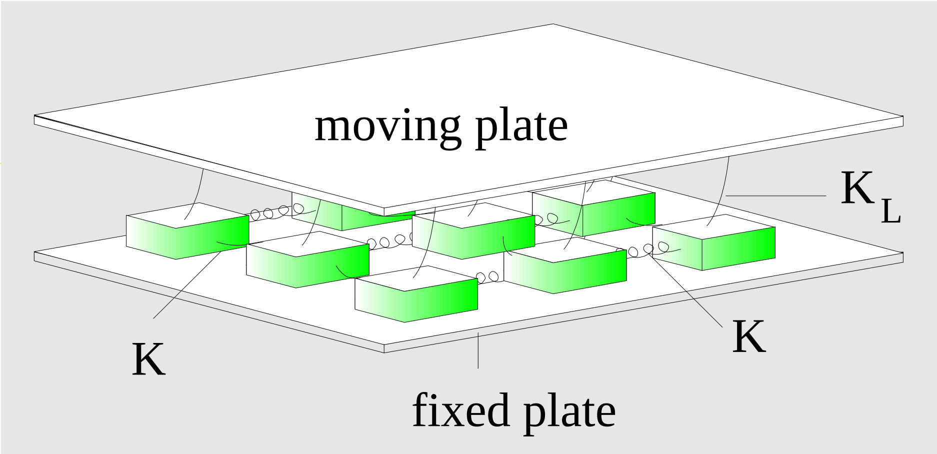
⋆ Computer simulations:
All blocks connected by flat springs to a moving plate

- stick-slip dynamics: slow driving (pull) + fast avalanches (shocks)
  - ⋆ The force on the block(s) increases (linearly) very slowly
  - ⋆ At some time (for some block) the force exceeds the static frictional force
  - ⋆ Then, that block moves fast, changing the force over the neighbor blocks and so on

  “Size” of the earthquake $\simeq$ number of sliding blocks

4. Self-organization towards criticality

Coupled-map lattice model

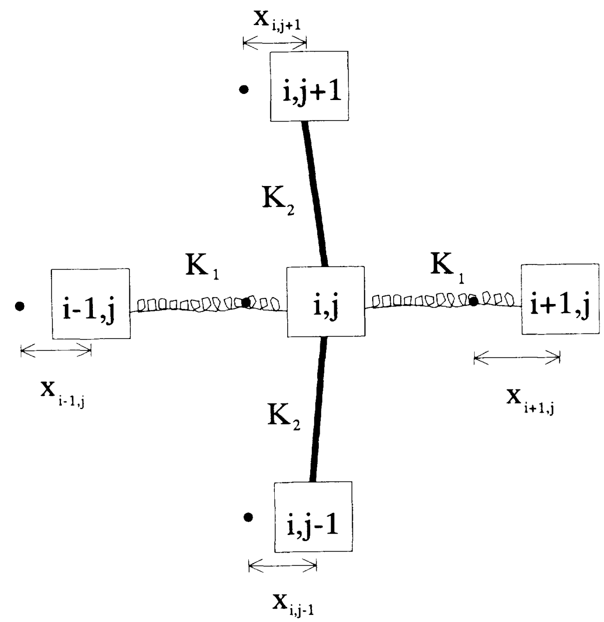


- Olami-Feder-Christensen (OFC) model, *Phys Rev Lett* 1992

  Two-dimensional version of Burridge-Knopoff model

  - ⋆ Coil (helical) springs connecting blocks in the direction of motion of the plate
  - ⋆ Flat (leaf) springs connecting blocks in the perpendicular direction (making the force then in the direction of motion also)
  - ⋆ In both cases the value of the elastic constants is K
  - ⋆ Flat springs connecting blocks with the upper moving plate with constants $K_L \neq K$

4. Self-organization towards criticality



- Let us define
  - ⋆ $F_{i,j} =$ Force on block i, j
  - ⋆ $x_{i,j} =$ Displacement in the direction of motion of i, j relative to the upper flat spring

  Also, the zero force between each pair corresponds to the lattice of upper springs. By Hooke's law $F_{i,j} =$

$$= -K(x_{i,j} - x_{i-1,j}) - K(x_{i,j} - x_{i+1,j}) - K(x_{i,j} - x_{i,j-1}) - K(x_{i,j} - x_{i,j+1}) - K_L x_{i,j}$$

$$= K \left(\sum_{nn(i,j)} x_{nn(i,j)} - 4x_{i,j} \right) - K_L x_{i,j}$$

If the upper plate moves with constant (small) velocity v then

$$\frac{dF_{k,l}}{dt} = -K_L \frac{dx_{k,l}}{dt} = K_L v \quad \text{for all } k, l$$

- When the force on some block i, j reaches the frictional threshold force F_{th} $\Rightarrow$ block i, j slips instantaneously to the position with of zero force, so

$$F_{i,j} \rightarrow 0 \qquad \text{(assumption of the model)}$$

Then, if we denote the new position of i, j as $x'_{i,j}$

$$0 = K \left(\sum_{nn(i,j)} x_{nn(i,j)} - 4x'_{i,j} \right) - K_L x'_{i,j}$$

where $nn(i, j)$ denotes the nearest neighbors of i, j. Substracting,

$$F_{i,j} - 0 = -(4K + K_L)(x_{i,j} - x'_{i,j})$$

- Therefore, the force on the $i+1, j$ neighbor (for instance)

$$F_{i+1,j} = K \left(\sum_{nn(i+1,j)} x_{nn(i+1,j)} - 4x_{i+1,j} \right) - K_L x_{i+1,j}$$

So, as $F_{i,j} = -(4K + K_L)(x_{i,j} - x'_{i,j})$ then $F_{i+1,j}$ changes to

$$F_{i+1,j} \rightarrow F_{i+1,j} + K(x'_{i,j} - x_{i,j}) = F_{i+1,j} + \frac{K}{4K + K_L} F_{i,j}$$

and the model is non-conservative, as $\alpha = K/(4K + K_L) < 0.25$ except if $K_L \rightarrow 0$

Summary of the rules of the OFC model

if $F_{i,j} < F_{th}$ for all i, j $\Rightarrow$ $dF_{i,j}/dt = K_L v$ with v very small

if $F_{i,j} \geq F_{th}$ for some i, j $\Rightarrow$ $\begin{cases} F_{nn(i,j)} & \rightarrow & F_{nn(i,j)} + \alpha F_{i,j} \\ F_{i,j} & \rightarrow & 0 \end{cases}$

The boundary conditions are disregarded

Note that there are 2 times scales:
The slow one is continuous, but the fast one is discrete

In practice, in simulations, don't use $dF_{i,j}/dt = K_L v$. Why?

Then, the slow time scale turns into discontinuous

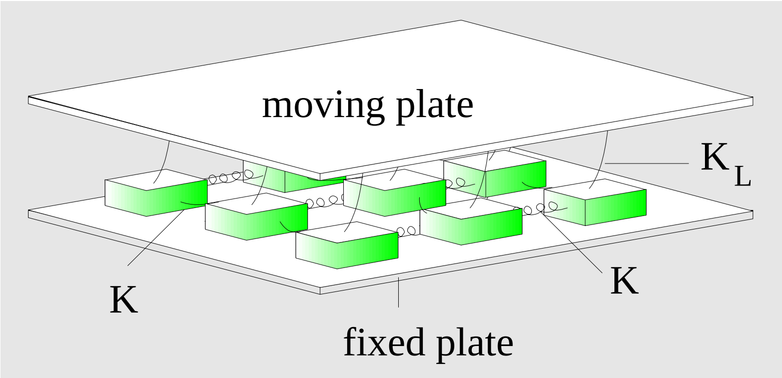
$\Rightarrow$ coupled map lattice model[3]

[3] Coupled map lattice = dynamical system on a lattice with continuous variables and discrete time

Earthquakes *can be* a SOC phenomenon

- Ingredients for SOC (and fulfillment in earthquakes)

Pruessner, private comm.



  - ⋆ Time scale separation ($\Rightarrow$ OK)
  - ⋆ Thresholds, interaction ($\Rightarrow$ OK)
  - ⋆ Avalanche dynamics ($\Rightarrow$ OK)
  - ⋆ Power-law distributions ($\Rightarrow$ OK)
    (with finite-size scaling)

  - ⋆ Underlying 2nd-order phase transition,
    reached by self-organization ($\Rightarrow$??)

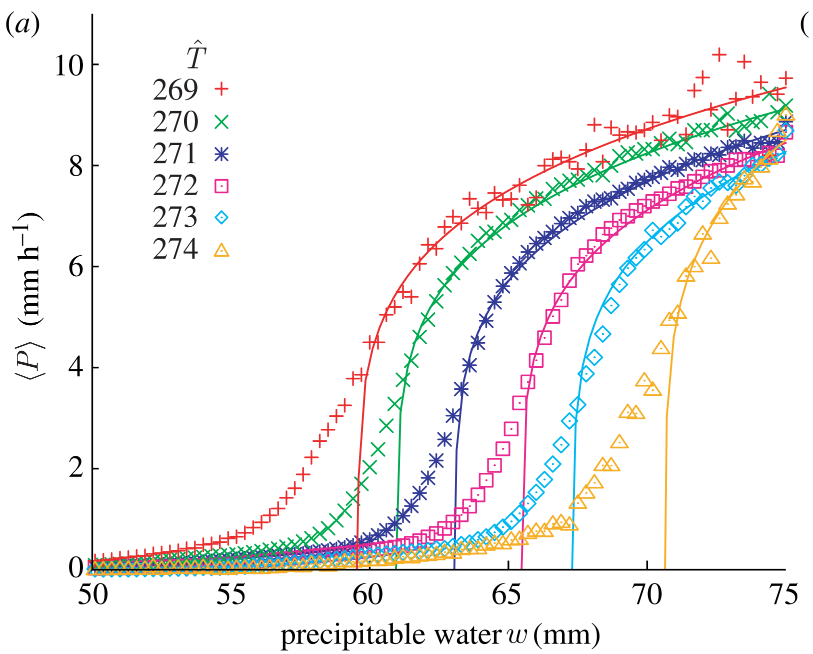
Think in the critical temperature of Fe, $T_c = 770°$C
or in the critical point of water, at $T_c = 374°$C and 218 atm

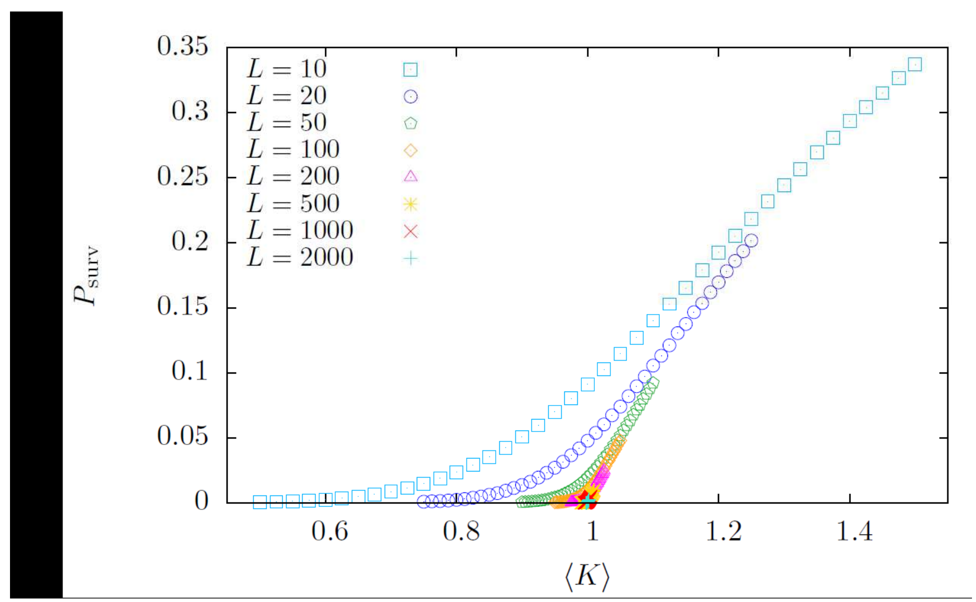
Andrews 1869 (for CO_2)

Other candidates for SOC

- For rain, Peters and Neelin have shown: *Nature Phys* 2006, Neelin et al. *Phil Trans R Soc A* 2008

  1. Existence in the atmosphere of a non-equilibrium stability-instability transition





Stanley, *Rev Mod Phys* 1999

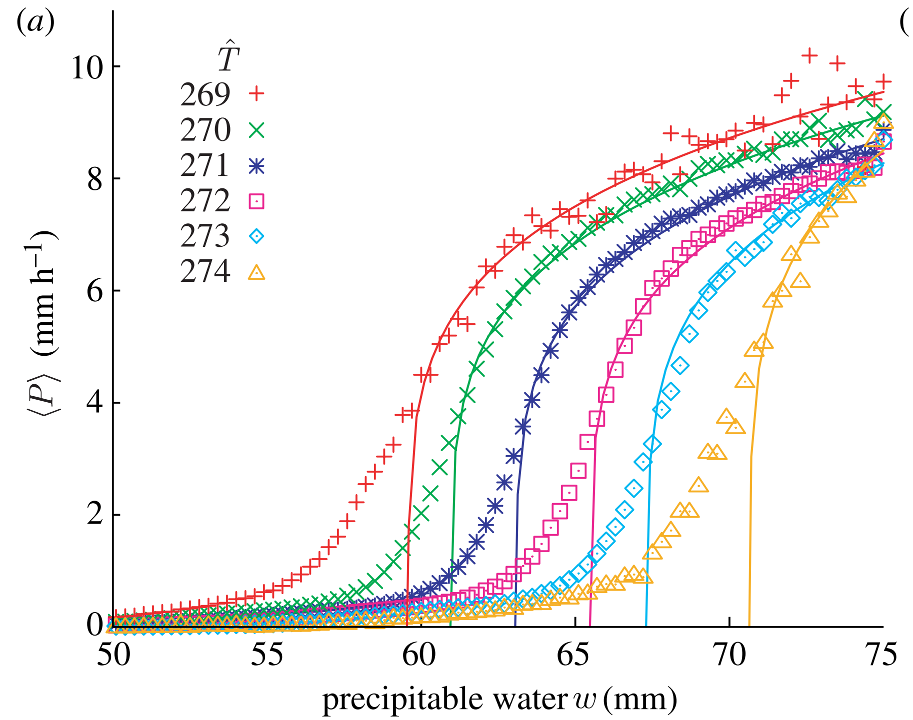
- Finite size effects

Finite size scaling: $\langle P\rangle = L^{-0.2/\nu} H[(w-w_c)L^{1/\nu}]$ (L system size)

$$H(x) \propto \begin{cases} |x|^{-\gamma} & \text{for } x\to -\infty \\ x^{0.2} & \text{for } x\to +\infty \end{cases} \underset{L\to\infty}{\Rightarrow} \langle P\rangle \propto \begin{cases} 0 & \text{for } w<w_c \\ (w-w_c)^{0.2} & \text{for } w>w_c \end{cases}$$

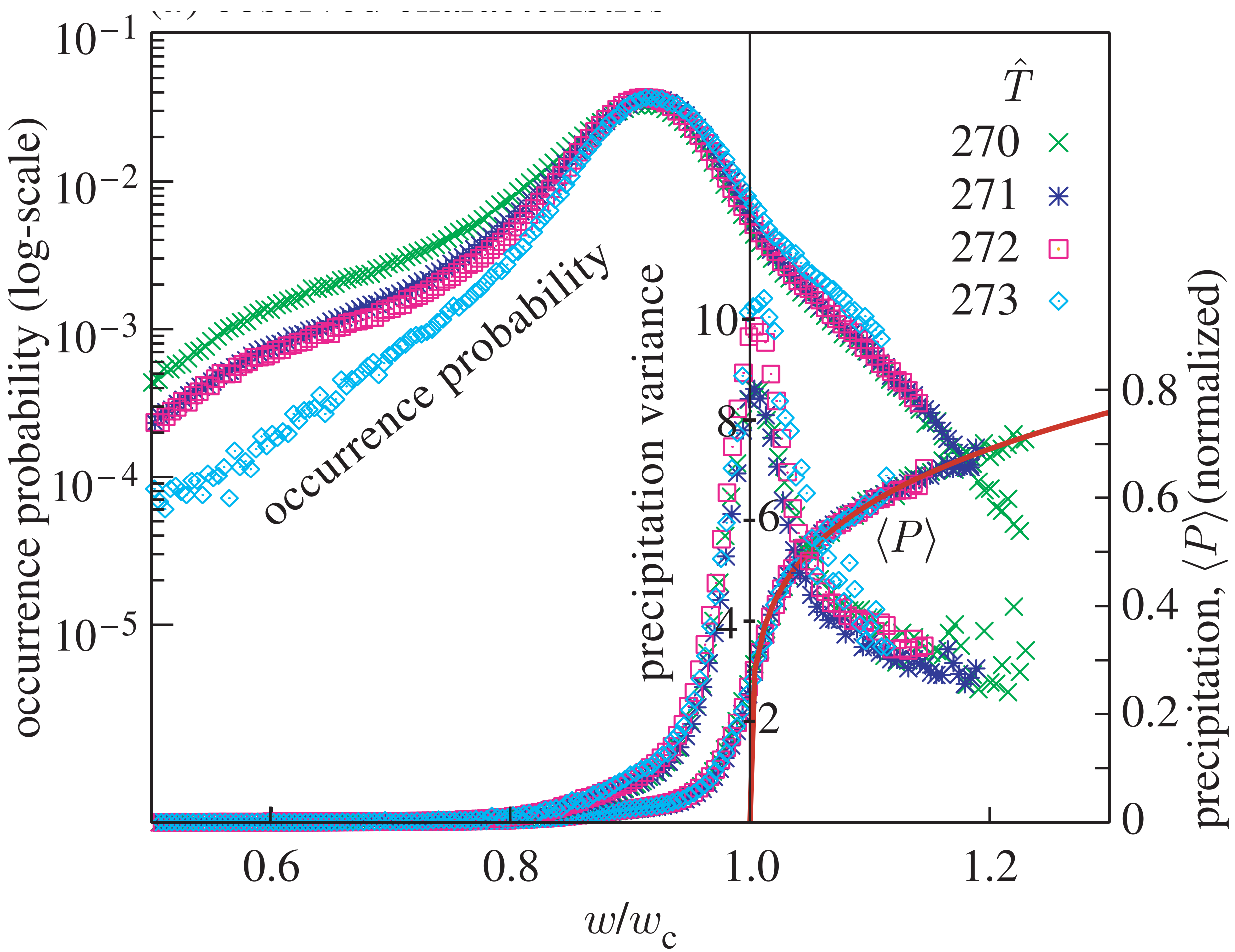
With critical point $w_c \simeq 63$ mm if $T = 271$ K, and so on

Phase transitions (abrupt changes) only exist in the limit $L\to\infty$



- Peters and Neelin have also shown: *Nature Phys* 2006, Neelin et al. *Phil Trans R Soc A* 2008

  2. The atmosphere is attracted towards the critical point of the transition



Continuous, non-upper-truncated power-law distributions

- Given by a probability density $D(x)$ with x real (continuous)

$$D(x) = \frac{B}{x^\gamma} \qquad \text{for } a \leq x < \infty$$

  with $a > 0$. Then, $\int_a^\infty D(x)dx = 1$ with $\gamma > 1$ implies

$$B = (\gamma - 1)a^{\gamma - 1}$$

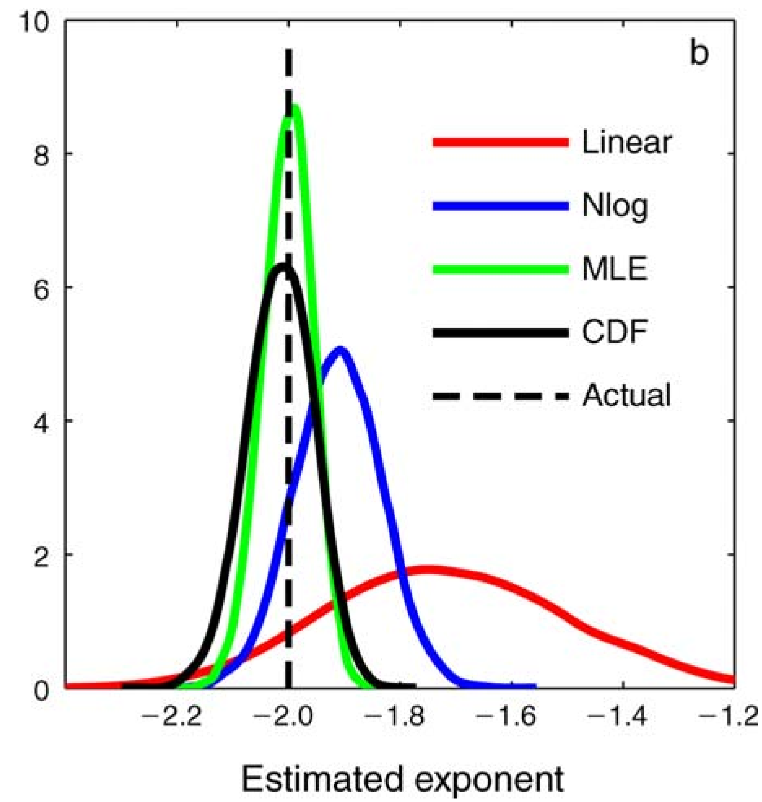
- In order to decide between competing explanations, universality classes, etc., it is important not only to determine if power laws hold, but also the precise value of the exponent γ

Fitting power-law distributions

- Most common method to fit power laws has been linear regression in log-log (needs to estimate first the empirical density $\Rightarrow$ delicate)

- Some authors have pointed out the superiority of maximum-likelihood (ML) estimation

  Goldstein et al. *Eur Phys J B* 2004
  Bauke, *Eur Phys J B* 2007
  White et al. *Ecol* 2008
  Clauset et al. *SIAM Rev* 2009



- ML estimators are: asymptotically unbiased and with lowest variance

  Invariant under re-parameterizations

Maximum likelihood (ML) estimation

- Given a dataset of size N, $x_1, x_2, \dots x_N$, the likelihood is the joint distribution

$$\mathcal{L}(\gamma) = D(x_1, x_2, \dots x_N; \gamma) = \prod_{i=1}^{N} D(x_i; \gamma)$$

(assuming independence). For a power law, $D(x) = B/x^{\gamma}$, the log-likelihood is

$$\ell(\gamma) = \frac{\ln \mathcal{L}(\gamma)}{N} = \frac{1}{N} \sum_{i=1}^{N} \ln D(x_i; \gamma) = \ln B - \frac{\gamma}{N} \sum_{i=1}^{N} \ln x_i = \ln B - \gamma \ln G$$

with G the geometric mean of the data and $B = (\gamma - 1)a^{\gamma - 1}$ (continuous case).
Maximizing

$$\frac{d\ell}{d\gamma} = 0 \qquad \Rightarrow \qquad \gamma = 1 + \frac{1}{\ln \frac{G}{a}}$$

- Note that the ML solution depends only on the geometric mean G

$$\gamma = 1 + \left(\ln \frac{G}{a}\right)^{-1}$$

So, any data, from any distribution, with the same G yields the same γ

Then, maximum likelihood should be called minimum unlikelihood

$\Rightarrow$ A goodness-of-fit test is necessary

Goodness-of-fit test

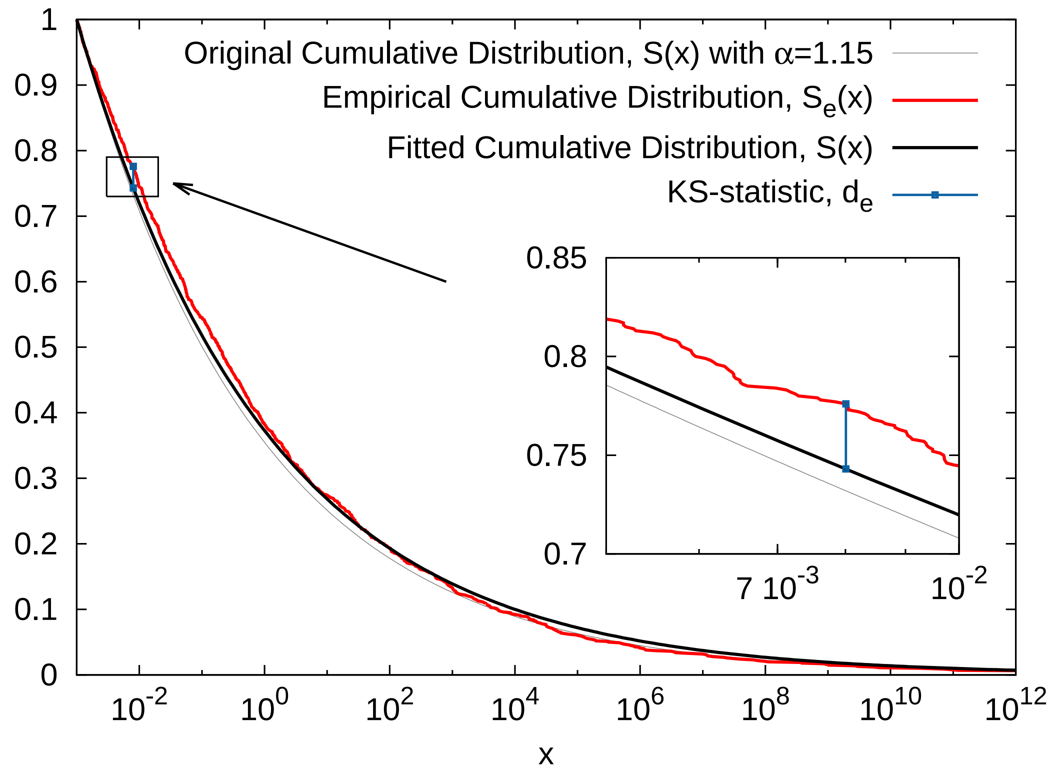
- In order to test the goodness of the (ML) fit let us consider Kolmogorov-Smirnov

  KS distance or KS statistic $=$ maximum difference for all x between empirical $S(x)$ and theoretical $S(x)$

  $d =$ max$\{S_{emp}(x) - S_{th}(x)\}$

  with $S(x) = \int_x^\infty D(x)dx$

  Care with $p-$value: Monte Carlo simulations



- The problem of power-law fitting is not in fitting the power law

  It is in the selection of $a =$ minimum value of x

Clauset's et al. recipe

Clauset, Shalizi & Newman *SIAM Rev* 2009

- ⋆ Take an arbitrary value of a ($=$ minimum x for which the power law holds)
  ⋆ Calculate fit by ML estimation $\Rightarrow$ yields exponent γ
  ⋆ Calculate Kolmogorov-Smirnov distance between empirical distribution and fit (no goodness-of-fit yet)

- ⋆ Select the value of a which minimizes Kolmogorov-Smirnov distance $d = d_{emp}$
  So, we come out with a fit given by a_{emp} and γ_{emp}

- ⋆ Calculate $p-$value of the fit by simulating N_{sim} synthetic samples:
  - $*$ Simulating a power law with exponent γ_{emp} for $x \geq a_{emp}$
  - $*$ Bootstrap of the empirical distribution for $x < a_{emp}$

  ⋆ Proceed with synthetic samples in exactly the same way as with the empirical

$\Rightarrow$ Each synthetic sample yields a value of d_{sim}

$\Rightarrow$ Calculate $p-$value as $\quad p \simeq \{\text{number of } d_{sim} > d_{emp}\}/N_{sim}$

- Justification of the minimization of d

  Under the null hypothesis, Kolmogorov-Smirnov distance goes as

  $$d \propto \frac{1}{\sqrt{N}}$$

  So, under the null hypothesis, the smaller a, the larger N and the smaller d

  But as soon as the null hypothesis fails, the fit deviates and d increases

  A sort of balance between the two effects is implicit

  Nevertheless, there is no reason why this deviation should compensate and overcome the reduction in d

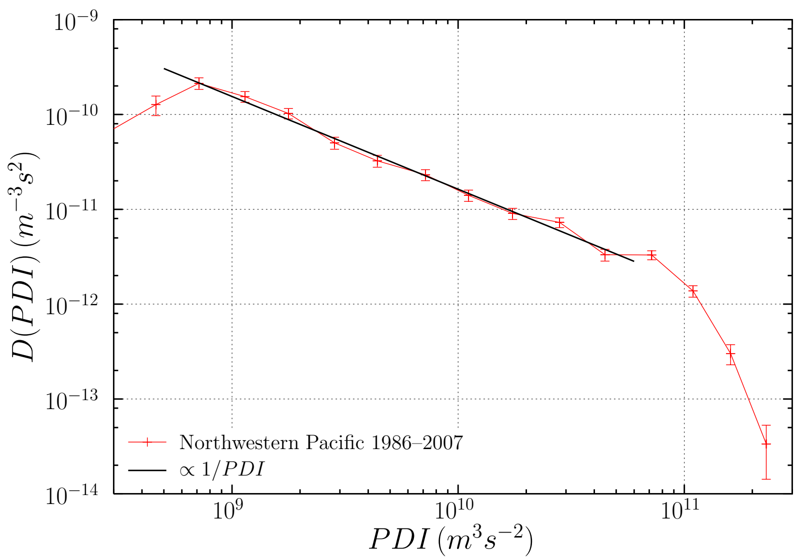
  (it would depend on the shape of the distribution for $x < a$)

Problems of Clauset et al.'s recipe

- The method performs bad when generalized to truncated power laws

$$D(x) \propto \frac{1}{x^\gamma} \qquad \text{for } a \le x \le b$$

This is common in complex systems, due to finite-size effects

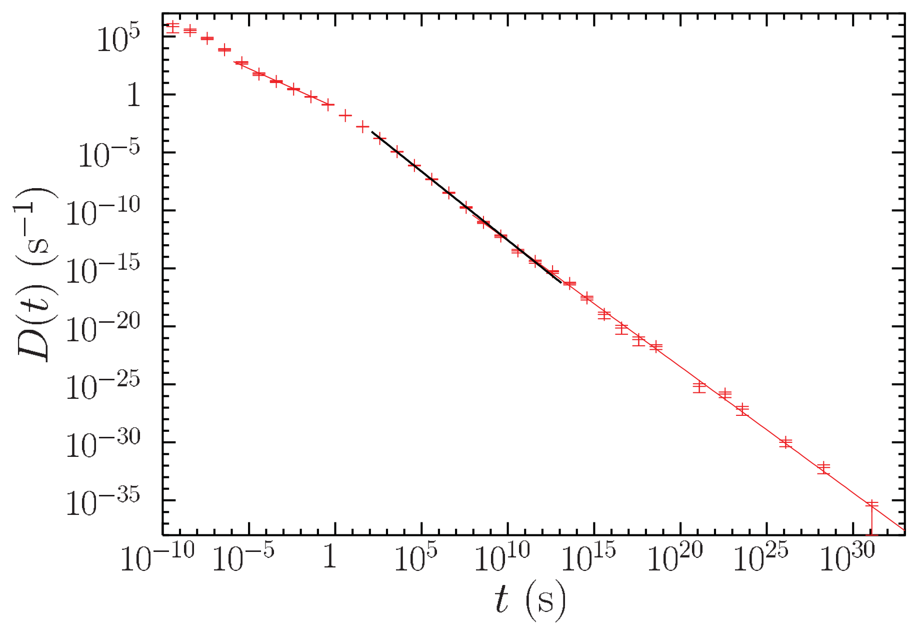


More problems of Clauset et al.'s recipe

A.C., Font & Camacho *Phys Rev E* 2011

- Consider nuclear half-lives: from below 10^{-16} s to 10^{23} yr $\sim$ 10^{31} s for ^{128}Te

  Clauset et al.'s recipe yields
  $a_{emp} = 30$ s and $\gamma_{emp} = 1.16$
  but $p = 0$

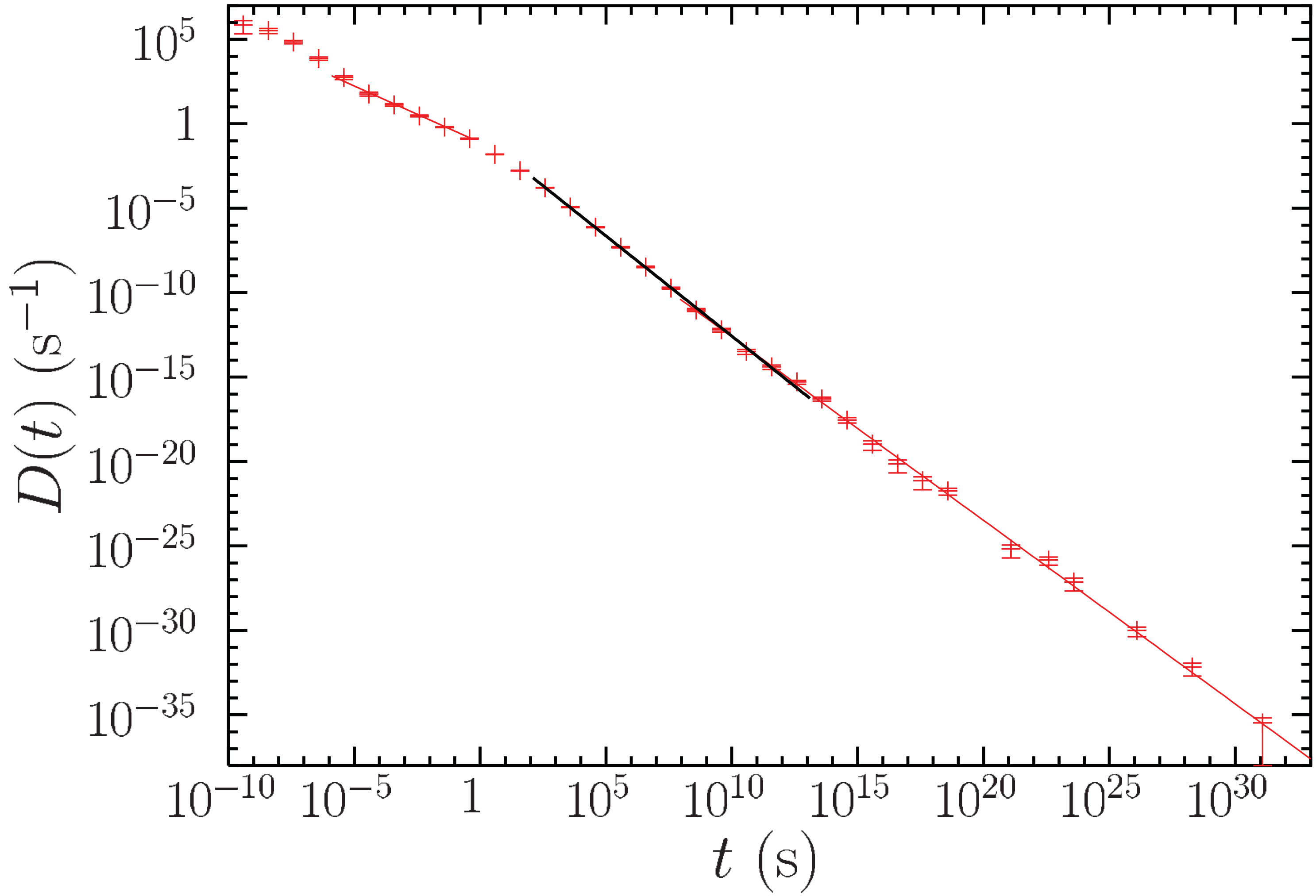


- Simulate power law for $x \geq 10^8$ s and bootstrap original data for $x < 10^8$ s

  In 80 % of the cases $p = 0$

  So, the recipe leads to the usual rejection of the power-law hypothesis when it is true!

  (p should be uniformly distributed between 0 and 1 under the null hypothesis)



⇒ Failure of the Clauset's et al. recipe

Alternative recipe

Peters et al. *J Stat Mech* 2010; Deluca & A.C. *Acta Geophys* 2013

- ⋆ Take an arbitrary value of a ($=$ minimum x for which the power law holds)
  ⋆ Calculate fit by ML estimation $\Rightarrow$ yields exponent γ
  ⋆ Calculate KS distance d between empirical distribution and fit
  (no difference with Clauset et al. yet)

- ⋆ Calculate a $p-$value for fixed a
  ∗ Simulate N_{sim} power-law synthetic samples with γ for $x \geq a$
  ⋆ Proceed with synthetic samples in exactly the same way as with the empirical

  $\Rightarrow$ Each synthetic sample yields a value of d_{sim}

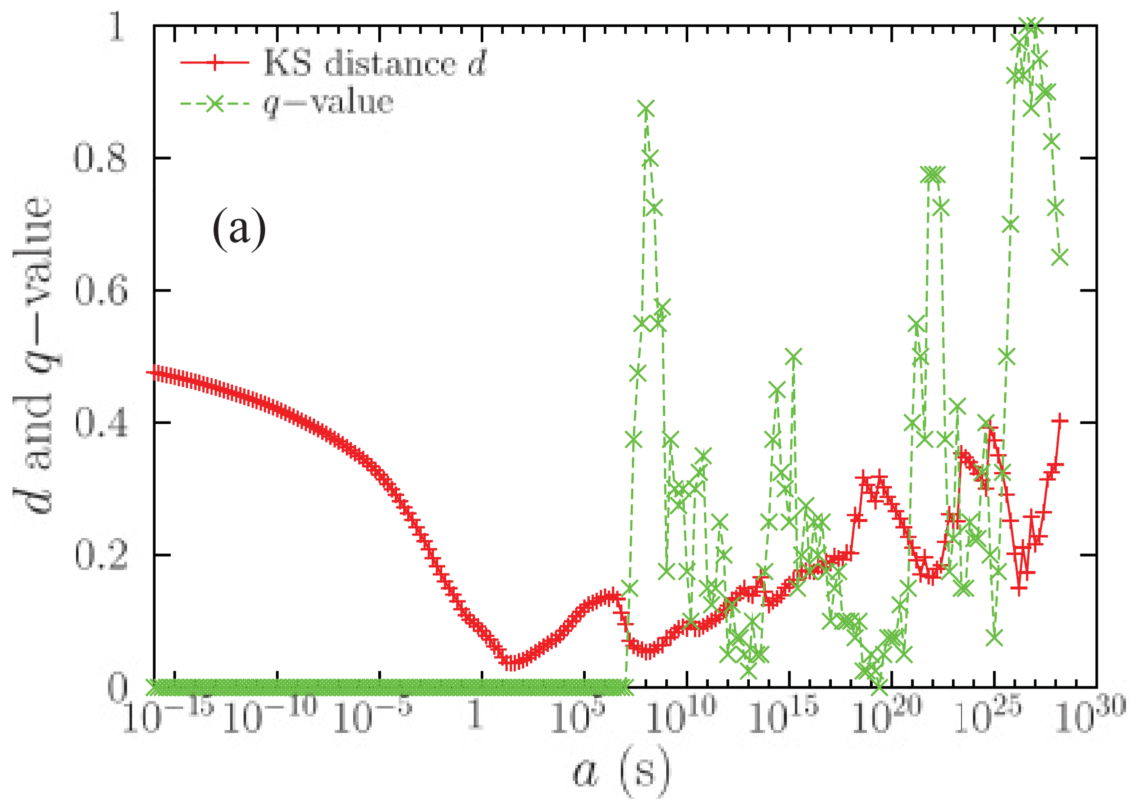
  ⋆ Calculate p as

$$p \simeq \frac{\text{number of } d_{sim} > d}{N_{sim}}$$

- Select the smallest value of a provided that $p > 0.20$ (e.g.)

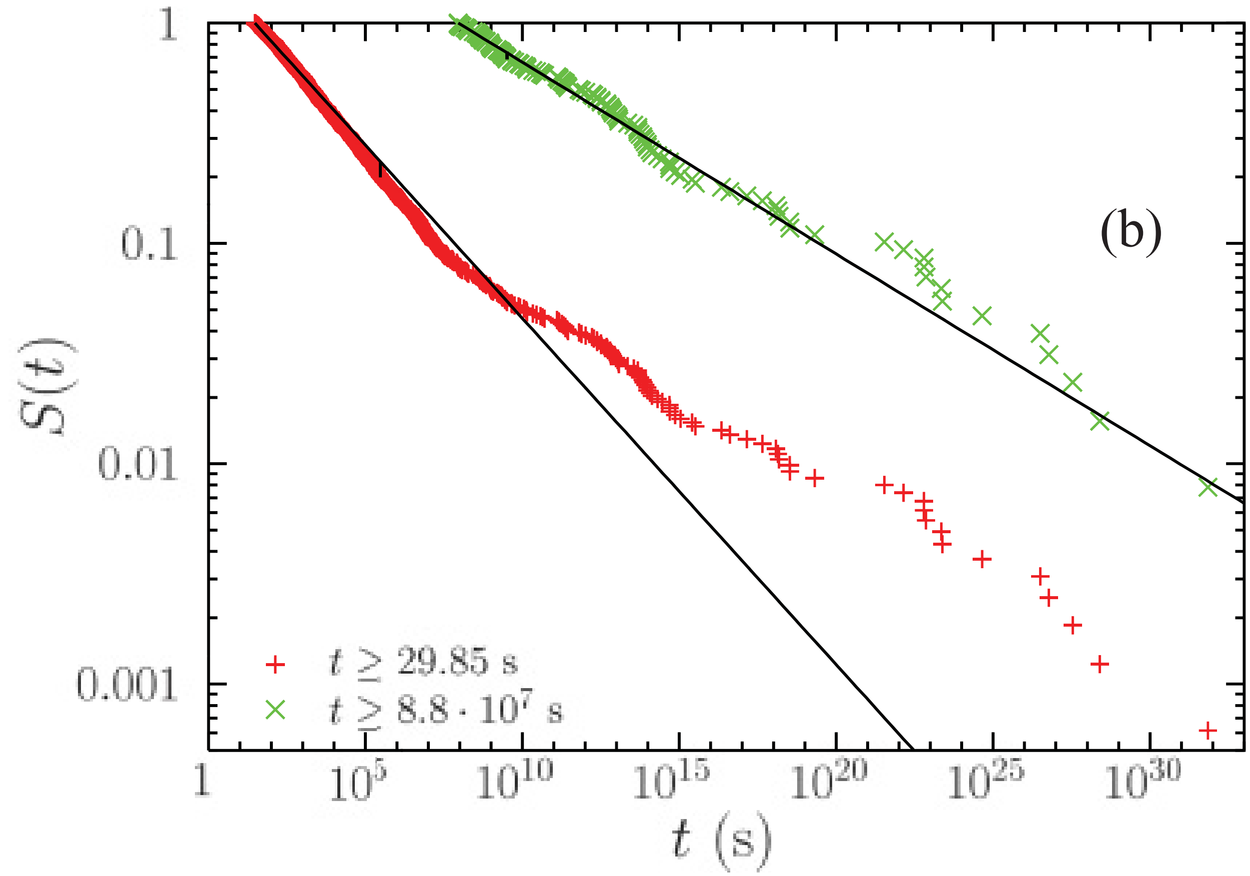
Performance

- Consider again nuclear half-lives: d and $p(=q)$ versus a

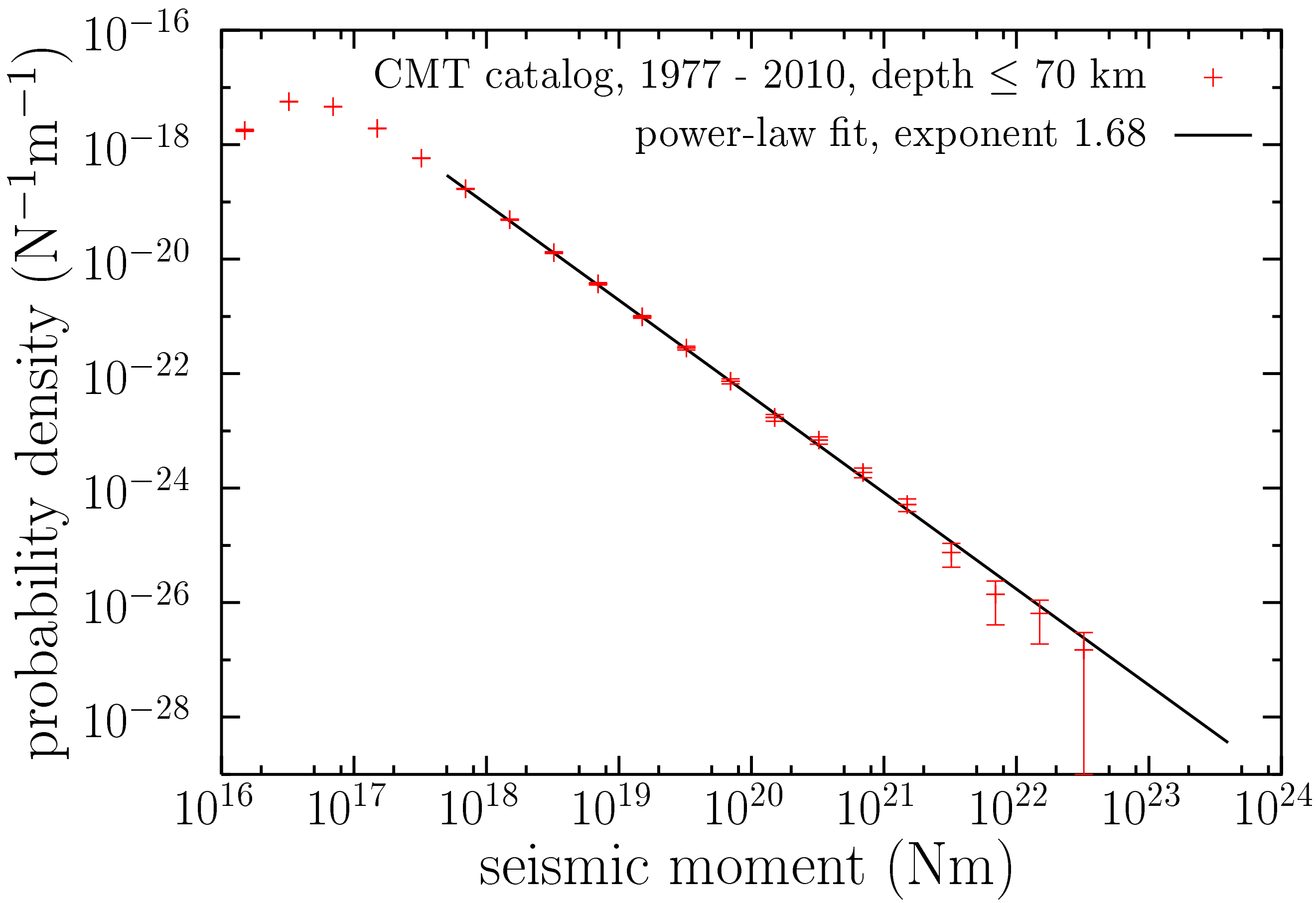


We obtain $a_{emp} = 3 \times 10^7$ s (~ 1 yr) and $\gamma_{emp} = 1.09$ (with $p > 0.20$)

- Comparison between Clauset et al.'s solution (red) (power-law rejected) and alternative (green) (for the (complementary) cumulative distribution)



- Global earthquakes revisited: power law cannot be rejected



But other fits are possible! $\Rightarrow$ LRT, or AIC, or BIC... Main et al. *Nature Geosci* 2008

(Upper) Truncated power laws

- Deviations from a power law arise for large x, due to finite size effects. So,

$$D(x) = \frac{B}{x^\gamma} \quad \text{with } a \le x \le b \quad \text{and } B = \frac{(\gamma-1)a^{\gamma-1}}{1-(a/b)^{\gamma-1}}$$

Be careful: b is not $b-$value. The log-likelihood is now

$$\ell(\gamma) = \ln B - \gamma \ln G = -\gamma \ln \frac{G}{a} - \ln a + \ln \frac{\gamma-1}{1-(a/b)^{\gamma-1}}$$

with G the geometric mean of data between a and b.

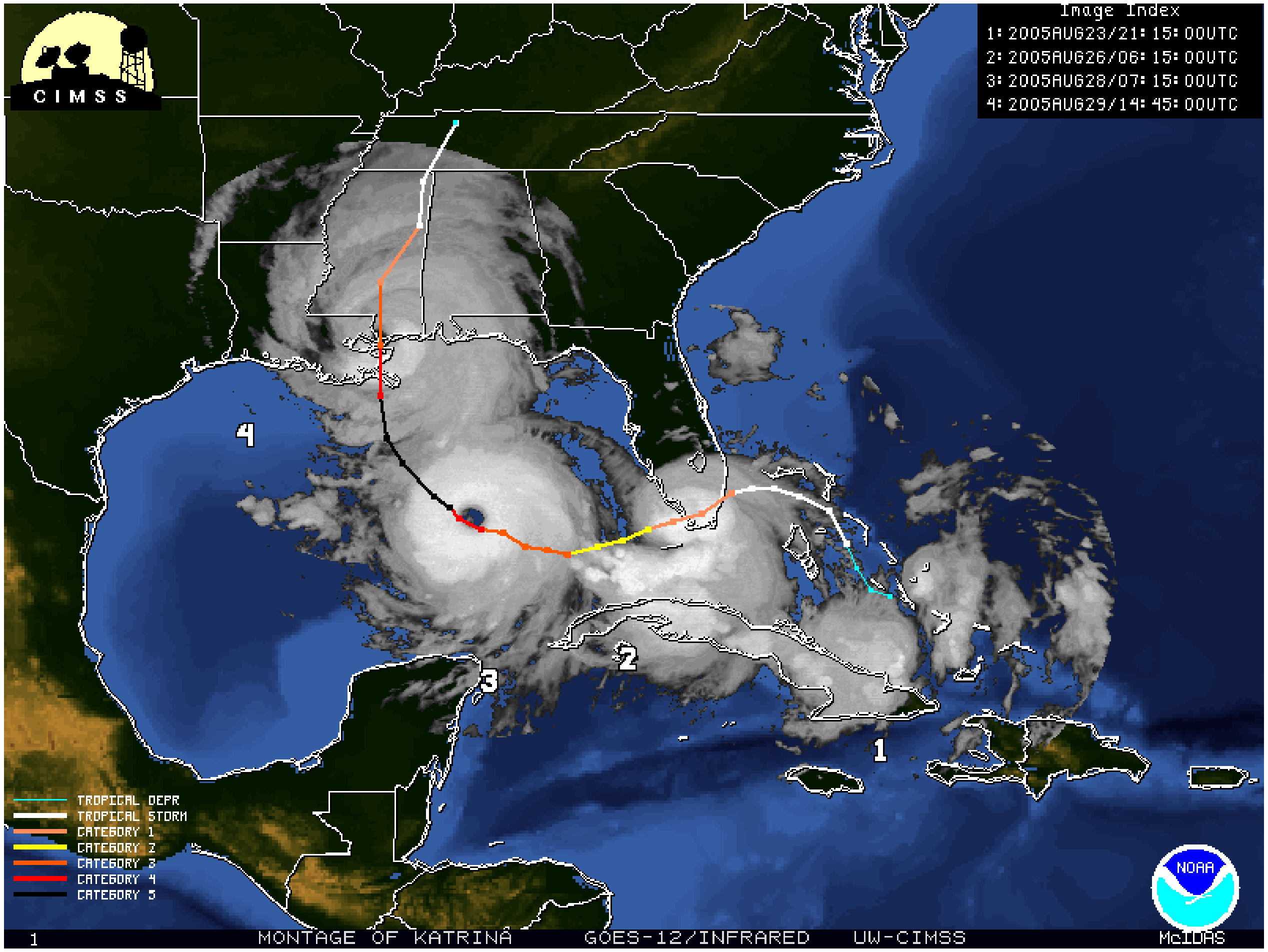
The log-likelihood needs to be maximized numerically

But the rest of the method is the same, swapping both a and b

Tropical cyclones

- (hurricanes, typhoons)

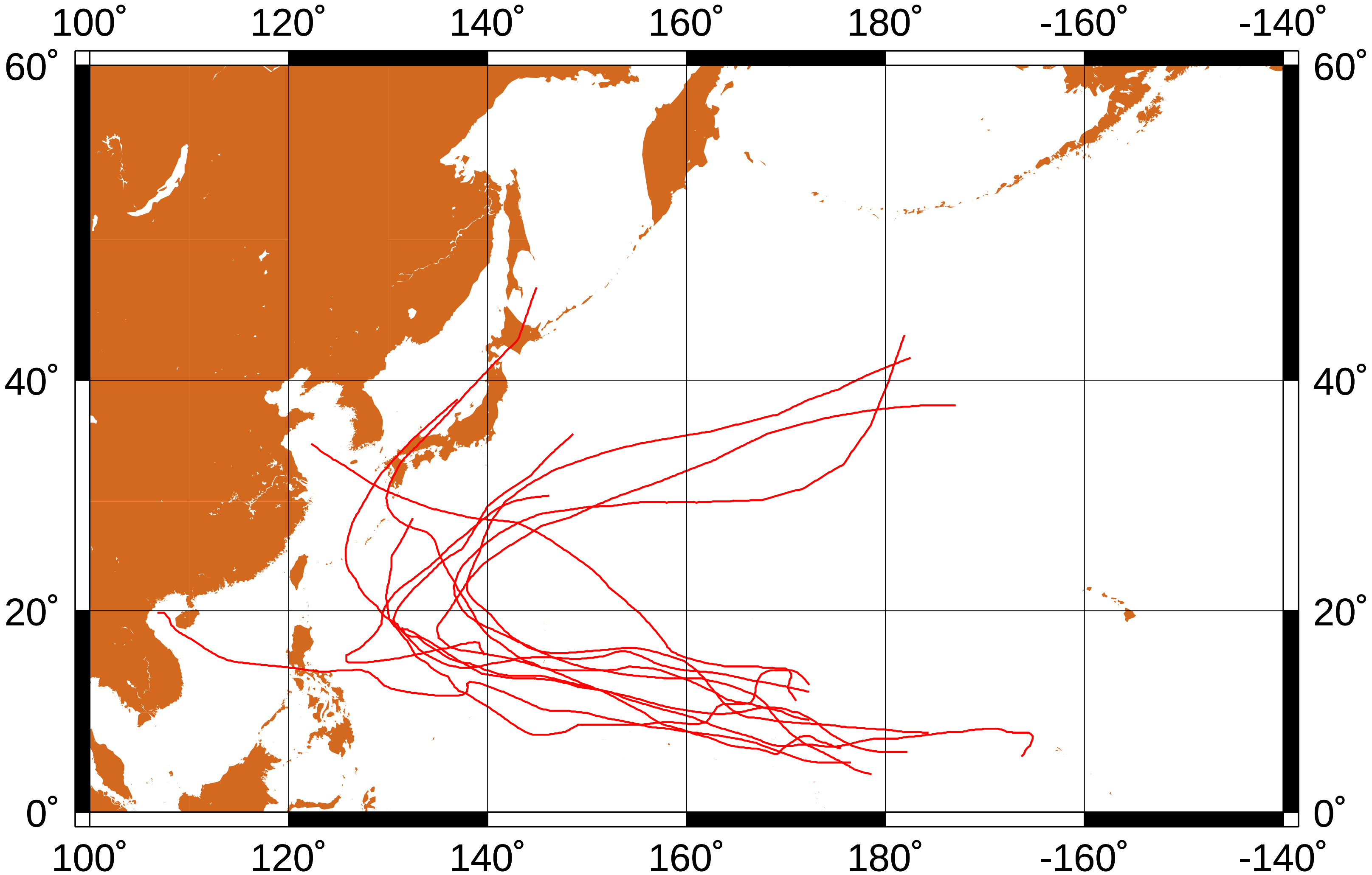
http://cimss.ssec.wisc.edu



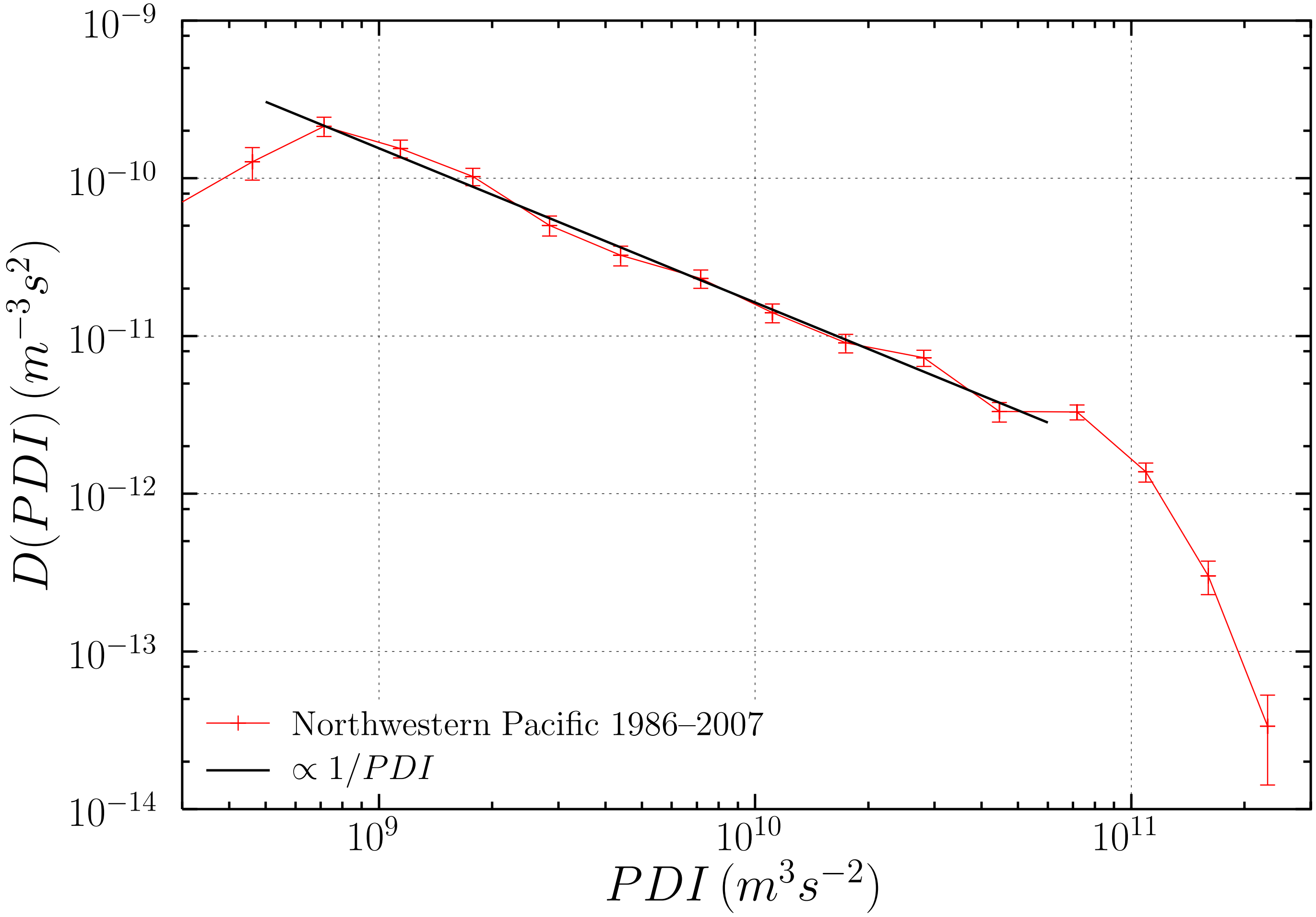
$$\text{Energy} \simeq \int\int C_D \rho |v(r,t)|^3 d^2 r dt$$

Bister & Emanuel, *Met Atm Phys* 1998

- Typhoons in the North Western Pacific (only the largest ones)

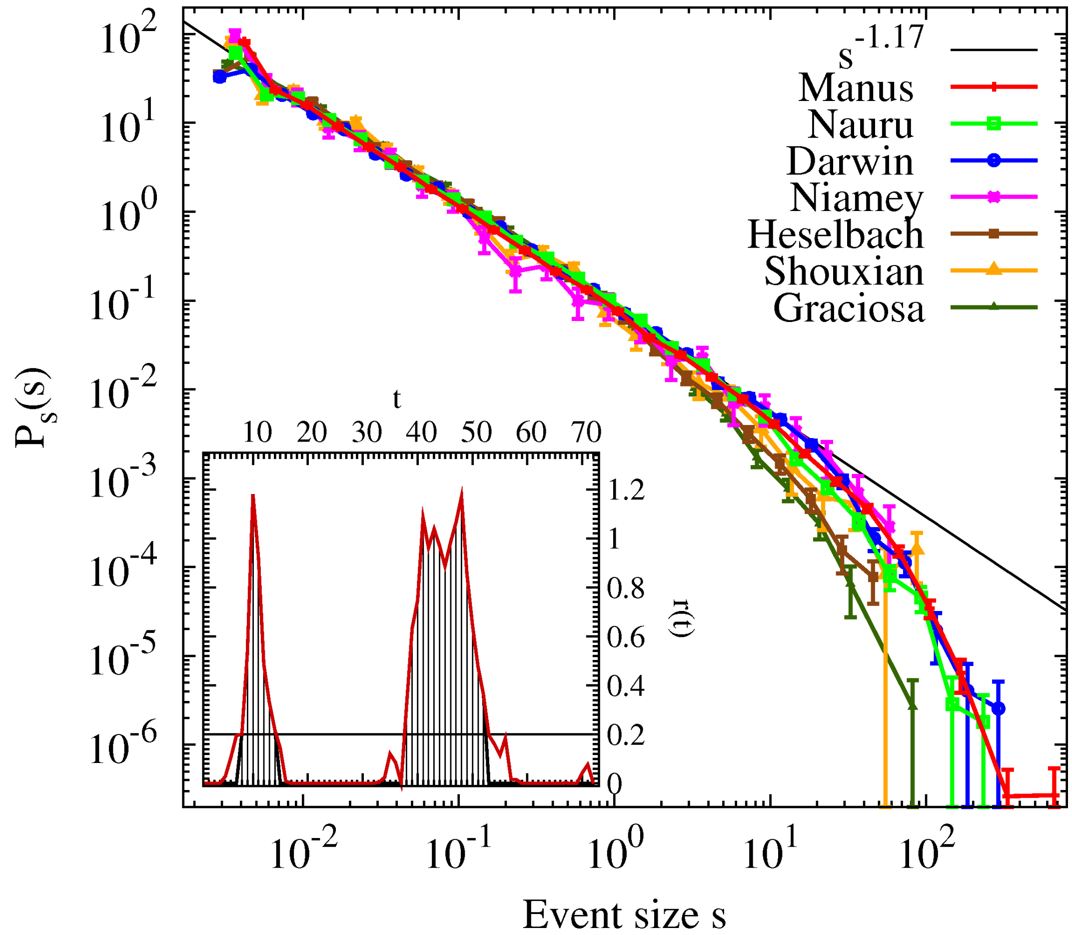


- Tropical cyclones (hurricanes, typhoons), A.C., Ossó, Llebot, *Nature Phys* 2010



- Rainfall

Peters et al. *Stat Mech* 2010



Discrete power laws

A. C., A. Deluca, & R. Ferrer-i-Cancho, *ArXiv* 2012

- The probability function is given by

$$f(n) = \frac{B}{x^\gamma} \quad \text{with } x = a, a+1, \dots, \text{ and } \gamma > 1 \quad \text{and } B = \frac{1}{\zeta(\gamma, a)}$$

where $\zeta(\gamma, a)$ is the Hurwitz theta function ($\zeta(\gamma, 1)$ is the Riemann function)

The log-likelihood is

$$\ell(\gamma) = -\ln \zeta(\gamma, a) - \gamma \ln G$$

which is more difficult to maximize

Care with the cumulative distribution function (for the KS test)

The simulation of the discrete distribution is more involving also

Bibliography

- K. Christensen & N. R. Moloney, *Complexity and Criticality*, IC Press (2005)
- A. Corral & F. Font-Clos, Criticality and self-organization in branching processes: application to natural hazards, in *Self-organized Criticality Systems*, edited by M. Aschwanden, OA Press (2013)
- A. Deluca & A. Corral, Fitting and Goodness-of-Fit Test of Non-Truncated and Truncated Power-Law Distributions, *Acta Geophysica* 61, 1351 (2013)
- A. Corral, Scaling in the timing of extreme events, *Chaos, Solitons, & Fractals* 74, 99 (2015)
- R. Garcia-Millan, F. Font-Clos, & A. C., Finite-size scaling of survival probability in branching processes, *Physical Review E* 91, 042122 (2015)

- `http://www.crm.cat/en/About/People/Researchers/ACorral`