


Entropic Distance for Nonlinear Master Equation

Physica A 474 (2017) 355; Phys.Rev.E 95 (2017) 032130; PLOS ONE 12 (2017) 0179656

T.S. Biró¹ Z. Néda² A. Telcs¹

¹MTA  Research Centre for Physics, Budapest, Hungary

²Hungarian Institute for Physics UBB Cluj, Romania

July 17, 2017

Talk by T.S.Biró at BGL17, Gyöngyös, August 24, 2017.



Entropic Distance and Divergence

Properties of metric distance:

- $\rho(P, Q) \geq 0$ for a pair of points P and Q .
- $\rho(P, Q) = 0 \Leftrightarrow P = Q$ (then and only then)
- $\rho(P, Q) = \rho(Q, P)$ symmetric measure
- $\rho(P, Q) \leq \rho(P, R) + \rho(R, Q)$ triangle inequality in elliptic spaces

Entropic divergence:

- $\rho(P, Q) \geq 0$ for a pair of distributions P_n and Q_n .
- $\rho(P, Q) = 0 \Leftrightarrow P = Q$ (then and only then)
- $\frac{d}{dt}\rho(P, Q) \leq 0$ if Q_n is the stationary distribution

Symmetrized Entropic Divergence

inherited properties

Entropic divergence $\rho(P, Q) = \sum_n \sigma(\xi_n) Q_n$ with $\xi_n = P_n/Q_n$.

Symmetrized kernel function:

$$\mathfrak{s}(\xi) := \sigma(\xi) + \xi \sigma(1/\xi). \quad (1)$$

Jensen inequality tells for $\sigma'' > 0$:

$$\sum_n \sigma(\xi_n) Q_n \geq \sigma \left(\sum_n \xi_n Q_n \right) = \sigma \left(\sum_n P_n \right) = \sigma(1). \quad (2)$$

For property 1 and 2 one sets: $\sigma(1) = 0$.

it follows $\mathfrak{s}(1) = 0$ and $\mathfrak{s}'' > 0$.

Symmetrized Entropic Divergence

emergent new properties

Derivatives:

$$s(\xi) = \sigma(\xi) + \xi \sigma(1/\xi)$$

$$s'(\xi) = \sigma'(\xi) + \sigma(1/\xi) - \frac{1}{\xi} \sigma'(1/\xi)$$

$$s''(\xi) = \sigma''(\xi) - \frac{1}{\xi^2} \sigma'(1/\xi) + \frac{1}{\xi^2} \sigma'(1/\xi) + \frac{1}{\xi^3} \sigma''(1/\xi). \quad (3)$$

Consequences:

- 1 $s(1) = 2\sigma(1) = 0$
- 2 $s'(1) = \sigma(1) = 0$
- 3 $s'' > 0 \Rightarrow \xi_m = 1$ is a minimum
- 4 $s(\xi) \geq 0$.

Entropic distance evolution

General P-Linear Discrete Markovian

Consider $\rho(P, Q) = \sum_n Q_n s\left(\frac{P_n}{Q_n}\right)$ and $\dot{P}_n = \sum_m (w_{nm}P_m - w_{mn}P_n)$.

Using $\xi_n = P_n/Q_n$ we obtain

$$\dot{\rho} = \sum_n s'(\xi_n) \dot{P}_n = \sum_{n,m} s'(\xi_n) (w_{nm} \xi_m Q_m - w_{mn} \xi_n Q_n). \quad (4)$$

Apply $\xi_m = \xi_n + (\xi_m - \xi_n)$ to get

$$\begin{aligned} \dot{\rho} &= \sum_n s'(\xi_n) \xi_n \sum_m (w_{nm} Q_m - w_{mn} Q_n) \\ &+ \sum_{n,m} s'(\xi_n) (\xi_m - \xi_n) w_{nm} Q_m. \end{aligned} \quad (5)$$

Entropic distance evolution

Taylor series remainder theorem in Lagrange form

Recall the Taylor expansion of the kernel function $s(\xi)$,

$$s(\xi_m) = s(\xi_n) + \boxed{s'(\xi_n)(\xi_m - \xi_n)} + \frac{1}{2}s''(c_{mn})(\xi_n - \xi_m)^2, \quad (6)$$

with $c_{mn} \in [\xi_m, \xi_n]$.

It delivers

$$\dot{\rho} = \sum_{n,m} [s(\xi_m) - s(\xi_n)] w_{nm} Q_m - \frac{1}{2} \sum_{n,m} s''(c_{mn}) (\xi_m - \xi_n)^2 w_{nm} Q_m. \quad (7)$$

With positive transition rates, $w_{nm} > 0$ the approach to stationary distribution,

$\dot{\rho} \leq 0$ is hence proven for all $s'' > 0$.

Without detailed balance

Example: Kullback–Leibler divergence

In case of $s(\xi) = -\ln \xi$, we have $s' = -1/\xi$ and $s''(\xi) = 1/\xi^2 > 0$.

The integrated entropic divergence formula (no symmetrization) in this case is given as

Kullback–Leibler divergence



$$\rho(P, Q) = \sum_n Q_n \ln \frac{Q_n}{P_n}. \quad (8)$$

For $P_n^{(12)} = P_n^{(1)} P_n^{(2)}$ also $Q_n^{(12)} = Q_n^{(1)} Q_n^{(2)}$ therefore we have $\xi_n^{(12)} = \xi_n^{(1)} \xi_n^{(2)}$. Aiming at $s(\xi^{(12)}) = s(\xi^{(1)}) + s(\xi^{(2)})$, the solution is $s(\xi) = \alpha \ln \xi$. For $s'' > 0$ it must be $\alpha < 0$, so o.B.d.A. $\alpha = -1$.

Entropic divergence as entropy difference

Example: logarithm

Entropic divergence from the uniform distribution $U_n = 1/W, n = 1, 2, \dots, W$:

Kullback–Leibler divergence



$$\rho(U, Q) = \sum_{n=1}^W Q_n \ln(WQ_n) = \ln W - \sum_n Q_n \ln Q_n = S_{BG}[U] - S_{BG}[Q] \quad (9)$$

with S_{BG} being the Boltzmann–Gibbs–Planck–Shannon entropy formula.

From the Jensen inequality it follows $\rho(U, Q) \geq 0$, so $S_{BG}[U] \geq S_{BG}[Q]$.

Entropic evolution

More general dynamics: P-nonlinear Markovian

Dynamical equation

$$\dot{P}_n = \sum_m [w_{nm} a(P_m) - w_{mn} a(P_n)]. \quad (10)$$

Stationarity condition

$$0 = \sum_m [w_{nm} a(Q_m) - w_{mn} a(Q_n)]. \quad (11)$$

Entropic distance formula

$$\rho(P, Q) = \sum_n \sigma(P_n, Q_n) \quad (12)$$

the dependence on Q_n can be fixed from $\rho(Q, Q) = 0$.

Without detailed balance

Change of entropic distance

$$\dot{\rho} = \sum_{m,n} \frac{\partial \sigma}{\partial P_n} [w_{nm} a(Q_m) \xi_m - w_{mn} a(Q_n) \xi_n] \quad (13)$$

with $\xi_n := a(P_n)/a(Q_n)$.

We put $\xi_m = \xi_n + (\xi_m - \xi_n)$ in the first summand:

$$\dot{\rho} = \sum_n \frac{\partial \sigma}{\partial P_n} \xi_n \sum_m \cancel{[w_{nm} a(Q_m) - w_{mn} a(Q_n)]} + \sum_{n,m} \frac{\partial \sigma}{\partial P_n} w_{nm} a(Q_m) (\xi_m - \xi_n) \quad (14)$$

In order to use the remainder theorem one has to identify



$$\frac{\partial \sigma}{\partial P_n} = s'(\xi_n) = s' \left(\frac{a(P_n)}{a(Q_n)} \right). \quad (15)$$

then $\dot{\rho} < 0$ for $s'' > 0$ and $P \neq Q$.

Without detailed balance

Example: q -Kullback–Leibler divergence

In case of $s(\xi) = -\ln \xi$, we have $s''(\xi) = 1/\xi^2 > 0$.

Now having a fractal nonlinear stochastic dynamics, $a(P) = P^\lambda$.

The integrated entropic divergence formula (no symmetrization):

Tsallis divergence,



$$\frac{\partial \sigma}{\partial P_n} = -\frac{Q_n^\lambda}{P_n^\lambda}, \quad \rho(P, Q) = \sum_n Q_n \ln_\lambda \frac{Q_n}{P_n}. \quad (16)$$

with

$$\ln_\lambda(x) = \frac{1 - x^{\lambda-1}}{1 - \lambda}. \quad (17)$$

Without detailed balance

Example: q -Kullback-Leibler divergence

In case of $s(x) = -\ln_\nu(x)$, we have $s'(x) = -x^{-\nu}$, $s''(x) = \nu x^{-\nu-1} > 0$.

Also having a fractal nonlinear stochastic dynamics, $a(P) = P^\lambda$.

The integrated entropic divergence formula (no symmetrization) becomes

Tsallis divergence, $q = \lambda\nu$



$$\rho(P, Q) = \sum_n \frac{Q_n}{1-q} \left[1 - \left(\frac{P_n}{Q_n} \right)^{1-q} \right] = \sum_n Q_n \ln_q \frac{Q_n}{P_n}. \quad (18)$$

Entropic divergence as entropy difference

Example: q-logarithm

Entropic divergence from the uniform distribution $U_n = 1/W, n = 1, 2, \dots, W$:

$$\rho(U, Q) = \sum_{n=1}^W \frac{Q_n}{1-q} \left[1 - (WQ_n)^{q-1} \right] = W^{q-1} (S_T[U] - S_T[Q]). \quad (19)$$

with S_T being the Tsallis entropy formula:

Tsallis entropy, $q = \lambda\nu$



$$S_T[Q] = \frac{1}{1-q} \sum_n (Q_n^q - Q_n) = - \sum_n Q_n \ln_q(Q_n). \quad (20)$$

From the Jensen inequality it follows $\rho(U, Q) \geq 0$, so $S_T[U] \geq S_T[Q]$. The factor W^{q-1} signifies non-extensivity.

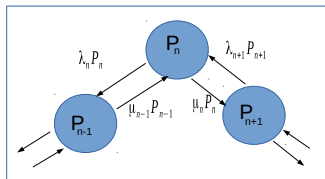
Schemes of Master Equations

Balanced vs One-Sided Growth

Symmetric Short Jumps: Drift + Diffusion

$$w_{nm} = \lambda_m \delta_{n+1,m} + \mu_m \delta_{n-1,m}$$

$$\begin{aligned} \dot{P}_n &= [(\lambda P)_{n+1} - (\lambda P)_n] \\ &- [(\mu P)_n - (\mu P)_{n-1}] \quad (21) \end{aligned}$$

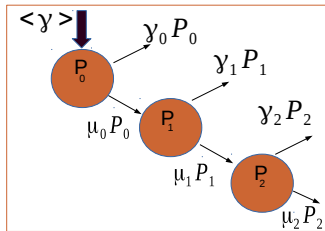


Unidirectional + Resetting

$$w_{nm} = \mu_m \delta_{n-1,m} + \gamma_m \delta_{n,0}$$

$$\dot{P}_0 = \langle \gamma \rangle - (\gamma_0 + \mu_0) P_0 \text{ and } \text{😊}$$

$$\dot{P}_n = \mu_{n-1} P_{n-1} - (\mu_n + \gamma_n) P_n \quad (22)$$



Unidirectional

sources of step-up, μ_n rates

- Propagation (hopping) on a chain: $\mu_n = \text{const.}$
- Rich gets richer: $\mu_n \propto (n + b)$
- Cumulative effect: $\mu_n \propto \sum_{i=1}^n \text{const} \propto n$
- Cancer growth: $\mu_n \propto \exp(n)$.

Resetting

sources of γ_n rates

- Loosing all money when having n : $\gamma_n = \text{const.}$
- Exponential dilution of sample space (cf. citations):

$$\frac{dN_n}{dt} = \mu_{n-1}N_{n-1} - \mu_n N_n;$$

with $P_n(t) = N_n(t)/N(t)$, $N = \sum_n N_n$ and $\gamma_n = \dot{N}/N = \text{const.}$

- Independent decay rate of all units: $\gamma_n \propto n$
- Evolutionary resets in number of species due to catastrophes: $\gamma \rightarrow 0^+$.

Short step-up + long hops to zero: stationary distribution

Stationary limit: $P_n(t) \rightarrow Q_n$, from $\dot{Q}_n = 0$ one obtains

$Q_0 = \langle \gamma \rangle_Q / (\gamma_0 + \mu_0)$ and

stationary ☺

$$Q_n = \frac{\mu_{n-1}}{\mu_n + \gamma_n} Q_{n-1} = \dots = \frac{\mu_0 Q_0}{\mu_n} \prod_{j=1}^n \left(1 + \frac{\gamma_j}{\mu_j} \right)^{-1}. \quad (23)$$

Constant rates

→ exponential distribution

Assume $\mu_j = \sigma$, attachment rate independent of number of links.

$$Q_n = Q_0 \prod_{j=1}^n \frac{\sigma}{\sigma + \gamma} = Q_0 (1 + \gamma/\sigma)^{-n}. \quad (24)$$

Geometrical sum for normalization. We obtain

Boltzmann–Gibbs exponential ☺

$$Q_n = \frac{1}{1 + \sigma/\gamma} e^{-n \cdot \ln(1 + \gamma/\sigma)}. \quad (25)$$

Linear preference, constant loss rate

→ Waring distribution

Linear preference in attachment: $\mu_j = \sigma(j + b)$ ($b > 0$).

$$Q_n = Q_0 \prod_{j=1}^n \frac{j-1+b}{j+b+\gamma/\sigma} = Q_0 \frac{(b)_n}{(c)_n}. \quad (26)$$

with $c = b + 1 + \gamma/\sigma$. Norm:

$$\sum_n Q_n = Q_0 (c-1)/(c-1-b) = 1.$$

Pochhammer ratio (Waring)



$$Q_n = \frac{c-1-b}{c-1} \frac{(b)_n}{(c)_n} \quad (27)$$

Matthias principle: tail of Waring

→ power-law!

The above result in the $n \rightarrow \infty$ limit:

Since

$$\lim_{n \rightarrow \infty} n^{c-b} \frac{\Gamma(n+b)}{\Gamma(n+c)} = 1, \quad (28)$$

we obtain

Pochhammer in $n \rightarrow \infty$ limit: **power-law!** ☺

$$Q_n \rightarrow \frac{\gamma}{\gamma + b\sigma} \frac{\Gamma(c)}{\Gamma(b)} n^{-1-\gamma/\sigma}. \quad (29)$$

Avalanche dynamics in the **large n limit!**

continuous variable: $x = n \cdot \Delta x$

- $P_n(t) = \Delta x \cdot P(n \cdot \Delta x, t)$ ensures $\sum_{n=0}^{\infty} P_n(t) = \int_0^{\infty} P(x, t) dx$.
- $\mu_n = \frac{1}{\Delta x} \cdot \mu(n \cdot \Delta x)$ and $\gamma_n = \gamma(n \cdot \Delta x)$ lead to

Continuum Master:



$$\frac{\partial}{\partial t} P(x, t) = -\frac{\partial}{\partial x} (\mu(x) P(x, t)) - \gamma(x) P(x, t). \quad (30)$$

with the stationary distribution

$$Q(x) = \frac{K}{\mu(x)} e^{-\int_0^x \frac{\gamma(u)}{\mu(u)} du}. \quad (31)$$

Particular continuous stationary distributions

with constant $\gamma(x) = \gamma$.

For constant rate $\mu(x) = \sigma$ **exponential**:

$$Q(x) = \frac{\gamma}{\sigma} e^{-\frac{\gamma}{\sigma} x}. \quad (32)$$

For linear preference $\mu(x) = \sigma(x + b)$ **Tsallis–Pareto**:

$$Q(x) = \frac{\gamma}{\sigma b} \left(1 + \frac{x}{b}\right)^{-1-\gamma/\sigma}. \quad (33)$$

For exponential dispreference $\mu(x) = \sigma e^{-ax}$ **Gompertz**

$$Q(x) = \frac{\gamma}{\sigma} e^{ax + \frac{\gamma}{a\sigma}(1-e^{ax})}. \quad (34)$$

Fluctuation – Dissipation

Conituous

Knowing / observing $Q(x)$ and $\gamma(x)$ one obtains

$$\mu(x) = \frac{1}{Q(x)} \int_x^{\infty} \gamma(u) Q(u) du = \langle \gamma \rangle_{\text{cut}}. \quad (35)$$

Analogy: multiplicative noise

Langevin: $\dot{p} + (\gamma p - \xi) = 0$; stochastic properties: $\langle \gamma p - \xi \rangle = K_1(p)$ and $\langle (\gamma p - \xi)(\gamma p - \xi)' \rangle = K_2(p)$.

Then the Fokker-Planck, $\frac{\partial f}{\partial t} = \frac{\partial}{\partial p}(K_1 f) + \frac{\partial^2}{\partial p^2}(K_2 f) = 0$ has the detailed balance distribution

$$Q(p) = \frac{K}{K_2(p)} e^{-\int_0^p \frac{K_1(q)}{K_2(q)} dq}. \quad (36)$$

The **Fluctuation–dissipation** theorem has the form

$$K_2(p) = \frac{1}{Q(p)} \int_p^{\infty} K_1(q) Q(q) dq. \quad (37)$$

Fluctuation – Dissipation

Discrete

Summing up the recursion from $n = m + 1$ to ∞ delivers

$$\mu_n = \frac{1}{Q_n} \sum_{m=n+1}^{\infty} \gamma_m Q_m. \quad (38)$$

Kubo formula: apply the above to constant γ and exponential distribution $Q_n = e^{-\beta\omega n} / Z$.

$$\mu_n = \frac{\gamma}{e^{\beta\omega} - 1}, \quad \mu(x) = \frac{\gamma}{\beta\omega} \quad (39)$$

Rate, Survival, Hazard

Connection to failure probability

Fluctuation-Dissipation vs. rate reconstruction vs. hazard

Cumulative hazard	$H(x)$
hazard (rate)	$h(x) = H'(x)$
PDF	$Q(x) = h(x) e^{-H(x)}$
Survival (rate)	$R(x) = \int_x^{\infty} Q(u) du = e^{-H(x)}$

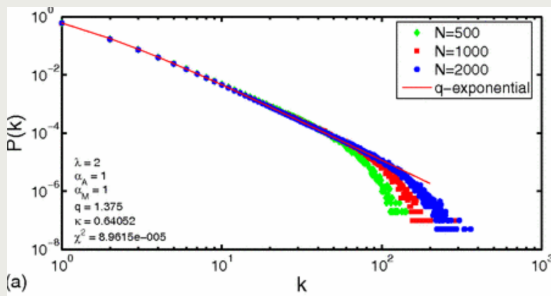
For $\gamma(x) = \gamma$ constant

$$\mu(x) = \gamma \frac{R(x)}{Q(x)} = \frac{\gamma}{h(x)}. \quad (40)$$

Networks: degree distribution

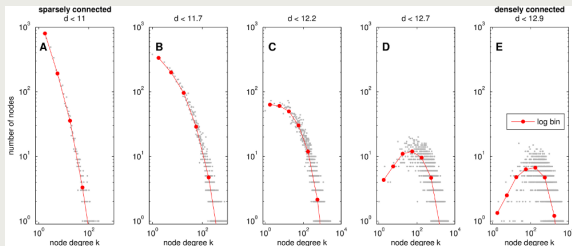
S.Thurner, F.Kyriakopoulos, C.Tsallis, PRE 76 (2007) 036111

- $\mu_n = \sigma(n + b)$ or $\mu(x) = \sigma(x + b)$: Matthew principle for adding the next connection to a node with n
- $\gamma_n = \gamma$: attack success ratio against a node with n connections
- $Q(x) \sim (1 + x/b)^{-\gamma/\sigma-1}$: stationary degree distribution (q-exponential)



Networks: degree distribution 2

M.Sholz, J.DataMining & Digital Humanities, 2015



With $\gamma(x) = \gamma \ln(x/a)$, $\mu(x) = \sigma x$ we get **log normal**

$$Q(x)dx = K e^{-\frac{\gamma}{2\sigma} t^2} dt \quad \text{with} \quad t = \ln(x/a). \quad (41)$$

Citations

Total number and fraction dynamics

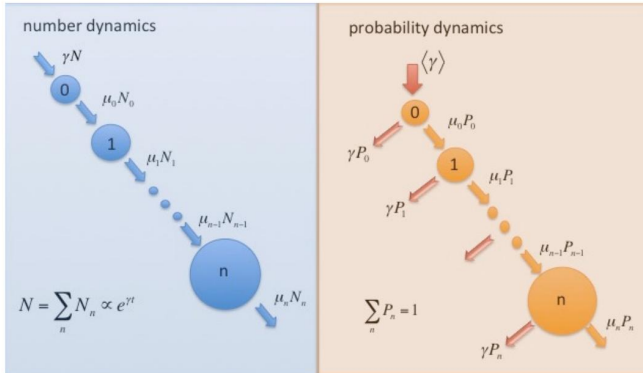


Figure 3. Schematic representation of the coarse-grained random growth model considered in the model. The panel on the left side indicates the growth process in the number of elements with n quanta: N_n . Due to the fact that the total number of elements is exponentially increasing, the probability P that an

Citations

Exponential growth

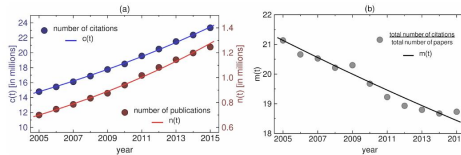


Figure 2. Results for the MEDLINE/PubMed database. **Figure 2a** illustrates the time evolution of the yearly indexed papers, $n(t)$, and the total number of citations, $c(t)$, introduced by them for each year in the 2005-2015 time interval. The trend $n(t)$ can be nicely fitted (red curve) with an exponential curve with $\gamma = 0.06$ using $t_0 = 2005$ and $n_0 = 699915$. Using $t_0 = 2005$, $n_0 = 699915$, $c_0 = 14792864$, $g = 1.4$ and $\gamma = 0.06$ ($\sigma = \gamma/g = 0.043$) the trend for $c(t)$ given by equation (2.3) can be fitted by choosing $b = 1.6$. **Figure 2b** illustrates the time evolution for the yearly incoming total number of citations divided by the total number of new papers, $m(t)$. Using the parameters from $n(t)$ and $c(t)$ the $m(t)$ trend given by equation (2.1) is plotted by the black curve.

Citations

Fraction of n times cited: Facebook and Web of Science

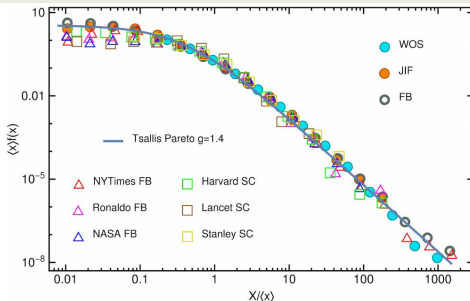


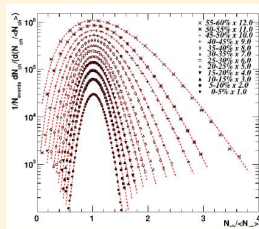
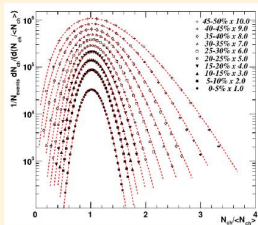
Figure 1. Rescaled distribution of the citation (share) numbers. $f(x)$ is the probability density (PDF) for one paper (post) to have x citations/shares. We present the $\langle x \rangle \cdot f(x)$ value as a function of $x/\langle x \rangle$ ($\langle x \rangle$ the mean value, or first moment of the PDF). For high citation number a clear power-law trend is visible. Different symbols are for different datasets as illustrated in the legend. The considered datasets are described in the Methods section. For high $x/\langle x \rangle$ a clear power-law trend is visible. The entire curve can be well-fitted with a TP distribution (1) with $g \approx 1.4$.

Hadronization

From QGP to n hadrons: NBD

PHENIX, PRC 78 (2008) 044902

Au + Au collisions at $\sqrt{s_{NN}} = 62$ (left) and 200 GeV (right). Total charged multiplicities.

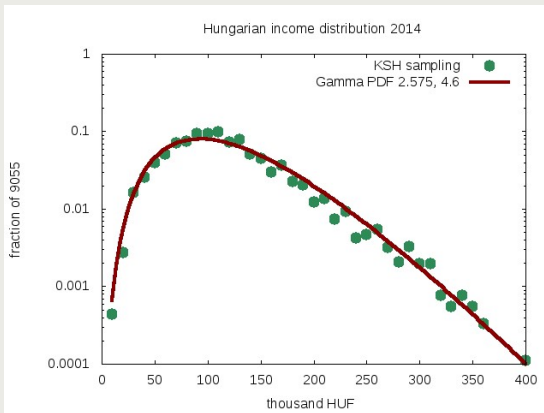


$$\gamma_n = \sigma(n - kf), \mu_n = \sigma f(n + k); \quad Q_n = \binom{n+k-1}{n} f^n (1+f)^{-n-k}.$$

Economy

Income distribution

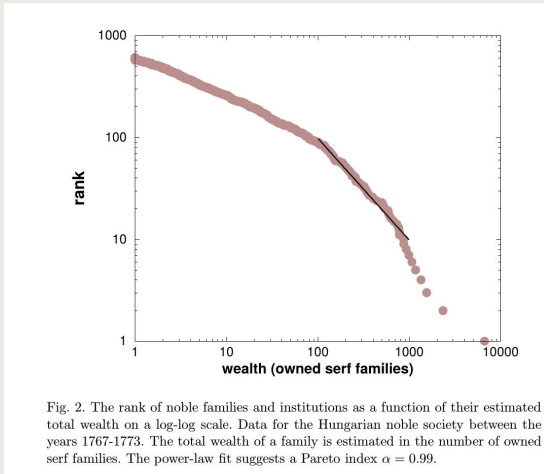
KSH data



$$\gamma(x) = \sigma(ax - c), \quad \mu(x) = \sigma x \quad Q(x) = \frac{a^c}{\Gamma(c)} x^{c-1} e^{-ax}$$

History

Medieval Servant Distribution G.Hegy, Z.Néda, M.A.Santos, Physica A 380 (2007) 271



Summary of Rates and PDF-s

$$\mu(x) = \gamma/h(x)$$

at constant aging γ

$\gamma_n, \gamma(x)$	$\mu_n, \mu(x)$	$Q_n, Q(x)$
const	const	geometrical \rightarrow exponential
const	linear	Waring \rightarrow Tsallis/Pareto
const	sublinear power	Weibull
const	quadratic polynomial	Pearson
const	exp	Gompertz
$\ln(x/a)$	σx	Log-Normal
linear	const	Gauss
$\sigma(ax - c)$	σx	Gamma

Deviation shrinks and moves as a soliton:

$$\dot{x}_c = \mu(x_c) !$$

Bibliographical Notes

- 1 Tsallis/Pareto degree distribution in huge networks**
Stefan Thurner, Fragiskos Kyriakopoulos, Constantino Tsallis: *Unified model for network dynamics exhibiting nonextensive statistics* Phys. Rev. E **76** (2007) 036111
- 2 Avalanche Master Equation for Citations**
A. Schubert, W. Glänzel: *A dynamic look at a class of skew distributions. A model with scientometric applications* Scientometrics **6** (1984) 149-167
- 3 Waring distribution** J. O. Irwin: *The Place of Mathematics in Medical and Biological Statistics* J. Roy. Stat. Soc. A **126** (1963) 1-45
- 4 Generalized Waring distribution** J. O. Irwin: *The Generalized Waring Distribution Applied to Accident Theory* J. Roy. Stat. Soc. A **131** (1968) 205-225
- 5 Rate reconstruction from a truncated integral** W. Glänzel, A. Telcs, A. Schubert: *Characterization by Truncated Moments and Its Application to Pearson-Type Distributions* Z. Wahrscheinlichkeitstheorie verw. Gebiete **66** (1984) 173-183
- 6 Avalanche Master eq. for networks** P. L. Krapivsky, S. Redner, F. Leyvraz: *Connectivity of Growing Random Networks* Phys. Rev. Lett. **85** (2000) 4629-4632
- 7 Waring Degree Distribution in Network – eq.8 in** P. L. Krapivsky, G. J. Rodgers, S. Redner: *Degree Distributions of Growing Networks* Phys. Rev. Lett. **86** (2001) 5401-5404
- 8 Hypergeometrical result in aging site network** S. N. Dorogovtsev, J. F. F. Mendes, *Evolution of networks with aging of sites* Phys. Rev. E **62** (2000) 1842-1845