Entropy and Wealth
A presentation explaining the paper with the same title
published in *Entropy*, 2021, https://doi.org/10.3390/e23101356

Demetris Koutsoyiannis and G.-Fivos Sargentis

October 2021
Entropy and wealth

\[ \Phi[x] := \mathbb{E}[-\ln P(x)] \]

\begin{align*}
\text{entropy} & \quad \text{uncertainty} \\
\text{δύναμις / potentia} & \quad \text{plurality of options} \\
\text{ε\acute{ν}τροπία} & \quad \text{freedom}
\end{align*}
Entropy and wealth

Title: Entropy and Wealth
Authors: Demetris Koutsoyiannis and G-Fivos Sargentis

Abstract: While entropy was introduced in the second half of the 19th century in the international vocabulary as a scientific term, in the 20th century it became common in colloquial use. Popular imagination has loaded “entropy” with almost every negative quality in the universe, in life and in society, with a dominant meaning of disorder and disorganization. Exploring the history of the term and many different approaches to it, we show that entropy has a universal stochastic definition, which is not disorder. Hence, we contend that entropy should be used as a mathematical (stochastic) concept as rigorously as possible, free of metaphorical meanings. The accompanying principle of maximum entropy, which lies behind the Second Law, gives explanatory and inferential power to the concept, and promotes entropy as the mother of creativity and evolution. As the social sciences are often contaminated by subjectivity and ideological influences, we try to explore whether maximum entropy, applied to the distribution of a wealth-related variable, namely annual income, can give an objective description. Using publicly available income data, we show that income distribution is consistent with the principle of maximum entropy. The increase in entropy is associated to increases in society’s wealth, yet a standardized form of entropy can be used to quantify inequality. Historically, technology has played a major role in the development of and increase in the entropy of income. Such findings are contrary to the theory of ecological economics and other theories that use the term entropy in a Malthusian perspective.

Keywords: entropy; wealth; income distribution; options; potentiality; principle of maximum entropy; Second Law; stochastic
Some of the paper’s contents:

- History of the entropy concept
- Definition of entropy
- The principle of maximum entropy
- The entropy maximizing distribution for constrained mean
- An entropic index of inequality, “standardized entropy”
- The exponential vs. Pareto distribution tails
- Application to societies’ income distribution
  - From the ancient classless society to modern stratified societies
  - Income redistribution in organized societies (toy model)
  - Empirical investigation of the distribution tail of wealth (i.e. elites)
  - Empirical investigation: USA, Sweden 2019
- Conclusions about the wealth distribution
Some of the paper’s refutations

- The misrepresentation of entropy in colloquial language
- The interpretation of entropy as disorder and disorganization
- The existence of an anti-entropic (negentropic) life principle
- Some results in entropic analyses in Econophysics
- The use of entropy in Ecological Economics
- The general use of the “80/20 rule” (misnamed as the “Pareto principle”)
- The “hard determinism” in Marxism
- The world-control aspect of the elites
- The Malthusian ideas
- The aspiration of degrowth

Pablo Picasso. Don Quixote and Sancho Panza. 1955
Dedication

This work is dedicated to the memory of Themistocles Xanthopoulos whose book series “Requiem with Crescendo?” triggered it.
Introduction

“*Wealth is not about having a lot of money; it’s about having a lot of options*”

Chris Rock (American comedian, writer, producer and director)

“Πάντα τὰ ἐμὰ μετ’ ἐμοῦ φέρω” (“*Omnia mea mecum porto*”; “*All that is mine I carry with me*”)

Bias of Priene (one of the seven Greek sages; 6th c. BC; quoted in Latin by Cicero, *Paradoxa Stoicorum* I, 8)
The name entropy

The name “entropy” has been introduced about 150 years ago as a scientific term but later its use became common of everyday language. We can find it in literature in poetry in press, and in web posts, but often its use is irrelevant to its real scientific meaning.


The most common use of the word entropy is when a writer wants to describe with an “intellectual” word a kind of disorder.

Wikipedia has a lemma for that: Entropy (order and disorder)

The colloquial use of the name entropy

According to data given by Word and Phrase Info and plotted in Figure, the frequency of “entropy” in fiction, for example, is not dramatically lower than in the academic texts.


Frequency of appearances of the words: entropy, disorder and uncertainty in the Corpus of Contemporary American English. The common and neutral word “and” was added as a proxy guide of the relative frequencies of the five indicated genres, which appear to be very close to each other.
The academic use of entropy

- The academic corpus can be investigated using bibliometric databases, such as Scopus.
- Expectation: the term “entropy” would more frequently appear in scholar articles in combination with terms such as “physics” or “thermodynamics”—this is the case for the recent years.
- Before 1960s the combination of term “entropy” with “society” or “social” was more frequent than the former.
- As seen in the figure, in the 21st century “entropy” is also used in combination with ecology and economics.

Relative frequency of appearances of the indicated key phrases in the article title, abstract and keywords of about 70 million articles written in English, which are contained in the Scopus database.

Scopus database, https://www.scopus.com/
The metaphorical use of entropy

- Out of its physical and stochastic context, the term “entropy” is typically used metaphorically and hence its meaning becomes ambiguous or diverse. For example, the term “social entropy”, in one of its earliest uses in scholarly publications [1] is equated to “dereliction, pollution and waste”, which are created by “economic activity” or by “society as consumers” and has to be minimized. Bailey in his book entitled “Social Entropy Theory” [2] tried to illuminate

  “the fundamental problems of societal analysis with a nonequilibrium approach, a new frame of reference built upon contemporary macrological principles, including general systems theory and information theory.”

- His interest is more to illuminate the “Social Entropy Theory” than define social entropy per se.

Entropy in social theories

- Even in an overview of his book “Social Entropy Theory, Bailey did not provide a definition of social entropy. In a critique of the book, Mayer finds the “unrelenting abstractness of Social Entropy Theory quite frustrating” and adds:

  “Never is the theory applied to real sociological data or anything like a real social situation.”


- Different aspects of entropy and energy in social systems have been examined in several papers.

Koutsoyiannis and Sargentis, Entropy and Wealth
In economics, Frederick Soddy (1877 –1956) and Nicholas Georgescu-Roegen (1906 –1994), fascinated by entropy in thermodynamics, sought analogies with economics and development in a Malthusian perspective.

Their interpretation, predominantly popular in social sciences and economics, wants entropy to be the waste produced when useful work is done.

Entropy used in Malthusian perspective (2)

- Ehrlich, P. R. The population bomb, Sierra Club/Ballantine Books, New York, 1968.

Abuse of the term entropy

- The use of the notion of entropy is mainly metaphorical, rich in imaginary interpretation and divergent.

The dominant view is that entropy epitomizes all “bad things” one can think in the universe, in life, in human societies and in economics.

- Our view is quite different.
  1. We insist that entropy should be used as a mathematical (in particular, stochastic) concept. We avoid using ambiguous terms such as “social entropy”. We claim that any interpretation of entropy should be as close to the mathematical definition as possible.
  2. On the second issue, we believe that the overloading of the concept of entropy with negative properties reflects misunderstanding of the underlying theory, guided by a deterministic world view.

Hieronymus Bosch. The Garden of Earthly Delights
https://en.wikipedia.org/wiki/Hieronymus_Bosch

Koutsoyiannis and Sargentis, Entropy and Wealth
The conception of entropy (1)

The name ἐντροπία (Greek for entropy) appears already in ancient Greek [1] (from the verb ἐντρέπειν, to turn into, to turn about) but was introduced in the international scientific vocabulary by Rudolf Clausius only in 1865 [2].

The rationale for introducing the term is explained in his own words, which indicate that he was not aware of the existence of the very word in ancient Greek [3]:

“We might call $S$ the transformational content of the body [...] But as I hold it to be better to borrow terms for important magnitudes from the ancient languages, so that they may be adopted unchanged in all modern languages, I propose to call the magnitude $S$ the entropy of the body, from the Greek word τροπή, transformation. I have intentionally formed the word entropy so as to be as similar as possible to the word energy; for the two magnitudes to be denoted by these words are so nearly allied their physical meanings, that a certain similarity in designation appears to be desirable.”

The conception of entropy (2)

- Clausius recognised that entropy is related to transformation and change, and the contrast between entropy and energy, where the latter is a quantity that is conserved in all changes. This meaning has more clearly been expressed in Clausius’ famous aphorism [4]:

“Die Energie der Welt ist constant. Die Entropie der Welt strebt einem Maximum zu.”

(The energy of the world is constant. The entropy of the world strives to a maximum).

- Hence, entropy and its ability to increase is the driving force of change.

- Mathematically, the thermodynamic entropy, \( S \), is defined in the same Clausius’ texts through the equation \( dS = \frac{\delta Q}{T} \), where \( Q \) and \( T \) denote heat and temperature.

- The definition, however, applies to a reversible process only. The fact that in an irreversible process \( dS > \frac{\delta Q}{T} \) makes the definition imperfect and affected by circular reasoning, as, in turn, a reversible process is one in which the equation holds.
In 1877, Ludwig Boltzmann (1844 -1906) gave entropy a probabilistic content as he linked it to probabilities of statistical mechanical system states, thus explaining the Second Law of thermodynamics as the tendency of the system to run toward more probable states, which have higher entropy.


In 1902, the probabilistic concept of entropy was advanced in thermodynamics by Josiah W. Gibbs (1839 – 1903).


In 1940, Claude Shannon (1916 – 2001) used an essentially similar, albeit more general, entropy definition to describe the information content, which he also called entropy.


In 1957, Edwin T. Jaynes (1922 –1998), introduced the principle of maximum entropy. This postulates that the entropy of a stochastic system should be at maximum, under some conditions, formulated as constraints, which incorporate the information that is given about this system.

Pioneers of entropy

Rudolf Clausius (1822 –1888)

Entropy as a probabilistic entity

1877

1865

Introduction of entropy in thermodynamics (and in the international scientific vocabulary)

Ludwig Boltzmann (1844 -1906)

Josiah W. Gibbs (1839 – 1903)

Advancement of the probabilistic concept of entropy in thermodynamics

1902

Claude Shannon (1916 – 2001)

More general probabilistic definition of entropy within information theory

1940


Introduction of the principle of maximum entropy

1957

Koutsoyiannis and Sargentis, Entropy and Wealth

19
Are thermodynamic and probabilistic entropy different?

- More than 150 years after the introduction of the entropy concept, its meaning is still debated.

- Note: the units of thermodynamic entropy are only an historical accident, related to the arbitrary introduction of temperature scales


- The connection of probabilistic and thermodynamic entropy is clearly implied by its pioneers, Boltzmann [1], Gibbs [2], Szilard [3] and von Neumann [4]. More recent accounts of the connection have been provided by Robertson [5] and Moore [6]. Furthermore, as has recently been shown [7,8], the thermodynamic entropy of gases can be easily produced by the formal probability theory without the need of strange assumptions (e.g. indistinguishability of particles).

Does entropy measure disorder? (1)

- In the public perception entropy has a negative content, and is typically identified with disorganization or disorder, and deterioration.

- Boltzmann did not identify entropy with disorder, even though he used the latter word in a footnote appearing in two papers of his, in which he speaks about the
  
  “agreement of the concept of entropy with the mathematical expression of the probability or disorder of a motion.”


- Wiener, the most influential scientist who supported the disorder interpretation. In 1946, he gave a keynote speech at the New York Academy of Sciences in which he declared that:

  “Information measures order and entropy measures disorder.”


- Also, in his influential book Cybernetics, he stated:

  “the entropy of a system is a measure of its degree of disorganization”

Does entropy measure disorder? (2)

- **Even in the 21st century, the disorder interpretation is dominant.** For example, Chaitin stated:

  > “Entropy measures the degree of disorder, chaos, randomness, in a physical system. A crystal has low entropy, and a gas (say, at room temperature) has high entropy.”


- **Atkins also explains entropy as disorder.** Additionally, he notes:

  > “the world is getting worse, that it is sinking purposelessly into corruption, the corruption of the quality of energy, is the single great idea embodied in the Second Law of thermodynamics.”

The very definition of entropy is inconsistent with a deterministic world view.

This entails difficulties in understanding entropy because our education is based on the deterministic paradigm.

Indeed, it is difficult to incorporate the clearly stochastic concept of entropy in a deterministic mindset.

Therefore, many have tried to find analogues more deterministic-friendly, identifying it with disorganization, disorder, deterioration.

All these have a negative connotation in the deterministic mindset.

But they are less appropriate and less rigorous as scientific terms and more appropriate in describing mental states and even more so in describing socio-political states (cf. Nazis’ “new order”, Kissinger’s “new world order” and Schwab’s “global order”).
Shannon in a deterministic mindset (2)

- In one of the earliest critiques of the disorder interpretation of entropy, Wright [1], makes a plea for moderation in the use of “intuitive qualitative ideas concerning disorder”. At a more absolute tone, recently Leff [2] states:
  
  “The too commonly used disorder metaphor for entropy is roundly rejected.”

- In an even more recent article, Styer [3] states:
  
  “we cannot stop people from using the word “entropy” to mean “disorder” or “destruction” or “moral decay.” But we can warn our students that this is not the meaning of the word “entropy” in physics.”

- Steyer attributes an excessive contribution to the misconception of entropy as disorder to autobiographical book “The Education of Henry Adams”. As quoted by Steyer, Adams contests Chaos and Anarchy, and states:
  
  “The kinetic theory of gas is an assertion of ultimate chaos. In plain words, Chaos was the law of nature; Order was the dream of man.”

- Undoubtedly, elites that want to control the world have exactly this dream. But this does not necessarily mean that the entire humanity has the same dream with the elites.

---

Entropy as a probabilistic concept

- When speaking about entropy, we should have in mind that the scale is an important element and that entropy per se, being a probabilistic concept, presupposes a macroscopic view of phenomena, rather than a focus on individuals or small subsets.

- If we viewed the motion of a particular dice throw, we might say that it was irregular, uncertain, unpredictable, chaotic, random.

- But macroscopization, by removing the details, may also remove irregularity. For example, the application of the principle of maximum entropy to the outcomes of a die results in equal probabilities (1/6) of each outcome. This is perfect order.

- Likewise, the maximum uncertainty in a particular water molecule’s state (in terms of position, kinetic state and phase), on a macroscopic scale results in the Clausius–Clapeyron law.

- As the accuracy of this law is so high, most people believe that the principle of maximum entropy is a deterministic law.
But if entropy is not disorder, what is it?

According to its standard definition, entropy is precisely the expected value of the minus logarithm of probability.

If this sounds too difficult to interpret, an easy and accurate interpretation is that entropy is a measure of uncertainty.

If “disorder” is regarded as a “bad thing”, for many the same is the case with uncertainty. The expressions “uncertainty monster” and “monster of uncertainty” appear in about 250 scholarly articles registered in Google Scholar.

However, if uncertainty is a monster, it is thanks to this monster that life is livable and fascinating.
Entropy: a measure of uncertainty (2)

- Uncertainty is not an enemy of science or of life; rather it is the mother of creativity and evolution.

- Without uncertainty, life would be a “universal boredom”, and concepts such as hope, will (particularly, free will), freedom, expectation, optimism, etc., would hardly make sense.

- A technocratic system where an elite comprising super-experts who, using super-models, could predict the future without uncertainty, would also assume full control on the society.

- Fortunately, this will never happen because entropy, i.e. uncertainty, is a structural property of nature and life. Hence, in our view, uncertainty is neither disorder nor a “bad thing”. How could the most important law of physics (the Second Law) be a “bad thing”?

Uncertainty is depicted by Marc Chagall’s Palette
In a deterministic world view, there is no uncertainty and there is no meaning in speaking about entropy.

If there is **no uncertainty**, each outcome is accurately predicted and hence there are **no options**.

In contrast, in an **indeterministic world**, there is a **plurality of options**. This corresponds to the Aristotelian idea of δύναμις (Latin: potentia—English: potency or potentiality).

This view, also is consistent with what was vividly expressed by Brissaud [67]:

“Entropy measures freedom, and this allows a coherent interpretation of entropy formulas and of experimental facts. To associate entropy and disorder implies defining order as absence of freedom”


The existence of options entails that there is freedom, in the following sequence:

\[ \text{entropy} \leftrightarrow \text{uncertainty} \leftrightarrow \text{plurality of options} \leftrightarrow \text{freedom} \]
Figure 4. An attempt at an artistic representation of the notion of entropy. Uncertainty is depicted by Marc Chagall’s Palette (adapted from [105]) and freedom by Marc Chagall’s Self-Portrait with Seven Fingers [106]; δύναμις (Greek) or potentia (Latin) is the Aristotelian idea of potency or potentiality.
Given the dominance of the negative connotations assigned to entropy, many thought that it makes it impossible to understand life, when the whole world is ruled by the Second Law of thermodynamics reflecting entropy’s tendency to become maximal.

Their perception is that maximum entropy points toward death and annihilation. Therefore, they have thought about an anti-entropic or negentropic principle governing life, biosphere, economy, etc., because these convert things which have less order, into things with more order.

https://en.wikipedia.org/wiki/The_Anatomy_Lesson_of_Dr._Nicolaes_Tulp
Schrödinger argued that “What an organism feeds upon is negative entropy”.

At the same time, he did not mention any other “life principle” additional to the Second Law that would drive life and evolution.

In a rather metaphysical context, assuming a non-statistical definition of negentropy, others see a negentropic principle governing life, biosphere, economy, etc., because these convert things which have less order, into things with more order.
The meaning of life

- If we see entropy as uncertainty, we also understand that life is fully consistent with entropy maximization.
- The human-invented steam engines (and other similar machines) increase entropy all the time, being fully compatible with the Second law, yet they produce useful work.
- Likewise, the biosphere increases entropy, yet it produces interesting patterns, much more admirable than steam engines. Life generates new options and increases uncertainty.

Compare Earth with a lifeless planet: Where is uncertainty greater? In which of the two planets a newspaper would have more events to report every day?
Final theses on entropy

- **Entropy is a stochastic concept** with a simple and general definition. Notably, according to its stochastic definition, entropy is a dimensionless quantity.

- As a stochastic concept, *entropy can be interpreted as a measure of uncertainty*, leaving aside the traditional but obscure and misleading “disorder” interpretation.

- The classical definition of thermodynamic entropy is not necessary; it can be abandoned and replaced by the probabilistic definition.

- Applied in thermodynamics, the thus defined entropy is the fundamental quantity, which supports the definition of all other derived ones.

- The entropy concept is equipped with the principle of maximum entropy, which states that entropy tends to take its maximum value that is allowed, given the available information about the system.

- The tendency of entropy to reach a maximum is the driving force of natural change.

- Life, biosphere and social processes are all consistent with the principle of maximum entropy as they augment uncertainty. Therefore, no additional “life principle” is necessary to explain them.
Mathematical formulation (1)

We consider a stochastic (random) variable $x$ (notice that we underline stochastic variables to distinguish them from regular variables) and we denote its distribution function (i.e. probability of non-exceedance) and its tail function (i.e. probability of exceedance), respectively, as:

$$F(x) := P\{x \leq x\}, \quad \bar{F}(x) = 1 - F(x) = P\{x > x\}$$

where $P$ denotes probability.

If the variable $x$ is discrete, i.e. it can take any of the values $x_j, j = 1, ..., \Omega$, with probability

$$P_j \equiv P(x_j) := P\{x = x_j\}$$

then the sequence $P_j$ defines its probability mass function. If the variable is continuous, i.e. it can take any real value (or a value in a subset of the real numbers), then we define the probability density function as the derivative of the distribution function.

$$f(x) := \frac{dF(x)}{dx}$$
Mathematical formulation (2)

- From general postulates about uncertainty, a unique (within a multiplicative factor) metric $\Phi$ results, serves as the definition of entropy.

- For a discrete stochastic variable:

$$
\Phi[x] := E[-\ln P(x)] = - \sum_{j=1}^{\Omega} p_j \ln p_j
$$

- For a continuous stochastic variable:

$$
\Phi[x] := E \left[ -\ln \frac{f(x)}{\beta(x)} \right] = - \int_{-\infty}^{\infty} \ln \frac{f(x)}{\beta(x)} f(x) \, dx
$$
Conservation Laws – and the exception of entropy

**Mass**

\[
\text{CH}_4 + 2\text{O}_2 \rightarrow \text{CO}_2 + 2\text{H}_2\text{O}
\]

**Momentum**

**Energy**

\[
E_p = m \ g \ h
\]
\[
E_k = 0
\]
\[
E_p + E_k = \text{constant}
\]

**Entropy**

\[
E_p = 0
\]
\[
E_k = m \ v^2/2
\]
Entropy maximizing distribution

- In studying the material wealth (or income) in a certain society, current or past, **we assume two characteristic quantities**: the mean $\mu$, which is related to the total energy available to the society, and an upper limit of wealth (or income) $\Omega$, which is mainly determined by the available technology (knowhow) and thus we call it technological upper limit. We define the ratio:

  \[ A := \frac{\Omega}{\mu} \geq 1 \]

- For any background measure $\beta(x)$, after incorporating the constraints to the entropy with Lagrange multipliers, the general solution of entropy maximizing density is:

  \[ f(x) = A \beta(x) \exp \left( - \sum_i b_i g_i(x) \right) \]
Entropy maximizing distribution with Lebesgue background measure (1)

Assuming a Lebesgue background measure with \( \beta(x) = 1/\lambda \), with \( \lambda \) being a monetary unit (e.g., \( \lambda = \text{USD} \ 1 \)), after algebraic manipulations, we find the entropy maximizing probability density to be:

\[
f(x) = \frac{b \ e^{-bx/\mu}}{\mu(1 - e^{-bA})}
\]

After algebraic manipulations, we find the maximum entropy:

\[
\Phi[x] = b(A) + \ln \left( \frac{\mu}{\lambda} \right) + \ln \left( \frac{A}{(A - 1) b(A) + 1} \right) =: \Phi(\mu, A)
\]
Maximum entropy as a function of (a) the technological limit $\Omega$ for constant mean $\mu = 1$ and (b) the mean $\mu$ for a constant technological limit $\Omega = 2.958$. The red dashed lines correspond to identical cases in the two panels. The green dotted line in panel (b) depicts the maximum entropy for the infinite technological limit.
Entropy maximizing distribution with hyperbolic background measure

- If we choose a hyperbolic background measure i.e., $\beta_H(x) = \frac{1}{\lambda + x}$, leaving $x$ unbounded, and constrain a generalized mean, consistent with the chosen background measure, then the entropy maximizing distribution comes to be of Pareto type, with density

$$f(x) = \frac{1}{\lambda_1} \left(1 + \xi \frac{x}{\lambda_1}\right)^{-\frac{1}{\xi}-1}, \quad x \geq 0$$

where $\xi := 1/b_1$ is known as the tail index and $\lambda_1 := \lambda \xi$.

- Maximized entropy for hyperbolic background measure

$$\Phi_H = 1 + \ln \xi = 1 - \ln(1 + \lambda/\mu)$$

- Entropy for the Lebesgue measure $\beta(x) = 1/\lambda$

$$\Phi = 1 + \xi + \ln \xi = 1 + \frac{1}{1 + \lambda/\mu} - \ln(1 + \lambda/\mu)$$

Note that: $\Phi_H \leq \Phi \leq 1 + \ln \left(\frac{\mu}{\lambda}\right) = \Phi_{exp}$
A graphic comparison of the exponential distribution (entropy $\Phi = 1$) with Pareto density and the two-parameter gamma density, the behaviour of which is opposite to Pareto.

Note that, in Pareto distribution and gamma distribution, entropy $\Phi = 0.884$. 
Entropic measure of inequality

While for a constant background density equal to the inverse of the monetary unit (i.e., $1/\lambda$ with $\lambda$ equal, e.g., to USD 1) the entropy provides a measure of society’s wealth (even if $x$ expresses income), if we change the background measure to the value $1/\mu$, where $\mu$ is the mean income, the consequently calculated entropy is a measure of inequality. Calling the latter quantity standardized entropy and denoting it as $\Phi_\mu(x)$ we get:

$$\Phi_\mu[x] = \Phi[x] - \ln \frac{\mu}{\lambda}$$
Entropy: From the ancient classless society to modern stratified societies (1)

- **Remember:** \( A := \frac{\Omega}{\mu} \geq 1 \)
- The mean \( \mu \), is related to the total energy available to the society.
- The upper limit of wealth (or income) \( \Omega \), is mainly determined by the available technology (knowhow) and thus we call it technological upper limit.

GDP per capita per year related to consumption of energy per capita per year and life expectancy; global average 1870–2011

Maddison Project Database. 2018.
https://www.rug.nl/ggdc/historicaldevelopment/maddison/releases/maddison-project-database-2018
An entropic view of the distribution of wealth through history

Koutsoyiannis and Sargentis, Entropy and Wealth
Entropy: From the ancient classless society to modern stratified societies (3)

Power of the largest prime movers in different eras


An entropic view of the distribution of wealth through history

\[ f(x) = \begin{cases} 
1 & \text{if } x = 0 \\
\mu & \text{if } x = 1 \\
A & \text{if } x = 2 \\
\Omega & \text{if } x = 3 \\
\phi & \text{if } x = 4 \\
0 & \text{otherwise} 
\end{cases} \]
Entropy: From the ancient classless society to modern stratified societies (4)

Power of the largest prime movers in different eras


An entropic view of the distribution of wealth through history

\[ f(x) \]

- \( \mu = 1 \)
- \( \lambda = 2.958 \)
- \( \Omega = 2.958 \)
- \( b = 0.703 \)
- \( \Phi = 0.922 \)
Entrophy: From the ancient classless society to modern stratified societies (5)

Power of the largest prime movers in different eras


An entropic view of the distribution of wealth through history

\[ f(x) \]

- \( \mu = 1.5 \)
- \( A = 4 \)
- \( \Omega = 5 \)
- \( b = 0.898 \)
- \( \Phi = 1.383 \)
**Entropy: From the ancient classless society to modern stratified societies (6)**

- **Power of the largest prime movers in different eras**


![Graph showing the power of the largest prime movers over time](image)

- An entropic view of the distribution of wealth through history

  - $\mu = 2$
  - $A = \infty$
  - $\Omega = \infty$
  - $b = 1$
  - $\Phi = 1.693$
In prehistoric societies the wealth could be measured in terms of the available energy per person. We identify this with the income $x$ and we set $x = 0$ to represent the energy that assures covering of the basic bodily energy needs.

With absence of technology (i.e., with technological limit $\Omega = 0$), the value $x = 0$ is the only option and the entropy is $-\infty$. The single allowed option and the $-\infty$ entropy, signifies a classless society.
Most probably, this corresponds to the so-called hunter-gatherer society, admired by the Marxist literature as the ancient communal ownership and aspired also for the future in a modern form.

Apparently, Marx and Engels were faithful to the deterministic scientific paradigm of their era, which they attempted to transplant to history and sociology.
The popularity of their ideas even today reflects the fact that the deterministic paradigm remains quite strong. In it, \textit{entropy has no place}, let alone in its connection with \textit{freedom}.

\begin{itemize}
  \item Primitive classless society
\end{itemize}

\begin{align*}
\mu &= 0 \\
A &= 0 \\
\Omega &= 0 \\
b &= 1 \\
\Phi &= -\infty
\end{align*}
At later stages, as knowledge was developed, the technological limit $\Omega$ was increased. The mean wealth $\mu$ also increased. It is plausible to assume that the rate of increase of $\mu$ always followed that of $\Omega$.

For example, at some phase of the development, the mean wealth $\mu$ was half the technological limit $\Omega$.

At this phase, entropy maximization suggests that the wealth was uniformly distributed among people. **Uniformity means that poor and rich were equally probable—not that the wealth is equally distributed among people.**
Entropy in antiquity

- When the technological limit increased more than twice the mean wealth, entropy maximization suggests that the distribution became (bounded) exponential.
- This means that the poor are more and the rich fewer—but richer than before.
- This diagram shows an intense stratification, which has been observed in antiquity with societies controlled by elites, and intense inequality which could be translated also as slavery.
Entropy in Middle Ages

- When the technological limit increased further, entropy maximization suggests that the distribution continued to be (bounded) exponential.
- Again the poor are more and the rich fewer—but richer than before.
- This graph appears smoother than the previous, but still with intense stratification (monarchies, feudalism), which were the general motive of society worldwide before the French Revolution (1789) and the Industrial Revolution which followed.
In modern societies, the technology has increased the upper limit of possible wealth of individuals by so much that practically does not put any restriction in entropy maximization.

As a result, entropy maximization of the income distribution of individuals, with the only restriction being the total income of the society (or equivalently, the average income per individual), results in exponential distribution.

Koutsoyiannis and Sargentis, Entropy and Wealth
Entropy in technological revolution

- This appears to be the natural distribution that is connected to a stable economy.

- Careful inspection of the graphs shows that it is not only the richest who become richer as technology evolved. The poor also became fewer and the curves $f(x)$ moved to the right.

- This means that, everybody, on the average, became richer thanks to technological evolution.
Introduction to empirical investigation

Two opposite forces modify the natural (exponential) distribution:

1. An organized society redistributes income and wealth through their transferal from the richer individuals to the purer by means of several mechanisms (e.g. taxation). As a result of redistributions, (a) poverty below a low level is eliminated; (b) the middle class, becomes more populated and amplified; (c) the rich lose income.

2. On the other hand, the actions of economic elites, pursuing a greater share of the community's wealth, tend to modify mostly the income distribution tail, converting it from exponential to power-law (Pareto). The effects of this force on the income distribution are to: (a) increase the number of poor; (b) diminish the middle class; and (c) benefit the richest.

At the same time, the elites advance both the technological limit and the average wealth. Naturally, the advancement of technology and average wealth are the positive side of elites' action.
Data processing for empirical investigation

- For our purpose, we have searched for data given in higher resolution than usually tenths of people’s.
- In the retrieved data, the information about the tail (the very rich people) is missing as the data values end at some level \( c \) with the last bunch of data given as “\( c \) and over”.
- It is thus crucial to find a way to extrapolate the distribution function beyond \( c \) and estimate expectations based on this extrapolation.
- It is consistent with our theoretical framework to assume that beyond \( c \) the following approximation is suitable:

  \[
  f(x) = e^{-x/k+b} \Rightarrow F(x) = ke^{\frac{-x}{k}} = kf(x), \quad x \geq c
  \]

  Algebraic manipulations allows to estimate \( B_g \) for the expectation of any function \( g(x) \), by replacing \( B_0, k \) and \( b \) with their estimates. In particular, for the entropy we have

  \[
  \hat{B}_\phi = \hat{B}_0 \left( 1 - \hat{b} + \frac{c}{k} \right)
  \]
Application to societies’ income distribution (1)

Illustration of the entropic framework using income data from the USA for year 2019; (left) probability density; (right) tail function (probability of exceedance).
Application to societies’ income distribution (2)

Illustration of the entropic framework using income data from Sweden for year 2019; (left) probability density; (right) tail function (probability of exceedance).

μ = 53336 $\, \text{S\,Ek}^{-1}
\sigma/\mu = 1.11
Φ = 11.82
Φμ = 0.94
The failure of “80/20 rule”

- In both cases, the income distribution is consistent with the principle of maximum entropy, and in particular with exponential distribution.

- Yet the effect of the elites is visible, as the distribution tails exceed those of the exponential.

- On the other hand, the data do not support the well-known “80/20 rule”, which is consistent with the Pareto distribution (with a specific value of the tail index).

- Specifically, 80% of the income is not generated by 20% of the population, but by more than 40% thereof, which is fully consistent with the exponential distribution. Interestingly, the “80/20 rule” is often called the “Pareto rule”, but the historical investigation performed reveals that the Italian economist Pareto did not suggest that, but later authors loaded him with things that he never said.
Study of the elites (1)

- The most telling evidence about the type of the distribution and, hence, the appropriate background measure, is obtained by studying the distribution tail. To study the tail, we do not need to examine the entire population, i.e. the entire range of the variable $x$.

- In order to empirically study the tail of income distribution, we used data for the net worth of richest people of the world (billionaires) and the evolution thereof.

Conditional probability of exceedance of the annual income of the richest persons in the world. The income per person was found by subtracting the total net worth of a year from that of the previous year. For the years 2002 (lowest average income) and 2021 (highest average income) exponential and power–law trends, respectively, are also plotted with dashed lines of the same color (where in the left panel the green dashed line for 2002 is indistinguishable from the continuous line).
Based on these data and with focus on the distribution tail, we concluded that the exponential tail is not uncommon, while the Pareto tail appears particularly in anomalous periods.

Impressively, the latest period of pandemic resulted in unprecedent profits of the richest, with a clear Pareto tail.
The elites’ role

- It is reasonable to assume that the economic elites pursue a greater share of the community wealth.

- The means to increase elites’ profits certainly include political power and, more recently, an attitude to control the world. Their endeavor becomes more efficient and acceptable by the society by several means they use, such as by overstating existing or non-existing threats, and then by presenting themselves as philanthropists (e.g. by funding nongovernmental organizations dealing with these threats) and world saviours.

- Apparently, if they succeed in controlling the world, this will decrease entropy and hence delimit freedom. In turn, this will lead to decadence, whose signs are already visible in the Western World.
A toy model for income redistribution (1)

- One of the important roles of a state reflecting an organized society is the redistribution of income and wealth through their transfer from some individuals to others by means of several mechanisms such as taxation, public services, land reform, monetary policies, and other.

- Here we examine one of the mechanisms, i.e. taxation, by means of a simplified toy-model example, which illustrates how redistribution affects the entropy maximizing exponential distribution.

- In our toy model, we assume that the original income $x$ follows the exponential distribution with parameter $\lambda$ (equal to the mean $\mu$) and that the tax rate $p$ increases with the original income, according to the function:

  $$ p(x) = \begin{cases} 
  0 & x \leq x_0 \\
  p_u & x = x_u \\
  p_u \frac{x - x_0}{x_u - x_0}, & \text{otherwise}
  \end{cases} $$

Here $x_0$ is a low value of income, denoting the starting point of taxation, and $x_u$ is a high value of income, beyond which the tax rate takes a constant value $p_u$. 

Koutsoyiannis and Sargentis, Entropy and Wealth
A toy model for income redistribution (2)

- Now we apply the toy model assuming $\lambda = \mu = 1$, $x_0 = 0.2$, $x_u = 1.4$, $p_u = 0.4$, $a = 0.5$, resulting in $\mu_y = 0.32$, $a\mu_y = 0.16$.

- The figure shows the variation of the tax, the income minus tax, and the redistributed income vs. the original income.

- In the right figure, notice that for small income the tax is zero and thus the income minus tax equals the original income, while for large income it equals 60% of the original income.
A toy model for income redistribution (3)

The figure depicts the probability density of the redistributed income in comparison to that of the original income. The differences are that in the redistributed income:

- poverty below the level $a\mu_y = 0.16$ is eliminated;
- the middle class, corresponding to incomes up to the mean, is more populated and amplified;
- the rich lose income.

As a result of these, both the entropy and the coefficient of variation have decreased in the redistributed income; the former from $\Phi = 1$ of the exponential distribution to $\Phi = 0.55$ (or $\Phi_\mu$ from 1 to 0.73) and the latter from $\sigma/\mu = 1$ to $\sigma/\mu = 0.57$. 

Koutsoyiannis and Sargentis, Entropy and Wealth
Discussion and conclusions (1)

- **Social sciences are often contaminated by subjectivity and ideological influences**, which become apparent when examined from distance, in the light of history.

- Here we explore whether the maximum entropy, applied to economics and, in particular, to the distribution of a wealth-related variable, namely the annual income, can give an objective description.

- We show that under plausible constraints related to the mean income, the principle of maximum entropy results in exponential distribution, bounded from above if we consider an upper technological limit, or unbounded otherwise.

- Historically, technology has played a major role in development and increase of the entropy of income. Under current conditions, technology no longer imposes a bounding condition on the economy, yet it remains an important factor in increasing wealth.

Temperatures created by human actions in different eras.

We have shown that the entropy is one of the most misunderstood concepts with rich and diverse interpretations which are continuously debated.

With the transplantation of the scientific term to the colloquial language, the popular imagination has loaded “entropy” with almost every negative quality in the universe, in life and in society.

For example, Thesaurus.com lists as its synonyms the words breakup, collapse, decay, decline, degeneration, destruction, worsening, falling apart, while in the site wordhippo.com the synonyms listed, in addition to those, amount to hundreds of words with negative meaning, including deterioration, chaos, havoc, confusion, disorder, disorganization, calamity etc.

Also in scholarly articles, there is no shortage of negative associations.
Discussion and conclusions (3)

- There are historical reasons, why the concept generated so many negative connotations. However, in the end of the 1940s, entropy acquired a clear and universal stochastic definition which is not a related to disorder.

- Furthermore, in the end of the 1950s, it was complemented by the principle of maximum entropy, which lies behind the Second Law, and gives explanatory and inferential power to the concept.

- By now, 60 years after, one would expect that entropy should have aborted the negative meanings, and be recognized as the driving force of natural change and the mother of creativity and evolution.

- This did not happen. Instead, it has been used as a spectre in social sciences, including economics and ecology (where entropy is considered as the waste produced when useful work is done) to promote neo-Malthusian ideas.
Discussion and conclusions (4)

- Entropy maximizing distribution emerges when the background measure has constant density, while if a hyperbolic background measure is used, the resulting distribution is Pareto. Based on real-world data, and in particular, those of the world’s richest, in order to give a better idea on the distribution tail, we conclude that the exponential tail is not uncommon, while the Pareto tail appears particularly in anomalous periods. Impressively, the latest period of the pandemic resulted in unprecedented profits of the richest, with a clear Pareto tail.

- A constant (Lebesgue) density of the background measure is reasonable and that the entropy maximizing, under this measure, exponential distribution is connected to a stable economy.
Furthermore, we examined two different factors, both leading to reduction of entropy and modification of the stable exponential distribution, but in different directions.

1. The organized societies use mechanisms of income redistribution in order to minimize poverty and enhance the middle class.
2. Politico-economic elites try to increase their profits, thus pointing toward a Pareto distribution, which populates more the poor and the very rich and reduces the middle class. At the same time, the elites advance both the technological limit and the average wealth.
Discussion and conclusions (7)

- Using publicly available income data for USA and Sweden, we showed that the income distribution is consistent with the principle of the maximum entropy and in particular, with the exponential distribution. Yet the effect of the elites is visible as the distribution tails exceed that of the exponential.

- On the other hand, the data do not support the “80/20” rule, which is consistent with the Pareto distribution (with a specific value of the tail index).

- Specifically, the 80% of income is not generated by 20% of the population but by more than 40% thereof, in full consistence with the exponential distribution.

- We have shown also a standardized form of entropy can be used to quantify inequality.

- Overall, we have tried to dispel the “bad name” of entropy in social sciences, emphasized its connection with the plurality of options. In addition we have tried to clarify that increasing entropy is associated with increases in wealth.