A Theoretical Model Describing the Lorenz Curve And the Pareto Principle

Bertram Scharpf
info@bertram-scharpf.de

May 2015
Version 1.3.4 (07/05/2017)

Abstract

This paper starts off from theoretical physics. It gives a short outline of the basics of the second law of thermodynamics. It strives for an excessively illustrative presentation of the measure *entropy* and the *Boltzmann distribution*. Then the view moves away from physics, and the previous assumptions will be applied to an economic entity. The Boltzmann distribution evolves to a theoretical formula for an ideal Lorenz curve and to a theoretical value for Pareto's empiric 80/20.

Keywords: Lorenz, Pareto, Boltzmann, probability, distribution, physics, wages, income

JEL categories: C16, C51, E24, Y80

Pages: 11

1. Introduction

1.1. Task formulation

Around the turn of the century 1900 in economics there were discovered two statistical phenomena. First, the US-american Max Otto Lorenz developed a method for measuring and graphically presenting the income distribution of a given country, nowadays named after him the Lorenz curve. A few years earlier, the Italian Vilfredo Pareto discovered that 80% of the land was held by 20% of the population. For both statistical measurements there are yet mathematical models, but those aren't very predicative of the distributions origin.

A quarter of a century earlier, in physics rose a branch that covered distribution problems as well, the statistical thermodynamics. Here, it was majorly Ludwig Boltzmann who figured out excellent solutions. These solutions make an abundance of predictions for phenomena that were detected later.

This thesis sketches out briefly Boltzmanns "distribution of the highest probability method". Then it develops this further to theoretical formulas for the Lorenz curve and for the Pareto principle.

1.2. Motivation

This is not a scientific thesis.

I am not a member of any scientific institute and, besides a prediploma in physics, I do not hold any academic degree. This circumstance I owe, according to my own estimate, lesser to a lack of enthusiasm or to a lack of capability, but to the influence of my family background.

No matter which explanation model may bite here, this thesis stays merely a suggestion and does not claim any scientific acknowledgement.

2. Empiric results

2.1. The Lorenz curve

The Lorenz curve was developed in 1905 by the US-american economist Max Otto Lorenz (1876-1959). It appears in present days in the common offical statistics. The values being measured are the distribution of income.

The incomes are first sorted by amount and then ascendingly summed up. The curve lies in the first quadrant. It starts at (0|0) and it ends at (1|1). The horizontal axis represents the percentage of income recipients from 0% to 100%. The vertical axis represents the summed up percentage of incomes from 0% to the respective recipients percentage.

A straight line from (0|0) to (1|1) would mean a distribution in that every participant has the same income. Typically, the curve is bellied downwards, what means that there are less participants that earn more, and vice versa.

2.2. The Pareto rule

The Italian economist Vilfredo Pareto (1848-1923) noticed that landholding in Italy is distributed lopsidedly. He stated that about 20% is the populations share of holding about 80% of the land. This means vice versa that the remaining 80% of the population hold the remaining 20% of the ground.

In modern everyday life you can experience similar lopsided distributions in other spheres. Here are some examples that I did not measure but just roughly estimate.

- On your computer, in 80% of the time it is switched on, there run 20% of the installed software. The browser and the E-mail client run recurrently; applications for very special purposes you deploy rather seldomly.
- 80% of all Mozart performances play 20% out of the Köchel catalogue.
- 20% of all chess players hold 80% of all achieved Elo points.
- Your local pub makes 80% of its turnover during 20% of the opening time.

3. The distribution of the highest probability principle

3.1. A plain example

Imagine three jugglers on a stage. These jugglers have three balls of the same type, which they throw at each other. What is surely not an artisanal challenge for the jugglers will become a mathematical one for us now.

Say, at every point of time a ball could be definitely declared to belong to a certain juggler. Either he touches the ball or the ball moves towards him. The question that we now want to ask is not, whom of the three jugglers posseses how much or even which ball. We ask which distribution ist the most frequent to occur. Obviously, either of the three has exactly one ball, one of them has all three of them, or one has two, one has the other and one is without. Let's write all possible distributions into a table.

•	•	•	distribution 1-1-1 does occur only 1 time
• • •			distribution 3-0-0 does occur 3 times
	• • •		
		• • •	
• •	•		distribution 2-1-0 does occur 6 times
• •		•	
•	• •		
	• •	•	
•		• •	
	•	• •	
			10 distributions in total

Thereafter, in 60% of the cases the distribution is 2-1-0, in 30% it is 3-0-0 and in only 10% of the cases it is 1-1-1. The distribution of the highest probability thus is 2-1-0.

This example yet is revealing another aspect of the most probable distribution. If this distribution once has occured, it will be left again, if at all, for just a short time. In our example of 3 jugglers with 3 balls the system will return to 2-1-0 after the second throw, of which you can easily convince yourself by playing through the game. For games with more balls and more jugglers one may make the effort of a computer simulation; however, this falls beyond the scope of this thesis.

Before we start the calculus, let's introduce another term: "disorder". What we do is a search for the distribution of the highest disorder. Intuitively, we tend to perceive it as disorder when the balls are distributed best possible, that is when the distribution is 1-1-1. But as the balls are assumed to be undistinguishable, this is the distribution of the best prediction level, therefore of the *best order*. Even the distribution 3-0-0 is better ordered than the "slanted" distribution 2-1-0.

3.2. Formulation of the principle

We call the total count of jugglers $N = 3$ and the total count of balls $M = 3$. The count of jugglers that hold i balls shall be called the occupancy count a_i. The frequency of a set of occupancy counts a_i is

$$P = \frac{N!}{a_0! a_1! a_2! a_3!}$$

Now we write out all the values. First is $N! = 3! = 6$. Then for the different distributions comes out

distribution	a_0	a_1	a_2	a_3	$\prod_i a_i!$	P
1-1-1	0	3	0	0	$0! \cdot 3! \cdot 0! \cdot 0! = 6$	$\frac{6}{6} = 1$
3-0-0	2	0	0	1	$2! \cdot 0! \cdot 0! \cdot 1! = 2$	$\frac{6}{2} = 3$
2-1-0	1	1	1	0	$1! \cdot 1! \cdot 1! \cdot 0! = 1$	$\frac{6}{1} = 6$

Using that we can define the task we approach to as follows.

$$\boxed{\text{We have to search for the } \textbf{distribution } a_i \text{ where the function } P \text{ comes to its } \textbf{maximum}.}$$

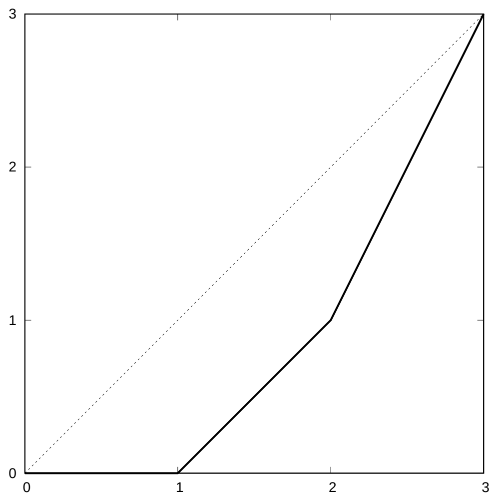
If we wanted to compute the number of all possible states (here 10), we would encounter a so-called partitioning problem. This is anything but trivial. Fortunately, this number is not relevant for our considerations. The value in percent of the whole number of possible distributions is not what we need.

3.3. First Lorenz curve

Before we proceed in searching the maximum, let's get a first taste of a Lorenz curve. We sort the three jugglers by the number of balls they own. Then the first one has none, the second one holds exactly one ball and the third one holds the last two of them. Now we're adding this up.

count	jugglers	balls	sum
0	no one	–	0
1	first	0	0
2	first and second	$0+1$	1
3	all three	$0+1+2$	3

Here comes the plot.



It does not actually look pretty, but maybe this is just the cause that one can quickly see what is the point we will come to.

3.4. The general case

Further on, we are interested in the most probable distribution in a general case of N jugglers with M balls. We are still looking for the maximum of the function

$$P = \frac{N!}{\prod_i a_i}$$

The number of participants N and the number of balls M both stay constant.

$$\sum_i a_i = N \qquad \sum_i i a_i = M$$

Using these constraints we can apply the method of Lagrange multipliers.

What we have to do next is to find the partial derivatives of P with respect to the variables a_i. This leads to several difficulties, because the a_i actually are discrete, and because the derivative of a factorial does not generate a usable equation.

Let's first solve the problem about the product of the factorials. Because the logarithm is a monotonically increasing function, we may not only search for a maximum of P but of $\ln P$ instead. Then the product turns into a sum.

$$\ln P = \ln N! - \sum_i \ln a_i!$$

Besides, this variable is (except for a factor) the widely known and narrowly understood **entropy**. Here's the exact definition.

$$S = k \ln P$$

The factor k will be introduced later; its name is the Boltzmann constant. With W instead of P this equation is engraved on Boltzmann's tombstone.

We proceed passing through the Lagrange method. If you are a mathematician, you have to be strong now: you will not be content how the limits are solved here. Sooner or later, one should dedicate a mathematically well-grounded presentation to this proceeding.

Let's now get to the partial derivatives of the a_i. To get those permitted we have to state another assumption. We suppose we deal with very, very, very many jugglers and also very, very many balls, to be precise the limit of $N \to \infty$ and $M \to \infty$, with M/N staying finite. Then it is allowed to take the a_i as continuous.

We set the parameters λ' and μ' in front of the constraints. The variables are primed because we're still facing several conversions and the parameters are slightly different from those we have to resolve later. Then we demand these derivatives to be zero.

$$\frac{\partial}{\partial a_j}(\ln N! - \sum_i \ln a_i! + \lambda' \sum_i a_i + \mu' \sum_i i a_i) = 0$$

With $\frac{\partial a_i}{\partial a_j} = \delta_{ij}$ (Kronecker delta: $\delta_{ii} = 1, \text{otherwise } 0$) this becomes

$$\frac{d(\ln a_i!)}{da_i} + \lambda' + \mu' i = 0$$

We solve the remaining derivatives using Stirlings approximation for large factorials in a special differential form.

$$\frac{d(\ln n!)}{dn} = \ln n \qquad \text{for} \qquad n \gg 1$$

Applying this we get the solutions

$$\ln a_i + \lambda' + \mu' i = 0, \qquad \Longrightarrow \qquad a_i = e^{-\lambda' - \mu' i}$$

Before we utilise the constraints, we will perform two further steps. First, instead of the very large occupancies a_i, we write the proportions of the whole.

$$p_i = \frac{a_i}{N}$$

Second, we part from the conception of discrete balls that are thrown to and from each other by the jugglers. From now on, we want to think of a row of systems that can contain an arbitrary quantitiy of a good. The theory we are treating here was developed in thermodynamics. There, an "ideal gas" is an entity of particles that can contain energy by their velocity and by their height in a gravity field. Later on, in the case of quantum theory, systems will surface again that take fixed-size portions of energy.

So, we switch from natural numbers i to a continuous ε. The indexed p_i become a function $p(\varepsilon)$. The expected value of the quantity of the good contained by a system shall be indicated by a term that is called "temperature" in physics. This is for reasons of science history, and the measured value T gets multiplied with a factor k, called the Boltzmann constant. The constraints now have the form

$$\int\limits_0^\infty p(\varepsilon)\,d\varepsilon = 1 \qquad \text{and} \qquad \int\limits_0^\infty \varepsilon p(\varepsilon)\,d\varepsilon = kT$$

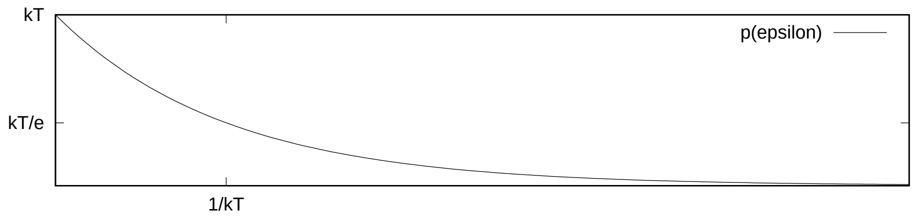
The sigma sign has become an integral that can be solved in a closed form. We again write the Lagrange conditions from the partial derivatives.

$$\ln p(\varepsilon) + \lambda + \mu\varepsilon = 0 \qquad \Longrightarrow \qquad p(\varepsilon) = e^{-\lambda - \mu\varepsilon}$$

The first constraint yields $e^{-\lambda} = \mu$, the second $\mu = \frac{1}{kT}$. Using this, we obtain the Boltzmann distribution.

$$p(\varepsilon) = \frac{1}{kT} e^{-\dfrac{\varepsilon}{kT}}$$

Here's the plotted graph.



4. The ideal Lorenz curve

4.1. Assumption about the income distribution

First we had spoken of balls that were thrown at each other by jugglers. This thought we had transferred to gas particles that are passing kinetic energy amongst them by elastic collisions. Our next step will be so heavily audacious that you will possibly cast doubt on the seriosity of this whole paper. Let's start off trying to see the total of all payable incomes as approximately constant. Let's further assume that the recipients of incomes, again approximately, stay the same persons and the same count. The thing that still is changing is the qualifications of the recipients, their workloads, and the respective popularities of their product. Then we may speak of the incomes as of monetary amounts that move between each two pay days from one pay slip to another.

Of course, this point of view is suitable for causing heaviest contradictions. In conclusion it reduces the individual persons with their painfully acquired abilities down to bare coincidence. That doesn't match with the modern picture of enlighted, consciously acting people that we have become fond of. However, in my opinion this picture itself is nothing less than a heavily simplifying theory.

The considerations taken here cover an imaginary society thats national product is stagnating. But in reality the average standard of living is increasing, as well as the technological evolution, and, last but not least, the life expectancy. Thus, there is enough room for assigning legitimacy to human diligence. Besides that, these considerations only treat material goods. Other aspects of quality of life like family, a clean environment, health, or freedom to travel do not neccessarily relate to high incomes.

There is one assumption that still has to be mentioned. We want to start from an economic order that allows incomes to be freely negotiable. "Free" means here, the labour market is taken to be completely free from influences like access limitations (guild and chamber memberships), inpermissible agreements (corruption, nepotism), and from governmental interferences (minimum wages, fiscal drag). Admittedly this is a model simplified to an extraordinary extent not occuring in reality. Beyond that, it attaches a philosophical explosive meaning to the word "free", because "free" means here both the absence of *outer* coercions and the prevention of dishonest doings *inside* the system.

If we agree to accept all these limitations, we may apply the method of the most probable distribution.

4.2. Derivation from the Boltzmann distribution

We start off from the Boltzmann distribution.

$$p(\varepsilon) = \frac{1}{kT}e^{-\dfrac{\varepsilon}{kT}} \qquad \text{with} \qquad p(\varepsilon = 0) = \frac{1}{kT} \qquad \text{and} \qquad p(\varepsilon \to \infty) \to 0$$

Then we form the inverse function.

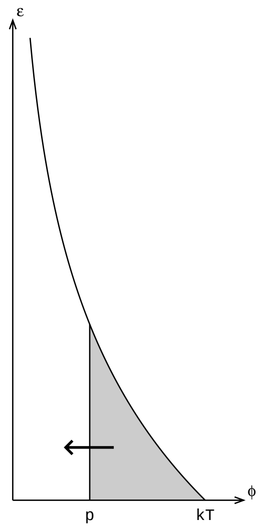
$$\varepsilon(p) = -kT\ln kTp \qquad \text{with} \qquad \varepsilon(p = \frac{1}{kT}) = 0 \qquad \text{and} \qquad \varepsilon(p \to 0) \to \infty$$

This is possible because the probability density is strictly monotonically decreasing.

The thing the statisticians do is, as mentioned above, to sort the incomes by amount and then sum them up. Now we will do the same with our function. We start at the probability density of a zero income, $p = \frac{1}{kT}$, and we integrate the incomes in the direction they increase, hence from right to left (minus!).

$$L(p) = -\int_{\frac{1}{kT}}^{p} \varepsilon(\varphi)d\varphi = -\int_{\frac{1}{kT}}^{p} (-kT\ln kT\varphi)\, d\varphi \qquad \text{with} \qquad \frac{1}{kT} \geq p > 0$$

And this is how it looks in the plot:



The value of L is increasing here with falling p. We will express the turn of the direction later. First, we take into consideration that the proportion of the total probability is seen as percentage, and we replace p by a ρ that is running from 1 to 0.

$$\rho = kT\varphi \qquad \Longrightarrow \qquad L(\rho) = \int_1^\rho \ln\rho\, d\rho \qquad \text{with} \qquad 1 \geq \rho > 0$$

Then the integral is easy to solve:

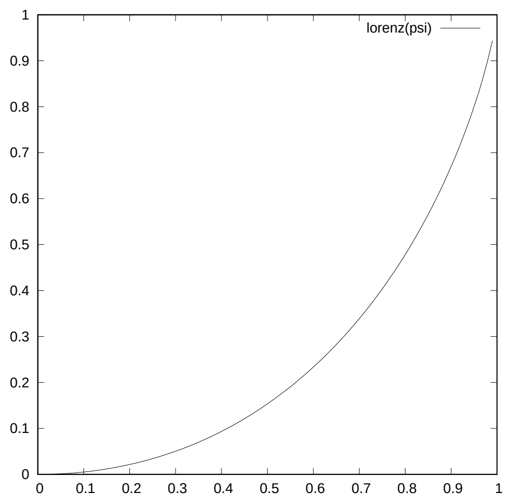
$$L(\rho) = \rho\ln\rho - \rho + 1 \quad = 1 - \rho(1-\ln\rho) \qquad \text{with} \qquad 1 \geq \rho > 0$$

Now let's change the point of view so that the counting of the percentage starts at small incomes.

$$\psi = 1 - \rho$$

$$L(\psi) = 1 - (1-\psi)(1-\ln(1-\psi)) \qquad \text{with} \qquad 0 \leq \psi < 1$$

The graphical plot shows amazingly pretty the picture that is know from the empirical measurement.



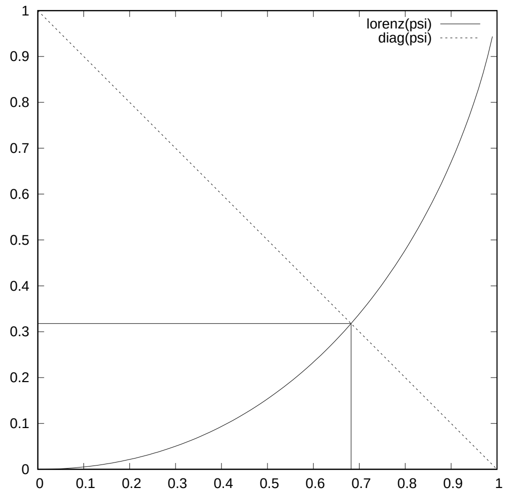
This curve has one further interesting aspect. The temperature has vanished from the formula and from the plot. The curve will always stay the same, no matter how high the expected value of the income of an individual may rise. Independently from the prosperity of a society there will alwas exist a majority with small incomes and a minority with high ones. The reason for this is not social injustice but simply that it is the distribution of the highest probability.

4.3. Search for the Pareto intersection

We state that the value 80%/20%, measured by Pareto, issued from an economy that did not fulfil roughly the idealised prerequisites we work with here. It is an empirical value that has to be different from the theoretical value because reality creates effects that could not be considered in our theory. So there remains the search for the ideal 80-20 value, thus the solution of the equation:

$$L(\psi) = 1 - \psi$$

In the plot:



Obviously we can write that shorter if we go back to ρ.

$$1 - (1 - \psi)(1 - \ln{(1 - \psi)}) = 1 - \psi$$

$$1 - \rho(1 - \ln \rho) = \rho$$

So we look for the zero of a function.

$$1 - \rho(2 - \ln \rho) = 0$$

This equation is not solvable in a closed form. A numerical approximation results in the about value:

$$\psi \approx 0.682155567100627, \quad L \approx 0.317844432899373$$

According to this we may say, if the Lorenz curve has reached its ideal shape, 68.2% of the labour forces draw 31.8% of the total income payments and vice versa.

This is a clearly more peaceable value than the one identified by Pareto. As we started from the assumption of an idealised market that is substantially more free than the realty market of Paretos times, this is a plausible result.

5. Conclusion

5.1. Result

A result of these calculations which is lesser important is that the bellied curve is roughly near the measured Lorenz curve. This could be reached by other approaches, too.

The important result is the message that this theory contains: Politicians on the soapbox perpetually complain that the gap between rich and poor is getting wider and wider. If the underlying explanation model, however, is permissible, this "gap" between rich and poor admittedly will not become narrower spontaneously from a certain point on, but most notably it will *not* get wider and wider automatically. If the "gap" between rich and poor should be observed to get wider in spite of this, the reasons have to be searched where the free market is limited. Artificial redistributions are counteractive.

The value of 80/20 that was supposed by Pareto for the property distribution draws a picture of the world that is more dramatic than how the world could be naturally. A distribution of 68/32 still is anything but cheerful, but it is a value more contenting by far.

A subject that deserves distinct treatment is the minimum wage. If you force the Boltzmann distribution to be zero on the far left, it will not shift itself to the right, but a peak will arise at the very start: the number of unemployed people will grow.

Beyond this, in my opinion, it is worth to deliberate further about the terms of freedom that were mentioned while scaffolding the model. They will allow to form guidelines for a prospering economic policy.

5.2. Future Prospects

We have treated the development of humankind in an elementary part of community life. This unravels further questions about long-term effects. If the principle of the most probable distribution is applicable to an economic evolution, one may be tempted to consider whether it could be applied to the whole technical evolution or even to the evolution of life altogether.

The principle of the most probable distribution remains to be one of the most exciting adventures in science and in philosophy.

5.3. Download sites and License

This document is available online. There is a German version and one in English. The English version is further available in letter paper format instead of A4. Currently these can be downloaded from two servers.

```
http://bertram-scharpf.de/books/bscharpf-lorenzpareto.pdf
http://bertram-scharpf.de/books/bscharpf-lorenzpareto-en.pdf
http://bertram-scharpf.de/books/bscharpf-lorenzpareto-en-letter.pdf
```

```
http://bertram-scharpf.spdns.de/books/bscharpf-lorenzpareto.pdf
http://bertram-scharpf.spdns.de/books/bscharpf-lorenzpareto-en.pdf
http://bertram-scharpf.spdns.de/books/bscharpf-lorenzpareto-en-letter.pdf
```

The source code is available as well.

```
https://github.com/BertramScharpf/lorenzpareto
```