

Scrutinizing Distributions Proves That IQ Is Inherited and Explains the Fat Tail

Michael Grabinski, Galiya Klinkova

Department of Business and Economics, Neu-Ulm University, Wileystr, Neu-Ulm, Germany Email: michael.grabinski@uni-neu-ulm.de, galiya.klinkova@uni-neu-ulm.de

How to cite this paper: Grabinski, M. and Klinkova, G. (2020) Scrutinizing Distributions Proves That IQ Is Inherited and Explains the Fat Tail. *Applied Mathematics*, **11**, 957-984.

https://doi.org/10.4236/am.2020.1110063

Received: September 4, 2020 Accepted: October 16, 2020 Published: October 19, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). http://creativecommons.org/licenses/by/4.0/

C ① Open Access

Abstract

The motivation of this paper is to show how to use the information from given distributions and to fit distributions in order to confirm models. Our examples are especially for disciplines slightly away from mathematics. One minor result is that standard deviation and mean are at most a more or less good approximation to determine the best Gaussian fit. In our first example we scrutinize the distribution of the intelligence quotient (IQ). Because it is an almost perfect Gaussian distribution and correlated to the parents' IQ, we conclude with mathematical arguments that IQ is inherited only which is assumed by mainstream psychologists. Our second example is income distributions. The number of rich people is much higher than any Gaussian distribution would allow. We present a new distribution consisting of a Gaussian plus a modified exponential distribution. It fits the fat tail perfectly. It is also suitable to explain the old problem of fat tails in stock returns.

Keywords

Fat Tail, Income Distribution, Chaos, Finance, IQ

1. Introduction

In finance, economics, and many social sciences distributions are important. However, there are two closely connected puzzling items. Firstly, there is an almost dogmatic assumption that there are Gaussian distributions only (with few exceptions). Secondly, there are partly strange methods to prove that something must have a Gaussian distribution. The mathematics of distributions is essentially a product of the 19th century, for an overview consider e.g. [1]. Since 100 years one can read in a *textbook* [2] on page 179 "*Everybody believes in the exponential law* [*i.e.* Gaussian distribution] *of errors: the experimenters, because they think it can be proved by mathematics, and the mathematicians, because*

they believe it has been established by observation". In non-mathematical sciences distributions became popular from roughly 1950+.

A typical paper from this time is [3]. Fama observed a "fat tail" in the distribution of stock market prices. This fat tail provoked an avalanche of publications which are impossible to mention completely, just as an example see [4] [5] [6]. What is most puzzling with it, is that the authors of [3] [4] [5] [6] (and many more) wrongly assumed a Gaussian distribution in the first place. Ignoring textbooks like [2] they "derived" the Gaussian distribution from a misinterpretation of the central limit theorem and assumed ergodicity without any justification. Even worse, once having accepted this derivation as correct, observations of stock prices and the like showing a fat tail are in contradiction to a Gaussian distribution. From this one has an experimental proof that some assumptions of Fama [3] must be wrong. Instead of starting all over, the fat tail is cherished as one of the greatest discoveries of the 20th century in finance.

Another reason for the wrongly assumed fat-tail in finance is that stock prices are fluctuating chaotically rather than randomly. The mathematical description of chaos is pretty old, and a modern summary (and application to physics) can be found in [7]. The application of chaos to business and economics is much younger, just as an overview see [8]-[15]. The distinction relevant here between randomness and chaos has been shown quite recently [16].

The name "fat-tail" originates from physics. However, the wrong doings mentioned above are not present there, for a quite recent example see [17].

Dealing with distributions other than Gaussian does rarely cause problems similar to the one mentioned here. As stated above, especially social sciences assume a Gaussian distribution for almost everything. Only if there is a proof for another distribution, it is not used. As an example consider the exponential distribution mostly applicated to describe queues [18] [19] [20] [21]. An already rarer example is the Poisson distribution [22]. There are also power law distributions but these are most common in physics within e.g. critical situations like phase transitions [23]. An application from physics (critical points) for herd behavior in financial markets can be found in [24].

So far for a brief summary of the use of distributions especially in non-mathematical sciences, the purpose of this paper is not to fix the mentioned problems especially when using a Gaussian distribution wrongly or without justification. This is hardly possible. Our goal is to explain the use of distributions in two general situations:

- An experiment or reality shows a strictly known distribution (e.g. Gaussian).
- An experiment or reality shows a strange situation (e.g. Gaussian with fat tail).

To give a complete answer to the two points would mean writing a textbook. Of course, such books have existed for roughly 100 years, see e.g. [2]. As stated above, they are rarely used in fields from economics over finance to psychology. Therefore we take just two examples which are discussed quite often.

Our first example is the distribution of the intelligence quotient normally re-

ferred to as IQ. There are lots of data worldwide, and they have one thing in common: They are perfectly Gaussian (unfortunately without fat tail). From this it is almost trivial to show in chapter 2 that the IQ is inherited or at least not changed by conscious action like training. This gives the first mathematical proof for an old and almost religious argumentation about nature versus nurture.

In chapter 3 we will scrutinize income distributions. We have chosen this example for two reasons. Firstly, it is an often debated subject especially since the best-selling book of Piketty [25], though we will not contribute to the political or moral debate of it. Secondly, income distributions also show something like a Gaussian distribution with a fat tail. In subchapter 3.3 we will derive a new model by using the remarks of Chapman [26] on historical data. It is also (part-ly) an explanation of the fat tail in finance. Though the general idea is identical, the effects of speculation make finance more complicated [27].

As a result we get a narrower distribution for the not-very-rich if the super-rich are allowed to have a wider distribution. In other words, without the super-rich there would be a less equal distribution within the "normal" people.

Fitting the data within subchapter 3.3 is extremely complicated. It shows the frontiers of numerical mathematics. Therefore we are deferring some of the mathematical derivations to chapter 4.

Chapter 5 gives a summary and ideas for further research.

2. What IQ Distribution Teaches Us

In **Figure 1** you can see the probability density for an IQ of German men (wide blue curve) and women (narrow red curve). Both curves peak at IQ = 100leading to an average IQ of 100. This is not identical but very similar in most developed countries. Sometimes the average IQ is higher but sometimes it is lower. In most third world countries it is even significantly lower. Because there is less education in these countries, some may argue that this is a proof for IQ being based on nurture. However, it could also be nature, as most of the inhabitants of a country live there for generations. So it is neither a proof nor a disproof.

Even much more universal than the average IQ in the developed world is the width of the distribution. It is wider for men than women. Again, this is no proof for nature or nurture. However, setting on nurture would mean that there is a *universal* difference in the education of boys and girls. Though there are differences in education, it would be at least puzzling that this difference is persistent in so many societies.¹

¹The different width in female and male IQ distribution is quite plausible in terms of classical evolution. At least during most of human history the roles of men and women were clearly defined. Men had to hunt and gather in order for a family to survive. Women had to take care of raising the kids. Especially in hunting there may be two successful strategies. Put most of your energy into your mussels or into your brain. (The brain uses a lot of energy) Raising kids a low IQ will maybe kill the kids. On the other side, a high IQ will not help very much but consumes much energy. Thus stone age women should have an almost constant average IQ.



Figure 1. Gaussian distribution of IQ of men ($\sigma \approx 16.2$) and women ($\sigma \approx 13.2$).

Though this paper is on mathematics rather than psychology, there is meanwhile an agreement under academic psychologists that IQ is inherited. In a review [28] one can read "... *IQ is fixed throughout your life: the only way you'll lose it is because of a brain injury*". One can find only small and rare deviations from it as stated in [29]. But this is no disproof. Especially, there is no recipe how to increase IQ, which would be necessary to prove nurture. There are even hints that it is epigenetic as stated in [30].

Nevertheless, there are *hardliners* (even in academic psychology) sticking to nurture instead of nature. Because the authors of this paper cannot read minds, we can only speculate about the reason for it. It looks like *ideology takes over science*. Of course, an inherited IQ has political consequences. It would be much harder to argue for a more equal income within countries and especially between developed countries and the rest of the world where average IQ is partly significantly lower. In order to find counter arguments against the mainstream of academic psychologists it is sometimes said that IQ is a poorly defined measure. One has to say that there are lots of advances ever since Binet and Simon developed the first IQ tests in the early 20th century. There are also broader measures like "fluid intelligence" which include the ability of abstract thinking and problem solving. Furthermore, many other data like the scores of Americas standardized entrance exams for university (e.g. GMAT) correlate nicely to the IQ. So we have extremely many sets of data which are all seeing nature instead of nurture.

Sometimes it is also stated that one can *learn* how to get a higher score in an IQ test. It would make the entire concept ridiculous. It is even easy to "prove" it by taking an IQ test several times. However, this is nothing but *corrupting the system*. Taking the same or a very similar math exam many times will also improve results.

Up to now we have just summarized about nurture versus nature in IQ. The motivation for this chapter was to find a clear-cut *mathematical* proof that IQ or

fluid intelligence is (essentially) inherited. However we are not sure whether it will convince the above-mentioned *hardliners*. To our own surprise it is extremely simple.

If IQ increases by training, it should be identical to learning in the sense of learning curves. Learning curves are used in (industrial) engineering and especially production management. Maybe first noted in [31], one can find the same approach in contemporary PhD theses like [32] or even applied to the learning of terrorists in a paper in Science [33]. Such forms of learning over time *t* show a decrease in e.g. cost in the form of

ť

$$-\alpha$$
 (1)

where α is a critical positive exponent. Purely numerical it fits most situations fairly well. Though widely used, it is wrong. Equation (1) is a result of a random walk [34]. It fits for unconscious learning only. This is like ants "learn" how to find food. It is also called trial and error. Obviously it does neither apply to learning how to produce cheaply nor to learning in order to increase IQ. Describing (conscious) learning correctly has been published surprisingly recently [35]. It has been extended to situations where two sides learn ([11] [36]) like in terrorist attacks where the aggressor learn how to attack and the protector how to defend.

Instead of Equation (1), learning takes the following form:

$$e^{-t/\tau} - 1 \tag{2}$$

where τ is a typical learning time. If IQ is essentially acquired by "learning," one would have the same picture in the IQ distribution. Proportionally to how much IQ points you have already gained, it will be more and more difficult to get an additional IQ point. We have a differential equation of the form $I\dot{Q} \propto -IQ$ leading to an exponential distribution of IQ:

$$p(IQ) = \lambda \cdot e^{-\lambda \cdot (IQ - IQ_0)}$$
 with $IQ \ge IQ_0$ (3)

where λ is a parameter determining the quality of the education program. A small λ means intense education for everybody and a big λ means no education. IQ₀ is the IQ at birth which may have a very narrow Gaussian distribution. The plot in **Figure 2** reveals the difference.

The difference between an inherited distribution of IQ (Figure 1) and a trained one (Figure 2) should be obvious. From this we have the clear result that reality (experiment) proves that IQ is essentially inherited. One of the most striking difference between Figure 1 and Figure 2 is the universal constant IQ_0 .

Some may say that there is both: nurture and nature. Of course there may be a Gaussian distribution of IQ_0 and some IQ achievements due to education. However the decision whether nature or nurture has (by far) the upper hand is easy. Does our measured IQ distribution look more like Figure 1 or Figure 2?



Figure 2. Exponential IQ distribution; $IQ_0 = 90$ and $\lambda = 0.1$ would mean $\langle IQ \rangle = 100$.

The above-mentioned hardliners may say that the skills to educate (*i.e.* λ) is Gaussian distributed within the parents. But this would still not lead to a Gaussian distribution. Though the distribution for high IQ values would look almost Gaussian, most strikingly this new distribution would still be very asymmetric and never show IQ < IQ₀ values in contrast to observations in the real world. Of course it is possible to destroy IQ either by physical or mental injury. As an example of the latter one, one may consider the tragic figure Kaspar Hauser² living in Germany around Nuremberg from about 1812 till 1833. However, there is never so much abuse to explain symmetry.

If it is possible to increase IQ (massively) by education or training, such methods could be applied to a centralized child education. It should have led to a massive increase in IQ in systems like the Soviet Union or mainland China.

All the above clearly states that IQ cannot be improved. It is a result of a random mix of the genes of the two parents. Due to the central limit theorem such randomness leads to a Gaussian distribution. It is very hard to imagine that any other mechanism creates a Gaussian distribution. Actually, mankind is not able to create (complete) randomness by e.g. computer programs. Astrophysicists are sometimes in need to scrutinize signals for *exact* randomness. They still have to rely on natural sources like radioactive decay in order to have a precise reference. Therefore the Gaussian distribution proves randomness and no conscious actions.

Some may argue now that the IQ is not inherited but a result of randomness ²He appeared in the Nuremberg area in the age of about 16 and could hardly speak. Obviously he had been a mistreated or severely abused child. He learned to speak and told people that he was held in a dark room by water and bread for almost his entire life though this cannot be true. There were also rumors that he was aristocratic and betrayed of his inheritance. However, a genetic analysis of 2002 disproved it once and for all. With learning to speak fluently he invented stories in order to gain the interest of other people. In some sense it shows that his IQ was present all the time and its usage had been recovered.

having nothing to do with nature or nurture. It would also lead to a perfect Gaussian distribution. However, there are correlations between the parents' and offspring's IQ. These can only be there if either nature or nurture plays a major role.

However it is impossible to judge whether this randomness is finished by conception. Something during the embryonic growth may contribute. At least for the trait homosexuality in women there seems to be strong evidence for it. Early childhood may also have an influence on IQ as long as it cannot be influenced consciously. To judge whether such unconscious influences exist is impossible to decide because the statistics are identical³.

In breeding animals genetic selection is meanwhile quite common. It is done by producing many embryos from one pair of parents in a Petry dish. The genes are scrutinized in order to find e.g. the embryo with the highest potential for a cow giving lots of milk. In humans such selection should theoretically be possible too. Though the genes for high IQ are not discovered yet, it will be possible someday. If done massively (albeit moral and ethical concerns), it would lead more and more to an exponential distribution of IQ. If done only by the rich who can afford it, it would lead to a mixture which is a Gaussian with a fat tail. It would be essentially the same model as we will suggest in subchapter 3.3.

3. Income Distributions

As a start consider the monthly net household income of 2017 in Germany as given in **Figure 3**. One may give this income distribution to graduate students (as we did) to find the best fitting Gaussian of it. Though it is a god math exercise, it does not make very much sense. The distribution is given. The necessary row data are arbitrarily fine. So why should one find a mathematical function which is at most a good approximation? As in the last chapter the goal is to learn something from the form of the distribution. From the exact Gaussian distribution of IQ one can derive that IQ is inherited only. Find a descend fit of **Figure 3**, and it may reveal why it comes to a difference in income.

Though this paper is on mathematics rather than economics, it is an interesting question especially in the age of globalization and Piketty's book [25] of popular science and also earlier [38] research. Please note that it is not only debated how narrow or wide an income distribution should be. Looking into the details, it is even not clear how uneven the distribution is [39]. Most scrutinized (and envied) are the super-rich as shown in e.g. [39]. However, their total wealth ³There is however at least weak evidence that both factors contribute. In all societies infidelity is considered bad. If nurture (even in the womb) is essential for IQ only, it would be uninteresting who impregnates whom (except for the pleasure). If it is purely nature, just the maid and not the companion should be selected. If both have some effect, it would be good if maid and companion are identical. So high IQ parents could pass the high IQ more likely to their children while infidelity would lead to a much narrower IQ distribution. Please note that such trait as fidelity can also be "inherited." It is known as cultural evolution, and is much less scrutinized than natural selection. It has been e.g. explicitly proven that in societies were female genital mutilation exists but not mutilated women are accepted too, there mutilated women have more kids so that this "culture" unfortunately spreads [37].



Figure 3. Distribution of net monthly household income in Germany 2017. Data from statista.com.

varies very much over the years. Furthermore, the ratio of income from wealth to work is often overstated.

Here we will consider income only, be it from wealth or work. Within our accuracy goal this distinction is unimportant. We will also always go for net income. Of course the net income depends on the political system and things like minimum wages and social support. On the other hand, people try to increase their net income. Again within our accuracy it does not matter very much.

In subchapter 3.1 we will quickly state how values like the one from **Figure 3** are classically fitted by a Gaussian distribution. Choosing a Gaussian is consistent with the results from chapter 2. The IQ is strictly Gaussian distributed. Other characteristics like strength or health are also Gaussian. In a just world income should therefore be distributed Gaussian. And as long as there is a free labor market without frictions (unfortunately not existing in the real world), justice will appear automatically. Redistribution may benefit the poor too little, but it will not enhance the income of the rich.

We will show that the classical approach using the mean and standard deviation is wrong for principle reasons. A least square fit or better least absolute value fit [16] gives different numbers. However, both ways cannot explain the existence of households with a monthly (net) income of \notin 10,000 or more. Though these are rare they exist in numbers unexplainable by any Gaussian distribution. And there are quite a few households having such income purely from labor. Even deviating from a Gaussian in a way that negative incomes are impossible does not fix the problem.

In subchapter 3.2 we explain why one should not take values from the columns of **Figure 3** for a fit. Such ups and downs come from tax rates and social support up to a certain income, etc. Anyway, it is completely uninteresting to get a most precise fitting mathematical function of **Figure 3**. It will be a polynomial of n^{th} degree were nothing can be learned from. But even with the results from subchapter 3.2, households of a monthly income of \notin 10,000 and over should not exist.

In subchapter 3.3 we will construct a new model by using the historical results of Chapman [26]. It is essentially a mixture of a Gaussian and exponential distribution. In doing so one will get quite some households with a monthly income of \notin 10,000 and over.

3.1. The Classical Gaussian Fit

Finding a Gaussian describing **Figure 3** is usually considered trivial. Just calculate the mean and the standard deviation in order to get μ and σ of the Gaussian distribution for the distribution p(E) of the net earnings

$$p(E) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$
(4)

Please note that we have only an income range in each column of **Figure 3**. This is particularly difficult for the first and the last column in **Figure 3**. They have (theoretically) average values within the intervals $(-\infty, 900 \notin)$ and $[6000 \notin \infty)$, respectively. Of course, the raw data leading to **Figure 3** will reveal the true averages. As an assumption one may take 450 \notin for the first and 7000 \notin for the last column. For all the others it is fair to assume the middle of the respective interval as the average. In doing so one will find

$$\mu \approx 2664.03 \notin \sigma \approx 1686.12 \notin$$
 (5)

Though these results are pretty simple to get, a least square fit of **Figure 3** with a Gaussian should be the more general choice. It would be the only way for a fit with an arbitrary distribution. Here one must minimize

$$\left(\frac{3792}{40205} - \int_{-\infty}^{900} dE \, p(E)\right)^2 + \dots + \left(\frac{2276}{40205} - \int_{6000}^{\infty} dE \, p(E)\right)^2 \to \min \qquad (6)$$

with respect to μ and σ contained in p(E) of Equation (4). The (numerical) solution of the minimization yields:

$$\mu \approx 2275.75 \notin \sigma \approx 1302.15 \notin \tag{7}$$

As one sees there is quite some difference between the values in Equation (5) and Equation (7). Please note that this has nothing to do with the assumption of $450 \in$ and $7000 \in$ average of first and last column, respectively. It is easily possible to show that *any* assumption for the averages of the first and last column of **Figure 3** will not make the values in Equation (5) identical to the one of Equation (7). In order to get identical values in Equation (5) and Equation (7) one has to assume that the averages of the first and last column are complex: $333.929 \notin \pm 1072.79 \notin i$ and $333.929 \notin \mp 1787.35 \notin i$ This is of course nonsense. With the imaginary parts even bigger than the real parts, one cannot speak of a small deviation due to the numerical calculation.

There is obviously a mistake in finding μ and σ in a Gaussian distribution by using the mean and standard deviation of the given data, even if the raw data (not clustered) were used. And this mistake can be quite big. We stress it here because this (wrong) procedure is standard for finding values of μ and σ in most non-mathematical sciences. The reason behind it is quite simple. If the given data are for sure *exactly* Gaussian, it is correct to assume that μ equals the mean and σ the standard deviation. However, this is something which will (almost) never be the case. The standard deviation is a non-linear function of the data. Although approximately Gaussian distributed data will be nicely fitted by a Gaussian, the standard deviation of these data is not necessarily an approximation for σ . It can be quite different as this example shows. Though the mean is a linear function of the data, it will not be identical to μ either. This has to do with the fact that σ and μ cannot be fitted independently.

Just for completeness we note that the least square fit is an approximation only as has been shown in [16]. Taking the squares yields positive values but it is arbitrary. Why not taking the fourth or sixth power? Correctly one has to take the absolute value:

$$\frac{3792}{40205} - \int_{-\infty}^{900} \stackrel{\epsilon}{dE} p(E) + \dots + \left| \frac{2276}{40205} - \int_{6000}^{\infty} \stackrel{\epsilon}{dE} p(E) \right| \to \min$$
(8)

This minimization is numerically quite challenging. Maybe that is the reason why the (wrong) least square fit and not the (correct) least absolute value fit is normally used. Of course, in many cases least square fit and least absolute value fit will lead to similar results. However, here it is not the case. Though numerically tough, it is a well-defined problem with a unique solution. For our values we will get:

$$\mu \approx 2096 \notin \sigma \approx 1228 \notin \tag{9}$$

The deviation of Equation (9) from Equation (7) is far from being negligible. And Equation (9) is even more different from Equation (5) than Equation (7). Though it is not the main part of this paper, we have two statements especially for non-mathematically sciences using statistics:

- Taking the standard deviation for *σ* and the mean for *μ* to fit a Gaussian distribution like in Equation (4) is generally wrong.
- The least square fit is an approximation only. The correct least absolute value fit will lead to quite different results especially in non-linear fits where the data vary over orders of magnitude.

Some critics might say that our Gaussian approach is faulty from the beginning. This is because a Gaussian distribution runs from minus to plus infinity. And negative incomes are impossible. Please note that this is always the case because nothing runs from minus to plus infinity. The IQ shows a perfect Gaussian distribution though there is no negative IQ. With income it is not absurd to assume negative values. With e.g. very low IQ and/or very poor health it is not possible to survive without support from the community which is nothing but a negative income. But be it as it may, of course one can start with a Gaussian running from zero to infinity. Because it needs a new normalization, Equation (4) will read now

$$q(E) = \frac{1}{\sigma\sqrt{2\pi} \cdot \left(1 + \operatorname{erf}\left(\frac{\mu}{\sqrt{2\sigma}}\right)\right)} \cdot e^{-\frac{(E-\mu)^2}{2\sigma^2}}$$
(10)

where $\operatorname{erf}(x)$ denotes the error function defined as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} dt \, e^{-t^{2}}$$
 (11)

within this approach it is also possible to get μ and σ from the mean and standard deviation. However, the mean and standard deviation are given by Equation (25) and Equation (26), respectively. Now we have to solve two coupled non-linear equations:

2664.03 €=
$$m(\mu, \sigma) \wedge (1686.12 €)^2 = s(\mu, \sigma)^2$$
 (12)

The solution of the couple Equation (12) is possible numerically only. As a result one will get

$$\mu \approx 1946.86 \notin \sigma \approx 2180.27 \notin (13)$$

Please note that getting μ and σ this way is incorrect for the same reason as the result in Equation (5) is wrong.

As stated above the correct way finding μ and σ is a least square fit. It takes the form

$$\left(\frac{3792}{40205} - \int_{0}^{900} dE \ p(E)\right)^{2} + \dots + \left(\frac{2276}{40205} - \int_{6000}^{\infty} dE \ p(E)\right)^{2} \to \min$$
(14)

The solution can be obtained numerically only:

$$\mu \approx 2104.98 \notin \sigma \approx 1620.90 \notin \tag{15}$$

As explained above and in [16], the least square fit is at most an approximation. To be precise one has to use a least absolute value fit. *I.e.* one has to solve the following:

$$\left|\frac{3792}{40205} - \int_{0}^{900} dE \, p(E)\right| + \dots + \left|\frac{2276}{40205} - \int_{6000}^{\infty} dE \, p(E)\right| \to \min$$
(16)

A numerical solution yields:

$$\mu \approx 1625.60 \notin \sigma \approx 1707.45 \notin (17)$$

This last result can be considered "exact" within our fit procedure. In the next subchapter we will learn that the fit procedure does not necessarily give a result from which one can learn something.

3.2. How to Fit Income Distributions

With Equation (17) we have the best possible Gaussian fit for the distribution of

Figure 3. The question is whether it is useful? The absolute value fit is a Gaussian distribution looking most similar to the given distribution in Figure 3. However, there may be some ups and downs in the distribution due to tax rates, minimum wages or social benefits. These are uninteresting here. We want to show whether our assumed Gaussian distribution will be in accordance to global measures such as Gini or median. One may think of many factors useful for this purpose. But they should also be available in standard data banks. For OECD countries many data are available for free in OECD. Stat which can be found in the internet under stats.oecd.org. From this we gathered data for four different countries. All are based on net monthly income per household in € at German PPP. (This is our contribution against the \$ dominance) The first two columns in Table 1 (number of households and mean income) are not statistical measures. They are gauge factors. As we always consider income per household, it would be smart divide any income by the particular mean income. It would lead to a mean income of 1 in every country. One would also get rid of any exchange rates or PPP. However, we avoided this approach in order to have results in real currency units which might be convenient for many readers especially economists.

The last two columns in Table 1 (median income and P90/P10)⁴ are the statistical measures we have chosen. The median is the "middle income." It is the income to choose if one has only one number instead of the entire distribution. Some people take the mean as an alternative measure for it. Why this is wrong can be found in [16]. Additionally or as an alternative, it would be good to consider the Gini coefficient in addition or instead of the median. However, including the Gini makes it numerically extremely complicated here. Especially for the extended model of subchapter 3.3 the authors were unable to perform the necessary numeric calculations. For the reason behind it please see Chapter 4. The P90/P10 ratio is a measure for the rich. Were the distribution completely symmetric like an ordinary Gaussian of Equation (4), the ratio would be one. We would love to have something like a P99/P1 ratio. But we did not find such numbers in free data banks.

One might argue that the standard deviation is also a global measure. So fitting with it like in subchapter 3.1 should not be a bad idea. Firstly, we have to

 # of howerholds	moon in como	madian in some	D00/D10

Table 1. Data from OECD.Stat 2016 taken from stats.oecd.org on 13/02/20, all at GER

	# of households	mean income	median income	P90/P10
GER	4,0749,525	2148.83€	1888.50 €	3.7
USA	126,220,000	2957.65 €	2394.29 €	6.1
UK	27,800,000	1915.89€	1567.63 €	4.2
DK	2,686,035	2213.79€	2006.73 €	2.9

⁴Most quantities like mean or median should be well-known. The P90/P10 ratio is the ratio of the income which is lower for 90% of the people to the income which is higher for 90% of the people. More can be found in Chapter 4, Equation (32) and Equation (33).

note that the standard division is a quite complicated expression as given in Equation (26) or even more complicated as indicated below Equation (26). Secondly, the standard deviation has a very limited meaning here. In measurements like the mass of an elementary particle one will expect *one* value. In several measurements one will get different results though they should be equal. It does make sense to build a mean. And one should test whether the measured data have a Gaussian distribution. (If not, something is systematically wrong) The Gaussian distribution should be narrow if the measurements are accurate. As a measure for accuracy one may take the easily obtainable standard deviation. However, in income distributions such parameter does not make sense at all. There is a reason why we have an income distribution. It is not an error in measurement. Here we assumed that it has to do with the distribution of skills such as IQ. The deviations will teach us (in subchapter 3.3) what other effects rather than skills contribute.

In an extreme socialist country, it may be stated that everybody should have the same income as ordered by a socialist income committee. In such country it would be a reasonable idea to measure the real income. The mean should be the value set by the committee and the standard deviation tells how good the socialist ideology has been implemented. This shows another misunderstanding of statistics in non-mathematical sciences. But as stated in the very beginning of the introduction, 100 years old books like [2] have never found readers there.

The fitting with median income and P90/P10 from **Table 1** must be at least an absolute value fit in accordance with Equation (16). However, as the quantities of **Table 1** have a different dimension (\in and no dimension) it is formally impossible to just minimized a sum of it. Depending on the chosen dimension (e.g. \in , Cent) the result will be strikingly different⁵. Therefore we have to take the *relative* deviation. Furthermore, the exact mean is a constraint. Put this together we have to minimize the following with respect to μ and σ :

і і

$$\left|\frac{n(\mu,\sigma) - \text{median}}{\frac{1}{2}(n(\mu,\sigma) + \text{median})}\right| + \left|\frac{p_{90/10}(\mu,\sigma) - \text{P90/P10}}{\frac{1}{2}(p_{90/10}(\mu,\sigma) + \text{P90/P10})}\right| \to \min$$
(18)

with the constraint

$$m(\mu, \sigma) = \text{mean} \tag{19}$$

 $n(\mu, \sigma)$, $p_{90/10}(\mu, \sigma)$, and $m(\mu, \sigma)$ must be taken from Equation (31), Equation (33), and Equation (25), respectively. The values for median, P90/P10, and mean come from **Table 1**. Equation (18) and Equation (19) look quite simple. However, inserting the functions from Equation (31), Equation (33), and Equation (25), it is already much more complicated. Theoretically one might solve Equation (19) for μ or σ and insert it in Equation (18). Unfortunately, there is neither an analytic solution nor one with arbitrary numerical accuracy. But of ⁵The problem is not only the dimension. Even with dimensionless quantities one can be very big (e.g. around 100) and the other one very small (e.g. around 1). Though both deviations are equally important, an ordinary absolute value fit will stress the deviation of the big quantity much more.

course, Equation (19) defines a function $\mu(\text{mean}, \sigma)$ or $\sigma(\text{mean}, \mu)$, at least piecewise. Inserting one of these functions in Equation (18) leads to one variable in Equation (18). Instead of minimizing one might set Equation (18) to zero. However, unlike polynomials highly non-linear functions generally need not to have zeros (even complex ones). And indeed at least with the values from **Table** 1 there will be no zeros. The classical approach to minimize Equation (18) by differentiation with respect to μ and σ and setting to zero is excluded because of the absolute values. Furthermore, differentiating the functions $n(\mu, \sigma)$ and $p_{90/10}(\mu, \sigma)$ gives far from trivial results. In addition there is the constraint of Equation (19).

Nevertheless, Equation (18) and Equation (19) are a well-defined problem with a solution. The constraint of Equation (19) makes it a one-dimensional problem in two dimensions. In Figure 4 we have e.g. plotted

. 1

$$\left| \frac{n(\mu, \sigma) - \text{median}}{\frac{1}{2} (n(\mu, \sigma) + \text{median})} \right| + \left| \frac{p_{90/10} (\mu, \sigma) - \frac{P90}{P10}}{\frac{1}{2} \left(p_{90/10} (\mu, \sigma) + \frac{P90}{P10} \right)} \right|$$
(20)

with the constraint of Equation (19) and the values of Germany from **Table 1**. The minimum of the red curve in μ - σ -space is our desired minimum. Determining this minimum numerically is not very complicated, once the functions of Equation (18) and Equation (19) are programmed. Using a software like Mathematica can be very helpful here. The minimum in **Figure 4** is at $\mu \approx 2106.3705 \in$ and $\sigma \approx 992.90175 \in$. The expression in Equation (20) takes the value of about 0.119. Instead of using the median one could have used the Gini $g(\mu, \sigma)$ of Equation (36). Though this will consume more CPU time, the result is surprisingly similar⁶. The minimum in a correspondingly changed Figure 4 will be at $\mu \approx 2106.3698 \in$ and $\sigma \approx 992.90455 \in$. The expression in Equation (20) (with $g(\mu, \sigma)$ instead of $n(\mu, \sigma)$) takes the value of about 0.163.

The similarity makes two things likely. Firstly, our fit procedure is not just luck. Secondly, the income distribution is essentially Gaussian as long as global measures like median or Gini are concerned. It is also intriguing to compare the result of the fit procedure of subchapter 3.1⁷ with our results here. It will show how wrong the approach of subchapter 3.1 is. The first column of **Table 2** gives the "exact" data from OECD.Stat. The calculated data from the fit here (fit 3.2) shows a quite good match. Because of the constraint in Equation (19), the mean is of course identical. The calculated data based on the results of the previous chapter (Equation (17)) deviate dramatically in the Gini and P90/P10 ratio. Taking them at face value, Germany's income distribution is as unequal as in the USA.

The results for all four countries are summarized in **Table 3**. Please note that ⁶Mathematically this is explained easily. In both version $p_{9010}(\mu, \sigma)$ was exactly 3.7, which implies a fixed value for μ and σ because of Equation (19).

⁷We are aware that the data from **Figure 3** are from 2017 while data from 2016 are used here. Unfortunately, we have for no data from 2016 for the Graphics in **Figure 3**. We also don't have data from 2017 for the quantities of **Table 1**. But the difference seems to be unimportant.



Figure 4. Plot of Equation (18) with constraint from Equation (19) with values for Germany.

Table 2. Comparison between classical fit (3.1) and new procedure (3.2) for GER.

	OECD.Stat	fit 3.2	fit 3.1
mean	2148.83 €	2148.83 €	2147.55 €
median	1888.50 €	2127.46 €	1993.36 €
P90/P10	3.7	3.7000	8.1030
Gini	0.294	0.2496	0.3508

Table 3. Summary of the four countries considered here.

	μ	σ	Gini this fit	Gini OECD.Stat
GER	2106.37 €	992.902 €	0.250	0.294
USA	2631.14 €	1918.25 €	0.315	0.391
UK	1852.56 €	967.013 €	0.267	0.351
DK	2202.08 €	848.398 €	0.213	0.261

 μ and σ are just fit parameters. They do not have the meaning of mean and standard deviation like in the Gaussian distribution of Equation (4). Due to the constraint of Equation (19) we have in effect only one fit parameter. With one parameter only, the Gaussian model used in this subchapter describes reality nicely. This is especially surprising as an income distribution is not a result of one or a few constants as it is often the case in e.g. physics. Here many million people interact, and all have a free will. Obviously we have a quite just world. As most skills (especially IQ) are Gaussian distributed, the income is in accordance with it.

Having a closure look at **Table 3** one can see that the Gini from our fit is always smaller than the observed one. The income in our modelled world is more equally distributed than reality shows. As stated in the introduction already, with a pure Gaussian distribution very wealthy people cannot exist. We have something like a fat tail. To solve this puzzle is the main point of the following subchapter.

3.3. The Extended Model

In the previous subchapter we have shown how to fit an income distribution with one effective fit parameter. The results are quite fine. However, they deviate for the rich. With the numbers of **Table 1** and **Table 3** it is easy to calculate the number of households (#*HH*) with a net income of 10,000 \in or more per month by integrating q(E) in Equation (10) from 10,000 \in to infinity. One will get the following:

$$#HH_{GER} = 3.86 \times 10^{-8}, #HH_{USA} = 8437$$

$$#HH_{UK} = 5.14 \times 10^{-10}, #HH_{DK} = 5.23 \times 10^{-14}$$
 (21)

A net household income of $10,000 \in$ is for sure not common in neither of the four countries as the median is roughly four times lower in each of the four. However, it does exist⁸, and for sure it is possible without any income from wealth. In contrast to it, Equation (21) teaches that such households should not exist in three out of four countries considered. And even in the USA the number is incredibly tiny. It should be bigger by at least a factor of 10^3 .

Though the necessary skills to create income are fairly well Gaussian distributed, at least higher incomes are much more likely than any Gaussian distribution would predict. One thing is the income of a *leader*. This is a person who has subordinates. And part of the value created by these subordinates will contribute to the income of the leader. Generally this is just because only the leader has the skills enabling the subordinates to create so much value. It is an explanation why leaders may have hourly wages several times the wages of their subordinates. However, these leadership skills will also show a Gaussian distribution. It will not lead to the observed *fat tail* in income distributions from work. It is also impossible that the households with lower income are betrayed by their bosses. It would be possible in totalitarian states but for sure not in the four OECD countries considered here. Democracy and a working labor market will always lead to justice. If a boss pays too little, the most skilled workers will leave making the company less profitable. It is the same as with ordinary goods. The market determines the price.

That people will allow for a redistribution, be it by tax or even free giving is not impossible even in market economies. Using an extended Edgeworth cube [40] it can be shown that it does make sense to consider "social peace" as a good people are willing to trade. Though this is a reasonable and likely effect, it would lead to an even narrower distribution especially for lower incomes and something like a "slim" tail.

However, without a working labor market everything is possible. Considering median incomes it is hard to imagine that the labor market is not working in this area. This is probably even accepted by trade unions and the like. They demand minimum wages, child support, etc. just because the free market is creating a too

⁸Unfortunately the authors have no data on households with an income of over $10,000 \in$ per month. However, it should be reachable with two academics working full time in above average positions in all four countries.

broad spread of incomes. A labor market works if there are many similar positions and many potential people able to fill the position. For people having several times the median income, there are less and less potential positions. Even if they imagine that they are creating much more value for their bosses, it will be much more difficult to find an alternative. One chance will be a spinoff. But it is rarely realistic. Rightly, there is also no lobby for people having several times the median income. They may suffer from injustice but not from financial hardship.

Now we have shown why a labor market may partly not work. It is another question who takes why an advantage of it. The answer to this question is only at first glance obvious. Unlike e.g. chimpanzees, humans are not homines oeconomici. They are not altruist either. They go for more money in order to become richer than their peers.⁹ It is the case especially within rich people. This is neither new nor is it just a gut feeling. Recently historic data from the second half of the 19th century has been analyzed in detail [26]. Such historic data have the advantage of being "natural." They are not biased by modern social policies. The essential result from [26] is, that the rich are even willing to give to the poorer ones as long as the rich remain distinct in income.

Putting this into our Gaussian distribution of Equation (10) would mean a sigma growing with income. Though we do not say that this ansatz is not worth pursuing, it has two disadvantages. Firstly, it comes a little bit unmotivated. Secondly, it bears technical problems. If $\sigma = \sigma(E)$ in Equation (10), the leading power of $\sigma(E)$ must be less than linear. In other words, $\sigma(E) \propto E^a$ with a < 1 in order to make normalization possible. With such low powers the effect is pretty tiny (besides making the math complicated especially for 1/2 < a < 1). This technicality can be fixed by introducing an income cutoff. Having a maximum possible income in the world is even realistic. Our income distribution like in Equation (10) is in that sense unrealistic because it will give a (very small) probability that someone has ten times the world income. On the other hand, setting a cutoff value seems arbitrary. It looks like an unmotivated fit parameter. Therefore we did not pursue this ansatz.

Our model used goes back to the effect that richer people tend to be leaders getting their money from advising subordinates. Getting up the hierarchy the number of people will be less and less. This alone would lead to an exponential distribution like in Equation (3). To make the number of assumptions as small as possible, we say that everybody tries to get money through subordinates. However, the will to do it *and* the possibility is proportional to income. This leads to an E^2 term in front of the exponential distribution:

$$E^2 \cdot e^{-\frac{E}{\lambda}}$$

We have to add this modified exponential distribution to our Gaussian one of Equation (10). After normalization we will have

⁹Of course there are many more models. Though our purpose is to show the mathematics behind modeling, the example presented here supports this assumption.

$$q(E) = \frac{1}{2\sigma\sqrt{2\pi}\cdot\left(1 + \operatorname{erf}\left(\frac{\mu}{\sqrt{2\sigma}}\right)\right)} \cdot e^{-\frac{(E-\mu)^2}{2\sigma^2}} + \frac{E^2}{4\lambda^3} \cdot e^{-\frac{E}{\lambda}}$$
(22)

Formally we have now an identical optimization as given in Equation (18) and Equation (19). Instead of q(E) from Equation (10) we have to use Equation (22) now. Our problem is the following:

$$\left|\frac{n(\mu,\sigma,\lambda) - \text{median}}{\frac{1}{2}(n(\mu,\sigma,\lambda) + \text{median})}\right| + \left|\frac{p_{90/10}(\mu,\sigma,\lambda) - P90/P10}{\frac{1}{2}(p_{90/10}(\mu,\sigma,\lambda) + P90/P10)}\right| \to \min$$
(23)

with the constraint

$$\lambda = \frac{2}{3}\operatorname{mean} + \frac{\mu}{3} \cdot \frac{\Gamma_r \left(-\frac{1}{2}, \frac{\mu^2}{2\sigma^2}\right) - 2}{1 + \operatorname{erf}\left(\frac{\mu}{\sqrt{2\sigma}}\right)}$$
(24)

ī.

 $n(\mu,\sigma,\lambda)$ and $p_{90/10}(\mu,\sigma,\lambda)$ are given in Equation (39) and Equation (40), respectively. This minimization problem is well defined. The constraint is even simpler as in Equation (19). However, the highly non-linear functions must be determined numerically which consumes quite some CPU-time and RAM. The reason behind it is stated in the next chapter. There we also explain why it is virtually impossible to use the Gini $g(\mu,\sigma,\lambda)$ in Equation (23).

Making a 3D-Plot of Equation (23) (with λ substituted via Equation (24)) shows the areas of local minima. One should not just calculate points and connect them. Gradients should be considered too. This will make sure that there is really and minimum. It will increase the number of points to be calculated by perhaps a factor of ten times ten. But it is necessary because the minimum will be typically at a non-analytic point due to the absolute values in Equation (23). A software like Mathematica is very helpful here. As the problem lies in the inverse functions, Mathematica analyses the original functions and tries to get at least piece-wise analytic inverse functions. This is of course not always possible. So one has to choose by hand which interval should be considered. Even this way it costs quite some CPU-time. And it is neither straight forward nor can it be automated. Having identified the area with the smallest local minimum it has proven practical to find its value iteratively by guessing the value for σ and then making a one-dimensional plot over μ which will yield a minimum at a certain value of μ . With this value of μ one can plot over σ , and so forth until sufficient accuracy has been reached.

In **Table 4** we summarized the data for our four countries. We are confident the digits shown there are accurate. But we are far from being able to produce results with arbitrary accuracy. Even this accuracy is pretty tedious and at least the authors see no way to automize it. The most interesting column in **Table 4** is the last one. Now the number of households with a net monthly income of 10,000 \in or more is at least about correct though most likely still understated. As indicated

	μ	σ	λ	# of households
GER	1719€	539€	859€	$1.44 \cdot 10^4$
USA	1772 €	928.2 €	1361€	$1.43 \cdot 10^{6}$
UK	1341€	416€	830€	6.96·10 ³
DK	1955€	340 €	824 €	$6.27 \cdot 10^2$

Table 4. Data for the coefficients and the number of housholds with $\ge 10^4 \notin$ per month.



Figure 5. Gaussian and exponential part of the income distribution for Germany.

in footnote 7, we do not know the exact number of these households. Else it would be smart to use it as a quantity to be fitted directly.

In **Figure 5** we have plotted the Gaussian (blue, narrow) and exponential (beige, wide) of the income distribution for Germany separately. Both parts peak at around the median income¹⁰. The total distribution is the sum of both curves. The exponential part clearly enlarges the peak at the median income. The number of households getting about the median income also increases. In addition one gets a "fat tail."

4. Derivation of Some Equations

In subchapter 3.1 we defined in Equation (10) a "Gaussian" distribution q which runs from zero to infinity. Of course it is straight forward to calculate the mean $m(\mu, \sigma)$ and variance (=square of standard deviation) $s(\mu, \sigma)^2$:

$$m(\mu, \sigma) = \int_{0}^{\infty} dE E \cdot q(E)$$
$$(\mu, \sigma)^{2} = \int_{0}^{\infty} dE (E - m(\mu, \sigma))^{2} \cdot q(E)$$

These integrals are tedious but straight forward to solve. A lengthy calculation yields

S

¹⁰Formally the Gaussian part (blue curve) peaks at μ and the exponential part (beige curve) at 2 λ . This makes the USA in **Table 4** a particular case.

$$m(\mu,\sigma) = \frac{\mu\left(2-\Gamma_r\left(-\frac{1}{2},\frac{\mu^2}{2\sigma^2}\right)\right)}{1+\operatorname{erf}\left(\frac{\mu}{\sqrt{2}\sigma}\right)}$$
(25)
$$s(\mu,\sigma)^2 = \frac{1}{\left(1+\operatorname{erf}\left(\frac{\mu}{\sqrt{2}\sigma}\right)\right)^3} \left(e^{-\frac{\mu^2}{2\sigma^2}}\sqrt{\frac{2}{\pi}}\mu\sigma\left(1+\operatorname{erf}\left(\frac{\mu}{\sqrt{2}\sigma}\right)\right)^2 + 2\left(\mu^2+\sigma^2+\operatorname{erf}\left(\frac{\mu}{\sqrt{2}\sigma}\right)\left(2(\sigma-\mu)(\mu+\sigma)+(\mu^2+\sigma^2)\operatorname{erf}\left(\frac{\mu}{\sqrt{2}\sigma}\right)\right) + \mu^2\left(4-\Gamma_r\left(-\frac{1}{2},\frac{\mu^2}{2\sigma^2}\right)\right)\Gamma_r\left(-\frac{1}{2},\frac{\mu^2}{2\sigma^2}\right)\right)\right)$$
(26)
$$-\left(5\mu^2+\sigma^2+(\mu^2+\sigma^2)\operatorname{erf}\left(\frac{\mu}{\sqrt{2}\sigma}\right)\left(2+\operatorname{erf}\left(\frac{\mu}{\sqrt{2}\sigma}\right)\right) - 4\mu^2\Gamma_r\left(-\frac{1}{2},\frac{\mu^2}{2\sigma^2}\right) + \mu^2\Gamma_r\left(-\frac{1}{2},\frac{\mu^2}{2\sigma^2}\right)^2\right)\Gamma_r\left(-\frac{1}{2},\frac{\mu^2}{2\sigma^2}\right)\right)$$

The error function erf has been defined in Equation (11) already. Γ_r is the regularized incomplete gamma function with

$$\Gamma_r(a,x) = \frac{\Gamma(a,x)}{\Gamma(a)}$$
(27)

where $\Gamma(a, x)$ and $\Gamma(a)$ are incomplete and "normal" gamma function, respectively with

$$\Gamma(a,x) = \int_{x}^{\infty} dt \, t^{a-1} e^{-t} \text{ and } \Gamma(a) = \int_{1}^{\infty} dt \, t^{a-1} e^{-t} + \sum_{k=0}^{\infty} \frac{(-1)^{k}}{k!(k+a)}$$
(28)

Please note the sum in Equation (28). Normally, the gamma function is displayed by an integral only. But this only works for positive arguments. In the entire paper the first argument of Γ_r is -1/2. So we have

$$\Gamma_r\left(-\frac{1}{2},x\right) = -\frac{1}{2\sqrt{\pi}}\Gamma\left(-\frac{1}{2},x\right) = -\frac{1}{2\sqrt{\pi}}\int_x^{\infty} dt \, t^{-3/2} e^{-t}$$
(29)

Equation (29) does not lead to much simplification in the numerical calculations.

To be consistent with an absolute value fit one should consequently write

$$s(\mu,\sigma) = \int_{0}^{\infty} dE \left| E - m(\mu,\sigma) \right| \cdot q(E)$$

This integral is easily solved by splitting it in one running from 0 to $m(\mu, \sigma)$ and one running from $m(\mu, \sigma)$ to ∞ . Though the solution is straight forward, it is much more complicated than Equation (26) and has about double its length. In subchapter 3.2 we used the median, which we will denote *n* here (because *m* is already used for the mean). For any distribution q(E) normalized to one the median is

$$n = Q^{-1}\left(\frac{1}{2}\right) \text{ with } \frac{\partial Q(x)}{\partial x} = q(x)$$
(30)

As always, the exponent -1 denotes the reverse function with $f^{-1}(f(x)) = x$. Applying this to the q(E) of Equation (10) (Gaussian from zero to infinity) leads to

$$n(\mu,\sigma) = \mu + \sigma\sqrt{2} \cdot \operatorname{erfc}^{-1} \left(1 - \frac{1}{2} \operatorname{erfc} \left(\frac{\mu}{\sigma\sqrt{2}} \right) \right)$$
(31)

where erfc is the complementary error function with $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$.

The P90/P10 (abbreviated as $p_{90/10}$) ratio one gets for a general (normalized) distribution q(E)

$$\int_{-\infty}^{\mathbb{P}_{10}} \mathrm{d}E \ p\left(E\right) = \frac{1}{10} \implies \mathbb{P}_{10} = Q^{-1}\left(\frac{1}{10}\right)$$

and

$$\int_{P90}^{\infty} dE \ p(E) = \frac{1}{10} \implies P90 = Q^{-1} \left(-\frac{1}{10} \right)$$

to

$$p_{90/10} = \frac{Q^{-1} \left(-\frac{1}{10} \right)}{Q^{-1} \left(\frac{1}{10} \right)}$$
(32)

where Q is defined as in Equation (30). Applied to the q(E) of Equation (10) (Gaussian from zero to infinity) leads to

$$p_{90/10}(\mu,\sigma) = \frac{\mu + \sigma\sqrt{2} \cdot \operatorname{erfc}^{-1}\left(\frac{1}{10} + \frac{1}{10}\operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right)}{\mu + \sigma\sqrt{2} \cdot \operatorname{erfc}^{-1}\left(\frac{9}{10} + \frac{9}{10}\operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right)}$$
(33)

Here erf and erfc are defined as in Equation (11) and Equation (31), respectively.

Though not used here, a few words about the Gini coefficient g. It is defined for distributions q(E) running from zero to infinity only.

$$g = 1 - 2\int_{0}^{1} \mathrm{d}f \, \frac{\int_{0}^{E(f)} \mathrm{d}\eta \, \eta \cdot q(\eta)}{\int_{0}^{\infty} \mathrm{d}\eta \, \eta \cdot q(\eta)} \tag{34}$$

with

$$E = E(f) \text{ as inverse function of } f(E) = \int_{0}^{E} d\eta q(\eta)$$
(35)

Taking the q(E) of Equation (10) leads to after a tedious but straight forward calculation

$$g(\mu,\sigma) = 1 - \frac{\left(\mu + 2\sigma\sqrt{\frac{2}{\pi}}e^{-\frac{\mu^{2}}{2\sigma^{2}}} + \mu \cdot \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right) + \frac{2\sigma \cdot \operatorname{erf}\left(\sqrt{2}\operatorname{erfc}^{-1}\left(1 + \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right)\right) - 2\sigma}{\sqrt{\pi}\left(1 + \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right)}\right) \left(\operatorname{erfc}\left(\frac{\mu}{\sigma\sqrt{2}}\right) - 2\right)}{\sqrt{\pi}\left(1 + \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right)} \left(1 + \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right)\right) \left(1 + \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right)\right) \left(1 + \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right) \left(1 + \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right)\right) \left(1 + \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right) \left(1 + \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right) \left(1 + \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right)\right) \left(1 + \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right)\right) \left(1 + \operatorname{$$

 Γ_r is defined as in Equation (27) and Equation (28). Though this expression looks pretty clumsy, it consists of functions which can be evaluated with arbitrary accuracy. This is in contrast to the Gini coefficient we would need in sub-chapter 3.3.

Calculating the mean $m(\mu, \sigma, \lambda)$ for the distribution of Equation (22) is simple because an integral over a sum is the sum of the integrals.

$$m(\mu,\sigma,\lambda) = \frac{\mu \left(1 - \frac{1}{2}\Gamma_r \left(-\frac{1}{2}, \frac{\mu^2}{2\sigma^2}\right)\right)}{1 + \operatorname{erf}\left(\frac{\mu}{\sqrt{2\sigma}}\right)} + \frac{3}{2}\lambda$$
(37)

The constraint from Equation (19) $m(\mu, \sigma, \lambda) = \text{mean}$ can be solved for λ .

$$\lambda = \frac{2}{3}\operatorname{mean} + \frac{\mu}{3} \cdot \frac{\Gamma_r \left(-\frac{1}{2}, \frac{\mu^2}{2\sigma^2}\right) - 2}{1 + \operatorname{erf}\left(\frac{\mu}{\sqrt{2}\sigma}\right)}$$
(38)

with Equation (38) the additional parameter λ can be eliminated.

To calculate the median $n(\mu, \sigma, \lambda)$ for the distribution of Equation (22) is formally like in Equation (30).

$$n = Q^{-1}\left(\frac{1}{2}\right)$$

with $Q(x) = 1 - \frac{e^{-\frac{x}{\lambda}}}{4\lambda^2} \left(x^2 + 2x\lambda + 2\lambda^2\right) - \frac{\operatorname{erfc}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)}{4 - 2\operatorname{erfc}\left(\frac{\mu}{\sigma\sqrt{2}}\right)}$ (39)

Unfortunately there is no closed inverse function of Q(x). Building an inverse function is numerically simple. But the amount of data is very big within our problem. Of course, we can insert λ from Equation (38) into Q(x) of Equation (39). This leaves us with two parameters μ and σ which are supposed to be determined eventually. In the sense of a Mont Carlo simulation one may assume 10^3 different values for each parameter. So we have 10^6 different functions Q(x). In order to build an inverse function, we may have to assume 10^3 different values for x. It leaves us with 10^9 values which must be calculated. As each calculation contains integrals, this will consume quite some computing power.

Building the P90/P10 ratio for the distribution from Equation (22) causes the same problem. Formally $p_{90/10}(\mu, \sigma, \lambda)$ is given like in Equation (32)

$$p_{90/10}(\mu,\sigma,\lambda) = \frac{Q^{-1}\left(-\frac{1}{10}\right)}{Q^{-1}\left(\frac{1}{10}\right)}$$
with $Q(x) = 1 - \frac{e^{-\frac{x}{\lambda}}}{4\lambda^2} \left(x^2 + 2x\lambda + 2\lambda^2\right) - \frac{\operatorname{erfc}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)}{4 - 2\operatorname{erfc}\left(\frac{\mu}{\sigma\sqrt{2}}\right)}$
(40)

The parameter λ can be eliminated with Equation (38). Again, we will need the inverse function of Q(x) given in Equation (40). In order to find it we also have to calculate 10⁹ data points. (However, it is mostly an identical calculation)

Just for completeness we will also show how to calculate the Gini coefficient for the distribution given in Equation (22). The Gini $g(\mu, \sigma, \lambda)$ is formally given in Equation (34). The necessary function f(E) in Equation (35) is easily determined to be

$$f(E) = 1 - \frac{e^{-\frac{E}{\lambda}}}{4\lambda^2} \left(E^2 + 2E\lambda + 2\lambda^2 \right) - \frac{\operatorname{erfc}\left(\frac{E-\mu}{\sigma\sqrt{2}}\right)}{4 - 2\operatorname{erfc}\left(\frac{\mu}{\sigma\sqrt{2}}\right)}$$
(41)

For obvious reasons f and Q are identical. Putting it together we have

$$g(\mu,\sigma,\lambda) = 1 - \frac{2}{m(\mu,\sigma,\lambda)} \int_{0}^{1} \mathrm{d}f \int_{0}^{E(f)} \mathrm{d}\eta \,\eta \cdot q(\eta)$$
(42)

 $m(\mu,\sigma,\lambda)$ is given in Equation (37). λ can be eliminated with Equation (38). E(f) is the reverse function from f(E) given in Equation (41), and $q(\eta)$ is defined in Equation (22). Inserting all this will make Equation (42) look much more complicated. But this is not the real problem. Getting the function E(f) one needs to make 10° calculations as stated above. Furthermore we have to take two integrals in Equation (42). Even going for a not too high accuracy, we may divide each integration interval into 100 pieces. This leads us with 10^{9+2+2} datasets. Even storing these 10^{13} data is critical. Therefore, a fit with the Gini instead or in addition to the median is impossible. Making some simplifications was not possible; at least the authors did not find any way. Finding the inverse function E(f) of f(E) may involve far less than the mentioned 10^9 data. A smart software like Mathematica is able to find gradients in f(E) and also proves continuity. With it, it is (mostly) able to construct an E(f) is useless for the two (numerical) integrations in Equation (42).

5. Conclusions and Further Research

We have shown how to use distributions, and what conclusions can be drawn. We have taken two examples and used mathematics which is well-known for over 100 years. This alone would disqualify our work as a journal publication. But we have chosen two particular examples: IQ distribution and income distribution. These examples belong to psychology and neighboring fields, and economics and finance and the like. These disciplines have in common that they use distributions frequently though they are not too close to mathematics. Over simplified two statements are prevalent there:

- Every distribution is a Gaussian distribution.
- The mean and the standard deviation determine μ and σ , respectively.

At most a X^2 test is applied in order to prove Gaussian behavior. Applying the central limit theorem wrongly sometimes produces Gaussian distributions which are by no means justified. The statements make life easy, but they are wrong and may lead to false conclusions. To show examples for it was the main motivation to write this paper.

Our first example attacks the first statement. IQ is distributed almost perfectly Gaussian. At first glance that seems to confirm the statement. However, if distributions are always Gaussian, a Gaussian IQ distribution is a tautology. A Gaussian distribution appears only if something happens by chance. It is very difficult to produce such distribution otherwise, as everybody might know once trying to fake lab data with the tools of the early 1980 ties. The accidently mixing of genes generates a certain IQ. In chapter 2 we concluded from this that IQ must be inherited or at least not being created by conscious actions. Though this is assumed by the vast majority of academic psychologist, we have presented mathematical proof for it.

As our proof is clear-cut, it is hard to imagine any further research. However, two things may be worth scrutinizing. One is the width of the IQ distribution. It differs for men and women. But even the total IQ distribution does most likely not have the same width in every country. The big problem is finding sources of data for it. Even for the average IQ in different countries there are no complete reliable data. In developing countries, it is particularly dim. Data for the width are not available, at least not for the authors. However, many factors might contribute to IQ and especially its distribution. Suspects are the frequency of marrying cousins, fidelity, religion, and many more.

In this paper we came also across cultural inheritance [37]. As at least some traits are inherited genetically, some are inherited culturally. It is much less scrutinized. The distribution of such traits might shed some light on the mechanism.

Our second examples are income distributions. Income is based on skills such as IQ. Many of them show a Gaussian distribution. Therefore it is a reasonable assumption that income shows a Gaussian distribution. Subchapters 3.1 and especially 3.2 confirm it in most points. Puzzling is a "fat tail". It means that there are far too many rich than any Gaussian distribution can predict. Before solving this problem, we showed in subchapter 3.1 that even if assuming a Gaussian distribution from minus to plus infinity, fitting μ by the mean and especially σ by the standard deviation can be tremendously wrong. In subchapter 3.2 we assumed a Gaussian distribution with positives incomes only. Such "half" Gaussian is a much more realistic assumption in many cases ranging from infection rates to certain measurement errors. There the standard deviation has nothing to do with σ . In Equation (26) we have displayed a lengthy formula for the standard deviation in such distribution. It is a (complicated) function of μ and σ . Though we got in subchapter 3.2 a decent fit with our half Gaussian, it is still far from possible to explain the fat tail.

In subchapter 3.3 we introduced in Equation (22) a distribution consisting of a Gaussian part and a (modified) exponential distribution. We were motivated to it by the evaluation of historical data by Chapman [26]. It led to a tremendously better description of the fat tails. The predicted number of households with a monthly income of €10,000 or above increased by a factor ranging from 10^{16} (DK) to 10^3 (USA) to much more realistic values. With monthly net incomes of €10,000 and especially above, the income from (inherited) wealth will be more and more important. Therefore it does not make sense to scrutinize very high incomes in detail within our model.

As a further research within income distributions one could extend the procedure of subchapter 3.3 to other countries and by using other measures to fit. As the (numerical) mathematics is very complicated, the authors will not presume actions in this direction. We also see no way to simplify or automate the numeric calculations, though it would be very welcome. Maybe tools of *big data* can help.

It would be worthwhile investigating in other areas where the results of [26] play a major (quantitative) role. Sloppily paraphrasing the results of [26] reads: "You are rich if you have more money than your neighbor". That envy influences behavior is well-known. We would like to look for quantitative effects other than income distributions. We give a warning to use envy as a variable which can be done e.g. in game theory. Assuming a value such as 3.7 for envy is possible but these envy values are not elements of a field (in a mathematical sense). Though calculations such as addition or multiplications are technically possible, the results are ludicrous [10].

A quite obvious extension of our results is fat tails in finance. The general mechanism must be the same. The income from stocks is nothing more than the sum of the values created by workers. Furthermore, even companies are neither homines oeconomici nor "machinae oeconomici". Companies are always led by humans. Therefore it comes as no surprise that many bosses want to make their company more profitable than a rival company, even if the total profit of both companies reduces this way. Quite a few (proud) stock holders will accept it.

The reason why we have not taken the fat tail from finance as an example has many sources. There are fundamental errors in the work of Fama [3]. To understand them one has to understand quite many newer publications such as [9] to [11] and [27]. This would have led away from the point we want to make. Commenting briefly, the following can be said. Fama assumes that there is a real fair value for e.g. stocks and the market prices fluctuates around it due to imperfect information. This would lead to a perfect Gaussian distribution and a fat tail would be a *big* surprise. But even the fellow Nobel laurate of 2013 (Robert Shiller)

disagrees with it. (Giving a shared Nobel prize to two contradicting theories in one year remains a conundrum to the authors) As explicitly shown in [9], the price of a stock does not fluctuate around an underlying value of the company considered. On average it overstates the company value several times. Up to now we are speaking of the fluctuation in time of a single stock. Assuming ergodicity the fluctuation should be the same over a portfolio of stocks. But assuming ergodicity is not justified even if Fama's approach was correct. The underlying true price is not something like an energy minimum. It changes in time differently for each stock. Furthermore the fluctuations in price are chaotically rather than by chance as has been shown in [11] and [27]. Chaotic fluctuations look like chance, but they are distinct [16] though this may or may not play a major role here. Chaotic fluctuations are deterministic instead random though they look random. Therefore ergodicity does not need to hold. The origin of these fluctuations is not an adjustment to true market prices but speculation. Though with some flaws the effect has been analyzed quantitatively in [41] by using Fourier analysis.

Acknowledgements

This publication is dedicated to the late Thomas Dierks. A distinguished scientist, and high school classmate and comrade of M.G.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Bronshtein, I.N., Semenddyayev, K.A., Musiol, G. and Muehlig, H. (2007) Handbook of Mathematics. 5th English Edition, Springer, Berlin Heidelberg.
- [2] Whittaker, G. and Robinson, E. (1924) The Calculus of Observations. Old Bailey, London.
- Fama, E.F. (1965) The Behavior of Stock-Market Prices. *Journal of Business*, 38, 34-105. <u>https://doi.org/10.1086/294743</u>
- [4] Boothe, P. and Glassman, D. (1987) The Statistical Distribution of Exchange Rates: Empirical Evidence and Economic Implications. *Journal of International Economics*, 22, 297-319. <u>https://doi.org/10.1016/S0022-1996(87)80025-9</u>
- Hsieh, D.A. (1988) The Statistical Properties of Daily Foreign Exchange Rates: 1974-1983. *Journal of International Economics*, 24, 129-145. https://doi.org/10.1016/0022-1996(88)90025-6
- Kaplan, P.D. (2012) Frontiers of Modern Asset Allocation. John Wiley and Sons, Hoboken. <u>https://doi.org/10.1002/9781119205401</u>
- [7] Schuster, H.G. (1984) Deterministic Chaos. PhysikVerlag, Weinheim.
- [8] Grabinski, M. (2004) Is There Chaos in Management or Just Chaotic Management? Complex Systems, Intelligence and Modern Technology Applications, Paris. <u>http://www.h-n-u.de/Veroeffentlichungen/CSIMTA%202004.pdf</u>

- [9] Appel, D. and Grabinski, M. (2011) The Origin of Financial Crisis: A Wrong Definition of Value. *PJQM*, 2, 33-51.
 <u>http://www.h-n-u.de/Veroeffentlichungen/PJQM_2_I_2011.pdf</u>
- [10] Klinkova, G. and Grabinski, M. (2017) Conservation Laws Derived from Systemic Approach and Symmetry. *International Journal of Latest Trends in Finance and Economic Sciences*, 7, 1307-1312. <u>http://h-n-u.de/systemic.pdf</u>
- [11] Klinkova, G. (2018) The Effects of Chaos on Business Operations. PhD Thesis, Neu-Ulm. <u>http://uni-neu-ulm.de/Klinkova_PhD</u>
- [12] Filipe, J.A., Ferreira, M.A.M., Coelho, M. and Pedro, M.I. (2010) Chaos, Anti-Chaos and Resources: Dealing with Complexity. *Aplimat-Journal of Applied Mathematics*, 3, 83-90.
- [13] Ferreira, M.A.M., Filipe, J.A., Coelho, M. and Pedro, M.I. (2010) Fishing Policies and the Contribution of Chaos Theory for Fisheries Management. *Journal of Economics and Engineering*, 2, 36-43.
- [14] Filipe, J.A. and Ferreira, M.A.M. (2013) Social and Political Events and Chaos Theory—The "Drop of Honey Effect." *Emerging Issues in the Natural and Applied Sciences*, 3, 126-137.
- [15] Grabinski, M. and Klinkova, G. (2020) Why Individual Behavior Is Key to the Spread of Viruses Such as Covid-19. *Theoretical Economics Letters*, **10**, 299-304. <u>https://doi.org/10.4236/tel.2020.102020</u>
- [16] Grabinski, M. and Klinkova, G. (2019) Wrong Use of Averages Implies Wrong Results from Many Heuristic Models. *Applied Mathematics*, **10**, 605-618. <u>https://doi.org/10.4236/am.2019.107043</u>
- Tallakstad, K.T., Toussaint, R., Santucci, S. and Måløy, K.J. (2013) Non-Gaussian Nature of Fracture and the Survival of Fat-Tail Exponents. *Physical Review Letters*, 110, Article ID: 145501. <u>https://doi.org/10.1103/PhysRevLett.110.145501</u>
- [18] Ferreira, M.A.M., Filipe, J.A. and Coelho, M. (2014) Performance and Differential Costs Analysis in a Two Echelons Repairs System with Infinite Servers Queues. *Advanced Materials Research*, **1036**, 1043-1048. <u>https://doi.org/10.4028/www.scientific.net/AMR.1036.1043</u>
- [19] Ferreira, M.A.M. and Filipe, J.A. (2015) Infinite Servers Queue Systems Busy Period—A Practical Case on Logistics Problems Solving. *Applied Mathematical Sciences*, 9, 1221-1228. <u>https://doi.org/10.12988/ams.2015.410808</u>
- [20] Ferreira, M.A.M. and Filipe, J.A. (2015) Occupation Study of a Surgery Service through a Queus Model. *Applied Mathematical Sciences*, 9, 4791-4798. <u>https://doi.org/10.12988/ams.2015.55390</u>
- [21] Filipe, J.A. and Ferreira, M.A.M. (2019) Solving Logistics' Problems through an Infinite Servers Queue Systems Approach. *International Journal of Business and Systems Research*, **13**, 494-507. <u>https://doi.org/10.1504/IJBSR.2019.102528</u>
- [22] Filipe, J.A. and Ferreira, M.A.M. (2017) In the Search for the Infinite Servers Queue with Poisson Arrivals Busy Period Distribution Exponential Behaviour. *International Journal of Business and Systems Research*, **11**, 453-467. https://doi.org/10.1504/IIBSR.2017.10007303
- [23] Grabinski, M. and Liu, M. (1990) A-B Interface of Superfluid ³He: Effect of Phase Coherence. *Physical Review Letters*, **65**, 2666. <u>https://doi.org/10.1103/PhysRevLett.65.2666</u>
- [24] Eguílus, V.M. and Zimmermann, M.G. (2000) Transmission of Information and Herd Behavior: An Application to Financial Markets. *Physical Review Letters*, 85,

5659-5662. https://doi.org/10.1103/PhysRevLett.85.5659

- [25] Piketty, T. (2014) Capital in the Twenty-First Century. Translated by Arthur Goldhammer from Le Capital au XXI siècle, Éditions du Seuil (2013). <u>https://doi.org/10.4159/9780674369542</u>
- [26] Chapman, J. (2019) Inequality and Poor Law Policy in Late-Victorian England. Annual Meeting of the Economic History Society, Belfast, 7 April 2019. https://eh.net/eha/wp-content/uploads/2018/06/Chapman.pdf
- [27] Klinkova, G. and Grabinski, M. (2017) Due to Instability Gambling Is the Best Model for Most Financial Products. *Archives of Business Research*, 5, 255-261. <u>https://doi.org/10.14738/abr.53.3029</u>
- [28] Fergusson, M. (2019) The Curse of Genius. *The Economist*, 1843, April 29th 2019 <u>https://www.economist.com/1843/2019/04/29/the-curse-of-genius</u>
- [29] Sternberg, R.J. (2008) Increasing Fluid Intelligence Is Possible after All. PNAS, 105, 6791-679. <u>https://doi.org/10.1073/pnas.0803396105</u>
- [30] Ben-Shahar, Y. (2017) Epigenetic Switch Turns on Genetic Behavioral Variations. PNAS, 114, 12365-12367. <u>https://doi.org/10.1073/pnas.1717376114</u>
- [31] Henderson, B.D. (1972) Perspectives on Experience. The Boston Consulting Group, New York.
- [32] Roscher, J. (2008) Bewertung von Flexibilitätsstrategien für die Endmontage in der Automobilindustrie. PhD Thesis, Stuttgart, 90f. https://elib.uni-stuttgart.de/bitstream/11682/4140/1/Diss_Roscher.pdf
- [33] Johnson, N., et al. (2011) Pattern in Escalations in Insurgent and Terrorist Activity. Science, 333, 81. <u>https://doi.org/10.1126/science.1205068</u>
- [34] Rudnick, J. and Gaspari, G. (2010) Elements of the Random Walk. Cambridge University Press, Cambridge.
- [35] Grabinski, M. (2007) Management Methods and Tools. Gabler, Wiesbaden, 204ff.
- [36] Klinkova, G. und Grabinski, M. (2012) Learning Curves with Two Frequencies for Analyzing All Kinds of Operations. Yasar University Publication, Bornova. <u>http://www.h-n-u.de/Veroeffentlichungen/learning.pdf</u>
- [37] Howard, J.A. and Gibson, M.A. (2017) Frequency-Dependent Female Genital Cutting Behavior Confers Evolutionary Fitness Benefits. *Nature Ecology & Evolution*, 1, 49. <u>https://doi.org/10.1038/s41559-016-0049</u>
- [38] Piketty, T. and Saez, E. (2003) Income Inequality in the United States, 1913-1998. *The Quaterly Journal of Economics*, 118, 1-39. https://doi.org/10.1162/00335530360535135
- [39] N.N. (2019) Measuring the 1 %. *The Economist*, Nov. 28th 2019. https://www.economist.com/briefing/2019/11/28/economists-are-rethinking-the-nu mbers-on-inequality
- [40] Kunze, O. and Schlatterer, F. (2018) The Edgeworth Cube: An Economic Model for Social Peace. *International Journal of Applied Behavioral Economics*, 7, 30-46. <u>https://doi.org/10.4018/IJABE.2018040103</u>
- [41] Schädler, T. and Steurer, E. (2019) Portfolio Selection Based on a Volatility Measure Adjusted for Irrationality. *Archive of Business Research*, 7, 278-283 <u>https://journals.scholarpublishing.org/index.php/ABR/article/view/6732/4245</u>