Annals of Operations Research manuscript No.
(will be inserted by the editor)

From Mean and Median Income to the Most Adequate Way of Taking Inequality Into Account

Vladik Kreinovich · Hung T. Nguyen · Songsak Sriboonchitta · Rujira Ouncharoen

Received: date / Accepted: date

Abstract How can we compare the incomes of two different countries or regions? At first glance, it is sufficient to compare the mean incomes, but this is known to be not a very adequate comparison: according to this criterion, a very poor country with a few super-rich people may appear to be in good economic shape. The importance of taking income inequality into account has been emphasized by Angus Deaton, the winner of the 2015 Nobel Prize in Economics. A more adequate description of economy is the *median* income. However, the median is also not always fully adequate: e.g., raising the income of very poor people clearly improves the overall economy but does not change the median. In this paper, we use known techniques from group decision making – namely, Nash's bargaining solution – to come up with the most adequate measure of "average" income: geometric mean. On several examples, we illustrate how this measure works.

We acknowledge the partial support of the Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Thailand. This work was also supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721.

V. Kreinovich
Department of Computer Science, University of Texas at El Paso
500 W. University, El Paso, TX 79968, USA
Tel.: +1-915-747-6951
Fax: +1-915-747-5030
E-mail: vladik@utep.edu

H. T. Nguyen
Department of Mathematical Sciences, New Mexico State University
Las Cruces, NM 88003, USA and
Faculty of Economics, Chiang Mai University
Chiang Mai, Thailand

S. Sriboonchitta
Faculty of Economics, Chiang Mai University
Chiang Mai, Thailand

R. Ouncharoen
Department of Mathematics, Chiang Mai University
Chiang Mai, Thailand

Keywords Income inequality · Nash's bargaining solution · Mean income · Median income

1 Mean income, median income, what next?

Mean income and its limitations. At first glance, if we want to compare the economies of two countries or two regions, all we need to do is divide, for each country, the total income by the number of people and compare the resulting values of mean income. If the mean income in country A is larger than the mean income in country B, this means that the economy of country A is in better shape than the economy of country B.

In many cases, this conclusion is indeed justified, but not always. The fact that the mean has limitations can be illustrated by a known joke: "What happens when Bill Gates walks into a bar? On average, everyone becomes a millionaire." This is a joke, but this joke reflects a serious problem: if a billionaire moves into a small and very poor country, the mean income in this country would increase but the country would remain very poor, contrary to the increase in the mean.

In other words, when comparing different economies, we need to take into account not only the total income, but also the degree of inequality in income distribution. The importance of taking income inequality into account has been emphasized by Angus Deaton, the winner of the 2015 Nobel Prize in Economics; see, e.g., [5].

Comment. In technical terms, we would like the proper measure of "average" income to not change much if we add of an *outlier* like Bill Gates. In statistics, the corresponding property of statistical estimates is known as *robustness*; see, e.g., [10, 31]. In these terms, the main problem of the mean is that it is not robust.

Medium income: a more adequate measure. To avoid the above problem, economists proposed several alternatives to the mean income. The most widely used alternative is the *median* income, i.e., the income level for which the income of exactly half of the population is above this level – and the income of the remaining half is below this level. For example, this is how the Organization for Economic Cooperation and Development (OECD) compares economies of different countries: by listing both their mean incomes and their median incomes; see, e.g., [25].

Median resolves some of the problems related to mean: for example, when Bill Gates walks into a bar, the mean income of people in the bar changes drastically, but the median does not change much.

Comment. The main problem with the mean, as we have mentioned, is that the mean is not robust. From this viewpoint, median – a known robust alternatives to the mean [10, 31] – seems a reasonable replacement of the mean.

Limitations of the median and remaining practical problem. While the median seems to be a more adequate measure of "average" income than the mean, it is not a perfect measure. For example, if the incomes of all the people in the poorer half increase – but do not exceed the previous median – the median remains the same. This is not a very adequate measure for governments that try to lift people out of poverty. Similarly, if the income of the poorer half drastically decreases, we should expect the adequate measure of "average" income to decrease – but the median remains unchanged.

Comment. After we reformulated the problem with mean in terms of robustness, a reader may be under the impression that robustness is all we seek. Alas, the above limitation shows that the problem of finding an appropriate measure of "average" income goes beyond robustness; namely:

– the main problem of mean is that it is *not robust* – it changes too much when we would like to change it a little bit;
– however, while the median *is* robust, it has another problem – it is "too robust": it changes too little (actually, not at all) when we would like it to change.

This example shows that we cannot solve our problem by simply reducing it to a known statistical problem of designing robust estimates, we do not need to solve the original problem of estimating the "average" income.

How this practical problem is resolved now. At present, economists propose different heuristic measures of "average" income which are supposedly more adequate than mean and median. There is no absolutely convincing arguments in favor of this or that measure; as a result, researchers use emotional and ideological arguments; see, e.g., [26].

What we do in this paper. In this paper, we show that under some reasonable conditions, it is possible to find the most adequate way how to take inequality into account when gauging the "average" income.

Comment. This paper is an extended version of the conference paper [15].

2 Analysis of the problem and the resulting measure

The problem of gauging "average" income can be viewed as a particular case of a problem of group decision making. For the problem of gauging "average" income – when taken "as is" – there is no immediate solution yet. Let us show, however, that this gauging problem can be reformulated in terms of a problem for which many good solutions have been developed – namely, the problem of group decision making.

To explain this reformulation, let us start with the simplest possible case our main problem: the case when in each of the two compared regions, there is perfect equality: all the people in the first region have the same income x, and all the people in the second region have the same income y. In this case clearly:

– if $x > y$, this means that the first region is in better economic shape, and
– if $x < y$, this means that the second region is in better economic shape.

What if we consider a more realistic case of inequality, when people in the first region have, in general, different incomes $x_1, \ldots, x_n$, and people in the second area also have, in general, different incomes $y_1, \ldots, y_m$? How can we then compare the two regions?

A natural idea is to reduce this comparison to the case when all the incomes are equal. In other words:

- first, we find the value x such that for the group of all the people from the first region, incomes $x_1, \ldots, x_n$ are equivalent – in terms of group decision making – to all of them getting the same income x;
- then, we find the value y such that for the group of all the people from the second region, incomes $y_1, \ldots, y_n$ are equivalent – in terms of group decision making – to all of them getting the same income y;
- finally, we compare the resulting values x and y: if $x > y$, then the first economy is in better shape, otherwise, if $x < y$, the second economy is in better shape.

Comment. Our main idea is to reduce the econometric problem of finding an adequate measure for "average" income" to a game-theoretic problem of cooperative group decision making. This idea is in line with the emerging view that game theory – which was originally invented as a *general* theory of group behavior – should be (and can be) successfully applied not only to situations of conflicting competition, but also to more general problems of economics – in particular, to problems of financial econometrics.

From the idea to the algorithm. To transform the above idea to the algorithm, let us recall a reasonable way to perform group decision making. In group decision making, we need to order situations with different individual incomes. To be more precise, in group decision making, we consider situations with different individual *utility values* $u_1, \ldots, u_n$ – since different people value different income levels differently; see, e.g., [7,16,24,28]. In this case, as shown by the Nobelist John Nash, under some reasonable assumptions, the most adequate solution is to select the alternative for which the product of the utilities $\prod\limits_{i=1}^n u_i$ is the largest possible; see, e.g., [16,23,28].

The utility is usually proportional to a power of the money: $u_i = C_i \cdot x_i^a$ for some $a \approx 0.5$; see, e.g., [12–14]. Substituting these utility values into Nash's formula, we get the product $\prod\limits_{i=1}^n C_i \cdot \prod\limits_{i=1}^n x_i^a$. In these terms, to find the value x for which the selection $(x_1, \ldots, x_n)$ is equivalent to x, we must find x for which

$$\prod_{i=1}^n C_i \cdot \prod_{i=1}^n x_i^a = \prod_{i=1}^n C_i \cdot \prod_{i=1}^n x^a.$$

Dividing both sides of this equality by the constant $\prod\limits_{i=1}^n C_i$ and extracting power a from both sides, we conclude that $\prod\limits_{i=1}^n x_i = \prod\limits_{i=1}^n x = x^n$. Thus, the value x which describes the income distribution $(x_1, \ldots, x_n)$ is equal to $x = \sqrt[n]{x_1 \cdot \ldots \cdot x_n}$ – the *geometric mean* of the income values. So, we arrive at the following conclusion.

Resulting measure of "average" income which most adequately described "average" income: geometric mean. Suppose that we need to compare the economies of two regions. Let us denote the incomes in the first region by $x_1, \ldots, x_n$ and the incomes in the second region by $y_1, \ldots, y_m$. To perform this comparison, we compute the geometric averages $x = \sqrt[n]{x_1 \cdot \ldots \cdot x_n}$ and $y = \sqrt[m]{y_1 \cdot \ldots \cdot y_m}$ of the two regions; then:

- if $x > y$, we conclude that the first region is in better economic shape, and

– if $x < y$, we conclude that the second region is in better economic shape.

From the mathematical viewpoint, comparing geometric means $\overline{x}$ and $\overline{y}$ is equivalent to comparing the logarithms of these means. Here,

$$\ln(\overline{x}) = \ln(\sqrt[n]{x_1 \cdot \ldots \cdot x_n}\,) = \frac{\ln(x_1) + \ldots + \ln(x_n)}{n}.$$

Thus, the logarithm of the geometric mean $\overline{x}$ is equal to mean value $E[\ln(x)]$ of the logarithm of the income – and therefore,

$$\overline{x} = \exp(E[\ln(x)]) = \exp\left(\int \ln(x) \cdot f(x)\,dx\right).$$

So, to compare the economies in two different regions, we need to compare the mean values $E[\ln(x)]$ of the logarithm of the income x in these regions.

Relation between the new measure and the mean income: an observation. It is well known that the geometric mean is always smaller than or equal to the arithmetic mean, and they are equal if and only if all the numbers are equal; see, e.g., [1,34].

Thus, the new measure of "average" income is always smaller than or equal that the mean income, and it is equal to the mean income if and only if all the individual incomes are the same – i.e., if and only if we have perfect equality.

3 The new measure is indeed more robust than the mean income

What is robustness: reminder. Robustness means that if we change each input x_i a little bit, the resulting characteristic of "average" income should not change much.

Effect of systematic and random changes. From the commonsense viewpoint, it is important to distinguish between systematic and random income changes.

Let us first consider the effect of systematic change in income. Systematic change means that each income is changed in exactly the same way. In precise terms, we assume that each income x_i is changed by the exact same percentage, i.e., that instead of the original income x_i, we now have a new income value $x'_i = x_i + r \cdot x_i = x_i \cdot (1 + r)$, for some small number r.

In this case, both the mean $\mu = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$ and the new characteristic $\overline{x} = \prod_{i=1}^{n} x_i^{1/n}$ change the same way:

- instead of the original value μ, we get the new mean $\mu \cdot (1 + r)$, and
- instead of the original value $\overline{x}$, we get the new value $\overline{x} \cdot (1 + r)$.

So, with resect to systematic changes, the new measure behaves exactly the same way as the mean – and, as one can easily see, the same as the median.

Let us now consider the effect of random changes. Let us assume that each original income x_i is replaced by the new value $x_i + r_i \cdot x_i$, where $r_1, \ldots, r_n$ are independent identically distributed random variables, with some mean r and standard deviation σ.

In general, we can have the non-zero mean $r \neq 0$. This general case can be viewed as a combination of two changes: a systematic change corresponding to the mean value r, and a purely random change corresponding to the differences $r_i - r$. Since we have already analyzed the effect of systematic changes, it is sufficient to consider the effect of purely random changes, for which the mean is 0. Thus, without losing generality, we can assume that the mean value of r_i is 0.

Towards a general formula for the effect of random changes. How will a general characteristic $y = f(x_1, \ldots, x_n)$ change under such a random change? When the values x_i change to the new values $x'_i = x_i \cdot (1 + r_i)$, the original value y changes to $y' = f(x'_1, \ldots, x'_n)$. We want to gauge the difference $\Delta y \stackrel{\text{def}}{=} y' - y = f(x'_1, \ldots, x'_n) - f(x_1, \ldots, x_n)$.

In terms of the changes $\Delta x_i \stackrel{\text{def}}{=} x'_i - x_i = r_i \cdot x_i$, we have $x'_i = x_i + \Delta x_i$ and thus,

$$\Delta y = f(x_1 + \Delta x_1, \ldots, x_n + \Delta x_n) - f(x_1, \ldots, x_n).$$

We are interested in the case when the changes are small. In this case, we can expand the above expression in terms of Δx_i and safely ignore terms which are quadratic (or higher order) in terms of these small changes. As a result, we get the formula

$$\Delta y = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \cdot \Delta x_i = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \cdot x_i \cdot r_i.$$

This formula is a linear combination of a large number of independent random variables r_i. Thus, according to the Central Limit Theorem (see, e.g., [32]), the distribution of Δy is close to Gaussian. A Gaussian distribution is uniquely determined by its mean μ_y and it standard deviation σ_y. Since the mean of each r_i is 0, the mean of the linear combination Δy is also 0, so $\mu_y = 0$. Thus, it is sufficient to only compute σ_y. Since the variables r_i are independent with the same standard deviation σ, we get

$$\sigma_y^2 = \sigma^2 \cdot \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)^2 \cdot x_i^2.$$

Resulting comparison. Let us apply this general formula to the function $f(x_1, \ldots, x_n) = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$ corresponding to the mean and to the function $f(x_1, \ldots, x_n) = \prod_{i=1}^{n} x_i^{1/n}$ describing the new measure.

For the mean-related function, $\frac{\partial f}{\partial x_i} = \frac{1}{n}$ and thus,

$$\sigma_y^2 = \sigma^2 \cdot \frac{1}{n} \cdot \left(\frac{1}{n} \cdot \sum_{i=1}^{n} x_i^2\right).$$

For the function corresponding to the new measure, we have $\dfrac{\partial f}{\partial x_i} = \dfrac{1}{n} \cdot \dfrac{f}{x_i}$, thus

$$\sigma_y^2 = \sigma^2 \cdot n \cdot \left(\frac{1}{n^2} \cdot f^2 \right) = \sigma^2 \cdot \frac{1}{n} \cdot \left(\prod_{i=1}^n x_i^{2/n} \right).$$

Thus, to decide which standard deviation is larger, we need to compare the arithmetic mean $\dfrac{1}{n} \cdot \sum\limits_{i=1}^n x_i^2$ of n values $x_1^2, \ldots, x_n^2$ with the geometric mean $\prod\limits_{i=1}^n (x_i^2)^{1/n}$ of these same numbers.

As we have mentioned in the previous section, the geometric mean is always smaller than or equal to the arithmetic mean, and the only case when these two means are equal is when all the values x_i^2 (and thus, all the values x_i) are equal; see, e.g., [1,34]. Thus, the new measure $\overline{x}$ is more robust than the arithmetic mean income μ.

4 First example of using the new measure of "average" income: case of low inequality

Case of low inequality: informal description. Let us first consider the case when inequality is low, i.e., when most people have a reasonable income, and the proportion of very poor and very rich people is not that large.

Towards a formal description. The fact that most incomes are close to one another means that most of these incomes are close to the mean income μ. In mathematical statistics, deviations from the mean are usually described by the standard deviation σ; see, e.g., [32]. In these terms, low inequality means that the standard deviation σ is small. Let us analyze what happens in this case.

Case of low inequality: analysis of the problem. As we have mentioned, the new inequality measure has the form $\overline{x} = \exp(E[\ln(x)])$. Thus, to compare the economies in two different regions, we need to compare the mean values $E[\ln(x)]$ of the logarithm of the income x in these regions.

Since the deviations from the mean $x - \mu$ are relatively small, we have can substitute $x = \mu + (x - \mu)$ into the formula for $E[\ln(x)]$ and ignore higher order terms in the expansion in $x - \mu$. According to the Taylor series for the logarithm, we have:

$$\ln(x) = \ln(\mu + (x - \mu)) = \ln(\mu) + \frac{1}{\mu} \cdot (x - \mu) - \frac{1}{2\mu^2} \cdot (x - \mu)^2 + \ldots$$

By taking the mean value of both sides and taking into account that $E[x - \mu] = \mu - \mu = 0$ and that $E[(x - \mu)^2] = \sigma^2$, we conclude that

$$E[\ln(x)] = \ln(\mu) - \frac{1}{2\mu^2} \cdot \sigma^2 + \ldots$$

Since we assumed that the deviations of x from μ are small, we can preserve only the first terms which shows the dependence on these deviations and ignore higher order terms in this expansion. As a result, we get an approximate formula

$$E[\ln(x)] \approx \ln(\mu) - \frac{\sigma^2}{2\mu^2}.$$

Thus, for $\overline{x} = \exp(E[\ln(x)])$, we get

$$\overline{x} = \exp(E[\ln(x)]) \approx \exp\left(\ln(\mu) - \frac{\sigma^2}{2\mu^2}\right) = \exp(\ln(\mu)) \cdot \exp\left(-\frac{\sigma^2}{2\mu^2}\right).$$

The first factor is equal to μ. To estimate the second factor, we can again use the fact that σ is small; in this case, we can expand the function $\exp(z)$ in Taylor series and keep only the first term depending on σ:

$$\exp\left(-\frac{\sigma^2}{2\mu^2}\right) = 1 - \frac{\sigma^2}{2\mu^2} + \ldots \approx 1 - \frac{\sigma^2}{2\mu^2}.$$

Substituting this expression into the above formula for $\overline{x} = \exp(E[\ln(x)])$, we conclude that

$$\overline{x} = \mu \cdot \left(1 - \frac{\sigma^2}{2\mu^2}\right) = \mu - \frac{\sigma^2}{2\mu}.$$

Thus, we arrive at the following conclusion.

Resulting formula. In the case of low inequality, the "average" income is equal to

$$\overline{x} = \mu - \frac{\sigma^2}{2\mu},$$

where μ is the average income and σ is the standard deviation.

Analysis of this formula. The larger the inequality, the larger the standard deviation σ, and the less preferable is the economy. The above formula provides an exact quantitative description of this natural qualitative idea.

Comments.

- The new measure takes inequality into account, and it avoids the ideological ideas of weighing inequality too much: if an increase in the mean income comes at the expense of an increase in inequality, this is OK, as long as the above combination of means and standard deviation increases.
- This example is one of the cases which shows that the new measure is more adequate than, e.g., the median. For example, if the incomes are normally distributed, then the median simply coincides with the mean, and so, contrary to our intuitive expectations, the increase in inequality does not worse the median measure of economics. In contrast, the new measure does go down when inequality increases.

5 Second example of using the new measure of "average" income: case of a heavy-tailed distribution

Heavy-tailed (usually, Pareto) distributions are ubiquitous in economics. In the 1960s, Benoit Mandelbrot, the author of fractal theory, empirically studied the price fluctuations and showed [17] that large-scale fluctuations follow the Pareto power-law distribution, with the probability density function $f(x) = A \cdot x^{-\alpha}$ for $x \geq x_0$, for some constants $\alpha \approx 2.7$ and x_0. For this distribution, variance is infinite. The above empirical result, together with similar empirical discovery of heavy-tailed laws in other application areas, has led to the formulation of *fractal theory*; see, e.g., [18,19].

Since then, similar Pareto distributions have been empirically found in other financial situations [3,4,6,8,20,22,27,30,33,35,36], and in many other application areas [2,9,18,21,29].

Formulation of the problem. Let us consider the situations when the income distribution follows Pareto law, with probability density $f(x)$ equal to 0 for $x \le x_0$ and to $A \cdot x^{-\alpha}$ for $x \ge x_0$.

Once we know x_0 and α, we can determine the parameter A from the condition that $\int f(x)\,dx = 1$. For the above expression, this condition leads to $A \cdot \dfrac{x_0^{-(\alpha-1)}}{\alpha - 1} = 1$, hence $A = (\alpha - 1) \cdot x_0^{\alpha-1}$.

For this distribution, we want to compute the mean income, the median income, and the newly defined "average" income.

Mean income. The mean income is equal to $\mu = \int x \cdot f(x)\,dx$, i.e., for the Pareto distribution, to

$$\int_{x_0}^{\infty} A \cdot x^{1-\alpha}\,dx = A \cdot \left.\frac{x^{2-\alpha}}{2-\alpha}\right|_{x_0}^{\infty} = A \cdot \frac{x_0^{2-\alpha}}{\alpha - 2}.$$

Substituting the above value of A, we conclude that the mean is equal to

$$\mu = \frac{\alpha - 1}{\alpha - 2} \cdot x_0.$$

Median income. The median income m can be determined from the condition that $\int_m^\infty f(x)\,dx = \dfrac{1}{2}$. For the Pareto distribution, this means

$$\int_m^\infty A \cdot x^{-\alpha}\,dx = A \cdot \frac{m^{-(\alpha-1)}}{\alpha - 1} = \frac{1}{2}.$$

Substituting the above expression for A into this formula, we conclude that $\dfrac{m^{-(\alpha-1)}}{x_0^{-(\alpha-1)}} = \dfrac{1}{2}$, hence $\dfrac{m^{\alpha-1}}{x_0^{\alpha-1}} = 2$, and $m = x_0 \cdot 2^{1/(\alpha-1)}$.

New measure of "average" income. For the new measure of average income $\overline{x}$, its logarithm is equal to the expected value of $\ln(x)$:

$$\ln(\overline{x}) = \int \ln(x) \cdot f(x)\,dx = \int_{x_0}^\infty \ln(x) \cdot A \cdot x^{-\alpha}\,dx.$$

This integral can be computed by integration by part; so, we get

$$\ln(\overline{x}) = \ln(x) \cdot \left.\frac{A \cdot x^{1-\alpha}}{1-\alpha}\right|_{x_0}^{\infty} - \int_{x_0}^\infty \frac{1}{x} \cdot \frac{A \cdot x^{1-\alpha}}{1-\alpha}\,dx =$$

$$\ln(x_0) \cdot \frac{A \cdot x_0^{-(\alpha-1)}}{\alpha - 1} - \int_{x_0}^\infty \frac{A \cdot x^{-\alpha}}{1-\alpha}\,dx =$$

$$\ln(x_0)\cdot\frac{A\cdot x_0^{-(\alpha-1)}}{\alpha-1}-\left.\frac{A\cdot x^{-(\alpha-1)}}{(1-\alpha)^2}\right|_{x_0}^{\infty}=$$

$$\ln(x_0)\cdot\frac{A\cdot x_0^{-(\alpha-1)}}{\alpha-1}+\frac{A\cdot x_0^{-(\alpha-1)}}{(1-\alpha)^2}.$$

Substituting the expression $A=(\alpha-1)\cdot x_0^{\alpha-1}$ into this formula, we get

$$\ln(\overline{x})=\ln(x_0)+\frac{1}{\alpha-1},$$

hence

$$\overline{x}=\exp(\ln(\overline{x}))=x_0\cdot\exp\left(\frac{1}{\alpha-1}\right).$$

Comment. When $\alpha\to\infty$, the distribution tends to be concentrated on a single value x_0 – i.e., we have the case of absolute equality. In this case, as expected, all three characteristics – the mean, the median, and the new geometric mean – tends to the same value x_0.

6 First auxiliary result: the new measure of "average" income may explain Deaton's empirical observation

Deaton's empirical observation. In his book [5], the Nobelist Angus Deaton mentions an interesting fact: that to understand the relation between income and other important characteristics q like life expectancy, clearer results are obtained if:

- instead of studying the correlation between q and the *income* x_i,
- we study the correlation between q and the *logarithm* $\log(x_i)$ of the income.

Deaton's observation explained. What a natural way to gauge the difference between two income levels $x<y$? A natural idea is to fix some kind of a step, and to gauge the difference between x and y as the number of steps needed to get from x to y.

The first step can be chosen arbitrarily: e.g., we go from the original income $u_0=x$ to a slightly larger number $u_1=x\cdot(1+\delta)$, for some small $\delta>0$. How do we define the next step, going from u_1 to some other value u_2? A reasonable idea is to select u_2 in such a way that

- the change from u_1 to u_2 would lead to exactly the same change in the "average" income $\overline{x}$
- as the original change from u_0 to u_1.

In other words, if we originally has two people with incomes $x_1=u_0$ and $x_2=u_1$, so that the overall "average" is equal to $\overline{x}=u_0^{1/n}\cdot u_1^{1/n}\cdot\prod\limits_{i=3}^{n}x_i^{1/n}$, then we should get the same new value of $\overline{x}$ in two situations:

- if we increase x_1 from u_0 to u_1, leaving all other incomes unchanged, and
- if we increase x_2 from u_1 to u_2, leaving all other incomes unchanged:

$$u_1^{1/n} \cdot u_1^{1/n} \cdot \prod_{i=3}^{n} x_i^{1/n} = u_0^{1/n} \cdot u_2^{1/n} \cdot \prod_{i=3}^{n} x_i^{1/n}.$$

Substituting $u_1 = u_0 \cdot (1+\delta)$ into the left-hand side of this formula, we conclude that the value $\overline{x}$ gets multiplied by a factor $(1+\delta)^{1/n}$. Thus, we must select u_2 for which

$$u_0^{1/n} \cdot u_2^{1/n} \cdot \prod_{i=3}^{n} x_i^{1/n} = u_0^{1/n} \cdot u_1^{1/n} \cdot \prod_{i=3}^{n} x_i^{1/n} \cdot (1+\delta)^{1/n}.$$

Dividing both sides of this equality by $u_0^{1/n} \cdot \prod\limits_{i=3}^{n} x_i^{1/n}$, we get $u_2^{1/n} = u_1^{1/n} \cdot (1+\delta)^{1/n}$, hence $u_2 = u_1 \cdot (1+\delta)$. Substituting $u_1 = u_0 \cdot (1+\delta)$ into this formula, we get $u_2 = u_0 \cdot (1+\delta)^2$.

Similarly, if we want the transition from u_2 to u_3 to have the same effect as the transitions from u_0 to u_1 and from u_1 to u_2, we need to take $u_3 = u_2 \cdot (1+\delta) = u_0 \cdot (1+\delta)^3$. In general, we get $u_k = u_0 \cdot (1+\delta)^k$.

Now we can compute the number of steps k needed to go from $u_0 = x$ to $u_k = y > x$: it is determined from the formula $y = u_k = u_0 \cdot (1+\delta)^k = x \cdot (1+\delta)^k$. Thus, $(1+\delta)^k = \dfrac{y}{x}$, and thus, the number of steps k is proportional to $\log\left(\dfrac{y}{x}\right) = \log(y) - \log(x)$.

If we fix the starting income x, then we get a natural measure of income y: the number of steps needed to reach y – which is thus a linear function of $\log(y)$. This explains why the logarithm of the money amount is a natural measure of income.

7 Second auxiliary result: the new measure of "average" income may explain the power-law character of income distribution

In the previous section, we analyzed how the new measure of "average" income $\overline{x} = \exp(\int \ln(x) \cdot f(x)\,dx)$ behaves in situations when the income distribution follows a power law.

Interestingly, the power law itself can be derived based on this inequality measure. Indeed, suppose that all we know about the income distribution is the value $\overline{x}$, and the lower bound $\delta > 0$ on possible incomes (this lower bound reflects the fact that a human being needs some minimal income to survive). There are many possible probability distributions $f(x)$ which are consistent with this information. In such situation, out of all such distributions, it is reasonable to select a one for which the entropy $S \stackrel{\text{def}}{=} -\int f(x) \cdot \ln(f(x))\,dx$ is the largest; see, e.g., [11].

To find the distribution that maximizes the entropy S under the constraints $\exp\left(\int \ln(x) \cdot f(x)\,dx\right) = \overline{x}$ and $\int f(x)\,dx = 1$, we can use the Lagrange multiplier technique that reduces this constraint optimization problem to the unconstrained problem of optimizing a functional

$$-\int f(x) \cdot \ln(f(x))\,dx + \lambda_1 \cdot \left(\exp\left(\int \ln(x) \cdot f(x)\,dx\right) - \overline{x}\right) + \lambda_2 \cdot \left(\int f(x)\,dx - 1\right),$$

for appropriate Lagrange multipliers λ_i. Differentiating this expression with respect to $f(x)$ and equating the derivative to 0, we conclude that

$$-\ln(f(x)) - 1 + \lambda_1 \cdot C \cdot \ln(x) + \lambda_2 = 0,$$

where $C \stackrel{\text{def}}{=} \exp(\int \ln(x) \cdot f(x)\,dx)$ and thus $C = \overline{x}$. Thus,

$$\ln(f(x)) = (\lambda_2 - 1) + \lambda_1 \cdot \overline{x} \cdot \ln(x).$$

Applying the function $\exp(z)$ to both sides of this equality, we conclude that $f(x) = A \cdot x^{-\alpha}$, where $A = \exp(\lambda_2 - 1)$ and $\alpha = -\lambda_1 \cdot \overline{x}$. So, we indeed get the empirically observed power law for income distribution.

Acknowledgments

The authors are thankful to all the participants of the 2015 Conference of the Thailand Econometric Society for valuable discussions.

References

1. E. F. Beckenbach and R. Bellman, Introduction to inequalities. Mathematical Association of America, New York (1975)
2. J. Beirlant, Y. Goegevuer, J. Teugels, and J. Segers. Statistics of extremes: theory and applications. Wiley, Chichester (2004)
3. B. K. Chakrabarti, A. Chakraborti, and A. Chatterjee, Econophysics and sociophysics: trends and perspectives. Wiley-VCH, Berlin (2006)
4. A. Chatterjee, S. Yarlagadda, and B. K. Chakrabarti, Econophysics of wealth distributions. Springer-Verlag Italia, Milan (2005)
5. A. Deaton, The great escape: health, wealth, and the origins of inequality. Princeton University Press, Princeton, New Jersey, and Oxford, UK (2013)
6. J. D. Farmer and T. Lux (eds.). Applications of statistical physics in economics and finance, a special issue of the Journal of Economic Dynamics and Control, Vol. 32, No. 1, pp. 1–320 (2008)
7. P. C. Fishburn, Nonlinear preference and utility theory. John Hopkins Press, Baltimore, Maryland (1988)
8. X. Gabaix, G. Parameswaran, P. Vasiliki, and H. E. Stanley, Understanding the cubic and half-cubic laws of financial fluctuations, Physica A, Vol. 324, pp. 1–5 (2003)
9. C. P. Gomez and D. B. Shmoys, Approximations and randomization to boost CSP techniques, Annals of Operations Research, Vol. 130, pp. 117–141 (2004)
10. P. J. Huber, Robust statistics. Wiley, Hoboken, New Jersey (2004)
11. E. T. Jaynes, Probability theory: the logic of science. Cambridge University Press, Cambridge, Massachusetts (2003)
12. D. Kahneman, Thinking: fast and slow. Farrar, Straus, and Girroux, New York (2011)
13. D. Kahneman and A. Tversky, Advances in prospect theory: cumulative representation of uncertainty, Journal of Risk and Uncertainty, Vol. 5, pp. 297–324 (1992)
14. J. Kozielecki, Psychological decision theory. Kluwer, Dordrecht, Netherlands (1981)
15. V. Kreinovich, H. T. Nguyen, and R. Ouncharoen, From mean and median income to the most adequate way of taking inequality into account, Proceedings of the 2015 Conference of Thailand Econometric Society, pp. 63–73 (2015)
16. R. D. Luce and R. Raiffa, Games and decisions: introduction and critical survey. Dover, New York (1989)
17. B. Mandelbrot, The variation of certain speculative prices, J. Business, Vol. 36, pp. 394–419 (1963)
18. B. Mandelbrot, The fractal geometry of nature. Freeman, San Francisco, California (1983)

19. B. Mandelbrot and R. L. Hudson, The (mis)behavior of markets: a fractal view of financial turbulence. Basic Books (2006)
20. R. N. Mantegna and H. E. Stanley, An introduction to econophysics: correlations and complexity in finance. Cambridge University Press, Cambridge, Massachusetts (1999)
21. N. Markovich (ed.), Nonparametric analysis of univariate heavy-tailed data: research and practice. Wiley, Chichester (2007)
22. J. McCauley, Dynamics of markets, econophysics and finance. Cambridge University Press, Cambridge, Massachusetts (2004)
23. J. Nash, The bargaining problem, Econometrica, Vol. 18, No. 2, pp. 155–162 (1950)
24. H. T. Nguyen, V. Kreinovich, B. Wu, and G. Xiang, Computing statistics under interval and fuzzy uncertainty. Springer Verlag, Berlin, Heidelberg (2012)
25. Organization for Economic Cooperation and Development (OECD), OECD Statistics, available at http://stats.oecd.org/
26. T. Piketty, Capital in the twenty-first century. Harvard University Press, Cambridge, Massachusetts (2014)
27. S. T. Rachev and S. Mittnik, Stable Paretian models in finance. Wiley Publishers, New York (2000)
28. H. Raiffa, Decision analysis. McGraw-Hill, Columbus, Ohio (1997)
29. S. I. Resnick, Heavy-tail phenomena: probabilistic and statistical modeling. Springer-Varlag, New York (2007)
30. B. Roehner, Patterns of speculation – a study in observational econophysics. Cambridge University Press, Cambridge, Massachusetts (2002)
31. P. Rousseeuw and A. Leroy, Robust regression and outlier detection. J. Wiley, New York (1987)
32. D. J. Sheskin, Handbook of parametric and nonparametric statistical procedures. Chapman & Hall/CRC, Boca Raton, Florida (2011)
33. H. E. Stanley, L. A. N. Amaral, P. Gopikrishnan, and V. Plerou, Scale invariance and universality of economic fluctuations, Physica A, 2000, Vol. 283, pp. 31–41 (2000)
34. J. M. Steele, The Cauchy-Schwarz master class: an introduction to the art of mathematical inequalities. Cambridge University Press, New York (2004)
35. S. V. Stoyanov, B. Racheva-Iotova, S. T. Rachev, and F. J. Fabozzi, Stochastic models for risk estimation in volatile markets: a survey, Annals of Operations Research, Vol. 176, pp. 293–309 (2010)
36. P. Vasiliki and H. E. Stanley, Stock return distributions: tests of scaling and universality from three distinct stock markets, Physical Review E: Statistical, Nonlinear, and Soft Matter Physics, Vol. 77, No. 3, Pt. 2, Publ. 037101 (2008)