Physica A 469 (2017) 824–847

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Inequality spectra

Iddo Eliazar

Smart Device Innovation Science Team, New Devices Group, Intel Corporation, Yakum, Israel

ARTICLE INFO

Article history:
Received 12 August 2016
Available online 25 November 2016

Keywords:
Statistical heterogeneity
Socioeconomic inequality
Lorenz curves
Gini index
Rényi index
Hill curves

ABSTRACT

Inequality indices are widely applied in economics and in the social sciences as quantitative measures of the socioeconomic inequality of human societies. The application of inequality indices extends to size-distributions at large, where these indices can be used as general gauges of statistical heterogeneity. Moreover, as inequality indices are plentiful, arrays of such indices facilitate high-detail quantification of statistical heterogeneity. In this paper we elevate from arrays of inequality indices to *inequality spectra*: continuums of inequality indices that are parameterized by a single control parameter. We present a general methodology of constructing Lorenz-based inequality spectra, apply the general methodology to establish four sets of inequality spectra, investigate the properties of these sets, and show how these sets generalize known inequality gauges such as: the Gini index, the extended Gini index, the Rényi index, and hill curves.

1. Introduction

Size distributions, i.e. statistical distributions of non-negative valued quantities, are abundant all across science and engineering. Examples of sizes include count, length, area, volume, duration, mass, and energy. The most widely applied measures of the statistical heterogeneity of a given size distribution are its standard-deviation and variance, the latter being the square of the former. The most common information-based measures of the statistical heterogeneity of a given size distribution are its perplexity and Boltzmann–Gibbs–Shannon entropy, the latter being the logarithm of the former.

Two main size distributions in economics and the social sciences are wealth and income distributions [1–3]. Economists and social scientists use gauges called *inequality indices* – rather than the aforementioned standard-deviation and perplexity – to measure the statistical heterogeneity of wealth and income distributions [4–6]. Indeed, economists and social scientists are interested in assessing the intrinsic socioeconomic inequality of wealth and income distributions, and find inequality indices to be better suited for this task. The *Gini index* is apparently the most commonly applied inequality index [7–9].

The standard-deviation and the perplexity are non-negative valued measures. Inequality indices, on the other hand, take values in the unit interval. Thus, the standard-deviation and the perplexity are analogous of temperature, whereas inequality indices are analogous of test scores. Moreover, the standard-deviation and the perplexity are unique, whereas inequality indices are plentiful, as we shall now explain.

The standard-deviation stems from Euclidean geometry – which, in the context of Euclidean spaces, is the only geometry that is founded on the notion of orthogonality. The perplexity follows from the Boltzmann–Gibbs–Shannon entropy – which is the only measure of information that stems from a given set of underlying 'natural' axioms [10–12]. Thus, both the standard-deviation and the perplexity are unique measures in their respective contexts.

Similarly to the Boltzmann–Gibbs–Shannon entropy, also inequality indices have to meet a given set of underlying 'natural' axioms. However, contrary to the Boltzmann–Gibbs–Shannon entropy, there are many – in fact infinitely many

E-mail addresses: iddo.eliazar@intel.com, eliazar@post.tau.ac.il.

http://dx.doi.org/10.1016/j.physa.2016.11.079

– inequality indices that meet these axioms [13]. Consequently, the socioeconomic inequality of wealth and income distributions can be measured in a variety of 'natural' ways.

Inequality indices were devised in the context of wealth and income distributions. However, from an abstract statistical perspective, inequality indices can be applied to size distributions at large. And indeed, the most popular inequality index – the aforementioned Gini index – is used in diverse fields of science. Recent examples of Gini-index applications outside economics and the social sciences include: bacterial chemotaxis [14], interface friction [15], crowd science [16], RNA regulatory mechanisms [17], stem cell differentiation [18], clean energy [19], cancer mutations [20], maternal and pediatric health and disease [21], cell transplants [22], and cosmological lensing [23]. For more applications of the Gini index see [24] and references therein.

In the dominion of size distributions, inequality indices offer the following multi-dimensional approach: instead of applying the standard-deviation or the perplexity and thus obtain a one-dimensional measurement of statistical heterogeneity—apply an array of inequality indices to obtain a multi-dimensional measurement of statistical heterogeneity. This multi-dimensional approach was detailed and advocated in [13].

To metaphorically picture the difference between the one-dimensional standard-deviation and perplexity approaches on the one hand, and the multi-dimensional inequality approach on the other hand, envisage a wall thermometer and a music equalizer. The wall thermometer is a metaphor for a single unbounded gauge—temperature, standard-deviation, perplexity, etc. The music equalizer, on the other hand, is a metaphor for an array of bounded gauges which are normalized to take values in the unit interval—inequality indices in our case.

The multi-dimensional approach can be elevated from arrays of inequality indices to *continuums* of inequality indices. One-parameter families of inequality indices include: (i) the *Atkinson index*, which is based on the Hölder mean [25]; (ii) the *extended Gini index*, which generalizes the Gini index [26–29]; (iii) the *Rényi index*, which is based on the Rényi entropy [30]; and (iv) the *hill curves*, which are based on the Lorenz set [31]. One-parameter families of inequality indices are henceforth termed *inequality spectra*.

On the one hand, inequality spectra offer a key feature which any single inequality index will fail to provide: sensitivity. On the other hand, as inequality spectra are controlled by a single parameter they offer yet another key feature: a graphic visualization. Specifically, in a graph of a given inequality spectrum the x-axis represents the control-parameter value, and the y-axis represents the inequality-index value. Inequality spectra are nonlinear transformations of size-distributions, and they provide a high-detail quantification of the intrinsic inequality of size-distributions.

In this paper we shall explore and study inequality spectra from a unified Lorenz-curves perspective. *Lorenz curves* are a widely applied methodology in the context of wealth and income distributions, and – just as in the case of inequality indices – they can be further applied in the context of size distributions at large [24,32–34]. Lorenz curves present the information that is coded by size-distributions in a universally calibrated socioeconomic perspective. Moreover, the Lorenz curves facilitate a *sociogeometric* perspective to the measurement of socioeconomic inequality [35,36].

Interestingly, Lorenz curves can be perceived as statistical laws defined on the unit interval. Exploiting this perspective, we may consider various functionals of these laws—henceforth termed *Lorenzian statistical laws*. And indeed, we shall use the moments, the Laplace transforms, the distribution functions, and the Rényi entropies of the Lorenzian statistical laws to establish four sets of inequality spectra. The path we shall follow in this paper is as follows.

We set off from the *Gini index* (Section 2), and from the notions of *Lorenz curves* and *sociogeometric inequality indices* (Section 3). Then, we devise a general method of constructing sociogeometric inequality indices from a given generalized-moment of the Lorenzian statistical laws (Section 4). Applying the general method of Section 4 to the moments of the Lorenzian statistical laws we establish the *moment spectra*, and show how these spectra generalize the Gini index and the extended Gini index (Section 5). Applying the general method of Section 4 to the Laplace transforms of the Lorenzian statistical laws we establish the *Laplace spectra*, and unveil the Poisson foundation of these spectra (Section 6). Also, applying the general method of Section 4 to the distribution functions of the Lorenzian statistical laws we establish the *Heaviside spectra*, and show how these spectra generalize the hill curves (Section 7). Finally, we address the Rényi entropy of the Lorenzian statistical laws and construct from it the *Rényi spectra*, which generalize the Rényi index (Section 8). We end the paper with a short conclusion (Section 9), followed by an Appendix.

In the recent years there is a growing interest – within physics communities in general and among the statistical-physics community in particular – in the study of economic and social systems and processes. This interest gave rise to the new fields of *econophysics* and *sociophysics* [37–43]. Commonly, statistical-physics methods are imported to economics and the social sciences, and are applied in order to model socioeconomic phenomena. In this paper we go the reverse direction: we take the socioeconomic notion of inequality indices, extend it to the more general notion of inequality spectra, and propose the application of inequality spectra to size-distributions at large—in particular to size-distributions in the physical sciences.

2. Gini index

In this paper we consider general size distributions with finite means, which we henceforth deem to be wealth distributions. Specifically, we consider the following wealth-distribution setting: an arbitrary human society comprising of a population of members, each society member is assumed to have a non-negative wealth, and the members' average