The Lorenz Curve: Evergreen after 100 years

Barry C. Arnold
University of California, Riverside

ABSTRACT

Lorenz’s (1905) suggestion of a graphical manner in which to compare inequality in finite populations in terms of nested curves turned out to be remarkably well accepted. We continue to learn about and to utilize concepts derived from and intimately related to this curve. There is still work to be done related to Lorenz ordering especially in higher dimensions. Moreover there exist many perhaps unexpected areas in which Lorenz ordering ideas can be profitably used. A small sample of such scenarios is included herein.

1 Introduction

In 1905, Max Lorenz proposed a simple graphical means to summarize the inequality of wealth in a finite population of individuals. Known subsequently as the Lorenz curve, it has survived well and indeed still occupies a preeminent place in discussion of the quantification of inequality. It was a simple, but a very good, idea. Subsequent investigations have provided useful interpretations of why it does so well in capturing our conceptions of what really constitutes inequality. Some of these insights will be discussed below. The mathematical concept known as majorization arrived somewhat later on the scene. Its close relationship with the ordering proposed by Lorenz in his pioneering paper has been apparent for many years, although it is difficult to pinpoint precisely when this nexus was first noted. Nevertheless, the deep understanding of the majorization partial order developed initially by Hardy, Littlewood and Polya (1927, 1934), has been seamlessly transformed to give us a spectrum of useful results and viewpoints on the “true nature” of Lorenz’s curve and its associated ordering of inequality in distinct populations or in populations viewed at different points in time. An important contribution to our understanding of the Lorenz curve and an important
reason for its continued and growing acceptance among economists, was found in Dalton’s (1920) careful discussion of criteria that might be arguably accepted as being clearly desirable features of any measure of inequality and of any method for comparison of inequality between populations.

The Lorenz curve may well have flourished without the contributions of Dalton, Hardy, Littlewood and Polya but it could not fail to flourish in the presence of such inputs. Indeed it seems that very few people question the fact that nested Lorenz curves signal a clear differential in inequality. What has kept the field of income inequality full of lively controversy is the question of what to do when, or how to interpret situations in which, Lorenz curves cross.

The first focus of the present paper will be on the role of the Lorenz curve in an income inequality context. But it must, and will, be remarked that inequality (often with a different name such as diversity, variability, etc.) is of interest in many other contexts. Moreover, thanks to the development of the mathematical consequences of majorization, it has been apparent that the majorization partial ordering or, almost equivalently, the Lorenz order has an ever expanding role to play in the study of a spectrum of optimization settings that involve unexpected appearances of Schur convex functions.

The need for a multivariate version of Lorenz’s curve and his partial order rapidly become apparent. But how to achieve this desirable goal proved to be elusive. In the mathematics literature, the parallel problem of defining multivariate majorization was receiving attention. In the mathematical context several variant interpretations of plausible multivariate majorization concepts were introduced but no productive suggestions were made regarding how to provide a multivariate version of the graphical tool that Lorenz had provided in giving us his “curve”. Graphical techniques are of course limited by the dimensions available to us when graphing, but it was clear that some graphical extension of Lorenz’s curve should be available at least in the bivariate case. Lorenz’s curve was to celebrate its 90th birthday however before such a suitable extension was made available. The resulting Lorenz zonoid and Lorenz zonoid ordering, provided initially by Koshevoy (1995) and
investigated thoroughly by Koshevoy and Mosler (Mosler (2002)), has infused new multidimensional vigor into the evergreen Lorenz curve concept. Lively future development can be confidently predicted as we argue about what to do when Lorenz zonoids are not nested, assuming that we generally agree that nested zonoids do reflect a basic inequality ordering in higher dimensions that successfully mirrors the univariate ordering originally proposed by Lorenz. “As the bow is bent concentration increases”.

2 Lorenz’s curve (1905)

It all goes back to a brief paper published by Max Lorenz in the Publication of the American Statistical Association (later to be known as the Journal of the American Statistical Society) in June 1905 (Lorenz (1905)). The paper is only 9 pages long. The major thrust of the paper is the presentation of discussion of the spectrum of summary inequality measures being used at that time. He recognizes the fact that logarithmic analysis of income distributions has been popularized (especially by Pareto) but he regards logarithmic curves as “treacherous”. In the last 3 pages of the paper he describes what will become the Lorenz curve. Actually there are only 35 lines of text and two diagrams devoted to the topic. It has all grown from that! These Lorenz curves provided originally by Lorenz will look a little strange to the modern reader. He has the axes interchanged (or rather, subsequently authors have interchanged the axes on his curves).

Lorenz began with a data set which provided, for a small selection of values, the proportion of the total population earning less than the given value together with the proportion of the total wealth of the population accruing to those individuals. The percentage of the population was plotted against the $y$ axis and the proportion of the total wealth was plotted against the $x$ axis. He then joined the points by a smooth curve, though he gave no hint as to how this interpolating curve was selected. Today we would likely interchange the axes in the diagram and interpolate linearly to obtain what we call the Lorenz curve.
If we have data available on every member of a finite population of \( n \) individuals then we can identify the Lorenz curve as being one defined by first ordering the wealths of the individuals from smallest to largest (denoted by \( x_{1:n}, x_{2:n}, \ldots, x_{n:n} \)) and then plotting the points
\[
\left( \frac{j}{n}, \frac{\sum_{i=1}^{j} x_{i:n}}{\sum_{i=1}^{n} x_{i:n}} \right), \quad j = 1, 2, \ldots, n.
\]
In addition we plot the point \((0,0)\) and linearly interpolate the \( n + 1 \) points to obtain the familiar bow shape curve. If two populations have nested Lorenz curves, i.e. if the bow it bent more for one of the populations, then a clear indication is provided that one population exhibits more inequality than the other. If we denote the two population vectors by \( \underline{x} \) and \( \underline{y} \) and their corresponding Lorenz curves by \( L_x(u) \) and \( L_y(u) \) then our assertion is that if \( L_x(u) \leq L_y(u), \quad \forall u \in (0,1) \) then \( \underline{x} \) exhibits at least as much inequality as \( \underline{y} \). It may be remarked that use of Lorenz curves allows comparison between populations of different sizes.

To Lorenz it was clear that the Lorenz ordering was a sensible way to quantify inequality orderings. But he really gave no insight as to why it should be self-evident or, if you wish, why it is even plausible. To justify the paramount role that the Lorenz order has played in inequality discussions, we need to move ahead to relate it to Dalton’s (1920) inequality principles. But before doing so, we will sidestep to consider the parallel concept of majorization.

3 Majorization

At some date prior to 1929, the partial ordering on \( \mathbb{R}^n \) known as majorization was introduced. By 1929, Hardy, Littlewood and Polya were showing its relationship to Schur’s averages but they did not use the name majorization. The sufficient conditions for majorization were known as the Muirhead condition (though Muirhood (1903) introduced the condition in the context of vectors of integers only).

For any vector \( \underline{x} \in \mathbb{R}^n \) we will denote its coordinates written in non-decreasing order by
We will write our results in terms of an ordering from smallest to largest (to mesh nicely with the Lorenz ordering associated with Lorenz’s curve) but it must be remarked that, in the mathematics literature, most of the discussion of majorization involves vectors arranged from largest to smallest.

Definition 3.1: Let \( \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \). We will say that \( \mathbf{x} \) is majorized by \( \mathbf{y} \) and write \( \mathbf{x} \prec \mathbf{y} \) if

\[
\sum_{i=1}^{k} x_{i:n} \geq \sum_{i=1}^{k} y_{i:n}, \quad k = 1, 2, \ldots, n - 1
\]

and

\[
\sum_{i=1}^{n} x_{i:n} = \sum_{i=1}^{n} y_{i:n}.
\]

In (1923) Schur introduced the concept of an “averaging”.

Definition 3.2: Let \( \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \). We will say that \( \mathbf{x} \) is an averaging of \( \mathbf{y} \) if there exists an \( n \times n \) doubly stochastic matrix \( P \) such that

\[
\mathbf{x} = P \mathbf{y}
\]

(the vectors involved here are interpreted as column vectors). Functions which are monotone with respect to majorization are called Schur convex functions.

Definition 3.3: Let \( A \subset \mathbb{R}^n \). A function \( g : A \to \mathbb{R} \) is said to be Schur convex on \( A \) if \( g(x) \leq g(y) \) for every pair \( x, y \in A \) for which \( x \prec y \).

The following basic results on majorization may be found in Hardy, Littlewood and Polya (1934) or, for a more modern treatment, in Marshall and Olkin (1979).

Theorem 3.4. Let \( \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \). The following are equivalent

(i) \( \mathbf{x} \prec \mathbf{y} \).

(ii) \( \mathbf{x} \) is averaging of \( \mathbf{y} \) (i.e. \( \mathbf{x} = P \mathbf{y} \) for some doubly stochastic matrix \( P \)).
(iii) \( \sum_{i=1}^{n} h(x_i) \leq \sum_{i=1}^{n} h(y_i) \) for every continuous convex function \( h : \mathbb{R} \to \mathbb{R} \).

(iv) \( \sum_{i=1}^{n} (x_i - c)^+ \leq \sum_{i=1}^{n} (y_i - c)^+ \ \forall c \in \mathbb{R} \) and \( \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \).

We will also use two results concerning the structure of \( n \times n \) doubly stochastic matrices. The set of all \( n \times n \) permutation matrices will be denoted by \( \mathcal{P}^n \), the class of all \( n \times n \) permutation matrices that involve an interchange of just two coordinates will be denoted by \( \mathcal{B}^n \) and the class of \( T \)-transition matrices (or more evocatively as we will see later, the class of Robin Hood matrices) denoted by \( T^n \) consists of all doubly stochastic matrices of the form

\[ \lambda I + (1 - \lambda)B \]  

(3.1)

where \( \lambda \in [0, 1] \) and \( B \in \mathcal{B}^n \).

Finally the class of all \( n \times n \) doubly stochastic matrices will be denoted by \( \mathcal{S}^n \). We have

**Theorem 3.5.**

(a) (Birkhoff, 1946). \( \mathcal{P}^n \) is the set of extreme points of \( \mathcal{S}^n \) and \( \mathcal{S}^n \) is the convex hull of \( \mathcal{P}^n \).

(b) (Hardy, Littlewood and Polya, 1934). \( P \in \mathcal{S}^n \) iff there exists a finite set of matrices in \( T^n \), say \( T_1, T_2, \ldots, T_m \) such that \( P = T_1 T_2 \ldots T_m \), (in fact \( m \) will be \( \leq n - 1 \) here).

Finally we remark that in the light of Theorem 3.5(b) it is possible to verify Schur convexity of a function by considering only the effect of changes in the first two coordinates. Thus a differentiable function \( g : I^n \to \mathbb{R} \), where \( I \) is an interval, is Schur convex if

(i) \( g \) is a symmetric function of \( x_1, x_2, \ldots, x_n \).

(ii) \( (x_1 - x_2)(\frac{\partial}{\partial x_1}g(x) - \frac{\partial}{\partial x_2}g(x)) \geq 0 \ \forall x \in I^n \)  

(3.2)
4 Dalton’s key principle

In his important paper on desirable properties of income inequality measures, Dalton (1920) proposed four principles. The important one for our purposes is the one that says that if we take a small amount from an individual in a population and give it to a relatively poorer individual in the population, the result will be a decrease in inequality. In other words, Robin Hood, in taking from the rich and giving to the poor, is indeed reducing inequality. In honor of the celebrated hero of Sherwood Forest, we will call this the Robin Hood principle (it is also known as the Dalton or the Pigou-Dalton principle) and the operation as a Robin Hood transfer.

Referring back to the discussion in Section 3, it is clear that the effect of a Robin Hood transfer on a population vector \( \mathbf{x} \) is to replace \( \mathbf{x} \) by \( T \mathbf{x} \) where \( T \) is a \( T \)-transform matrix (or Robin Hood matrix).

Moreover in the light of Theorem 3.5(b), we will have \( \mathbf{x} \prec \mathbf{y} \) iff \( \mathbf{x} \) can be obtained from \( \mathbf{y} \) by a finite series of Robin Hood transfers (from rich to poor).

5 The Lorenz order for finite populations

In the context of income and wealth inequality, it is natural to restrict attention to non-negative variables. We may then consider a vector \( \mathbf{x} \in \mathbb{R}^{n+} \) and associate a Lorenz curve with this vector using Lorenz’s definition (with linear interpolation between the points). We may note that for two vectors \( \mathbf{x}, \mathbf{y} \in \mathbb{R}^{n+} \), the Lorenz curve \( L_x(u) \leq L_y(u) \), \( \forall u \in (0,1) \) if and only if the standardized version of \( \mathbf{y} \) is majorized by the standardized version of \( \mathbf{x} \). Standardization in this setting involves dividing each coordinate of \( \mathbf{x} \) by the total \( \sum_{i=1}^{n} x_i \), \( \mathbf{y} \) being standardized in similar fashion so that the standardized vectors have coordinates which sum to one. Thus, many properties of the Lorenz order can be immediately deduced from corresponding results for majorization. There is an
advantage to use of the Lorenz order. It allows us to compare populations of different sizes. In fact it is possible and natural to extend the Lorenz order to compare quite arbitrary distributions, as we shall observe in the next section. From the results described for majorization we can conclude that if we wish to use a scale invariant measure of inequality and if we accept that a Robin Hood operation decreases inequality, we are forced to accept the Lorenz order (based on nested Lorenz curves) as an appropriate inequality partial ordering. The attractiveness of the Robin Hood principle and the fact that one can move from a lower Lorenz curve to a higher Lorenz curve via a finite sequence of Robin Hood operations, undoubtedly explains the almost universal appeal of the Lorenz order as a clear reflection of inequality ordering.

6 The general Lorenz order

Let $\mathcal{L}^+$ denote the class of all non-negative random variables with finite positive expectations. We define a partial order on $\mathcal{L}^+$ by first associating a Lorenz curve with each random variable in $\mathcal{L}^+$.

**Definition 6.1**: (Gastwirth, 1971). Let $X \in \mathcal{L}^+$ with distribution function $F_X$ and quantile function $F_X^{-1}$. The Lorenz curve $L_X$ corresponding to $X$ is defined by

$$L_X(u) = \left[ \int_0^u F_X^{-1}(x)dx \right] / \left[ \int_0^1 F_X^{-1}(x)dx \right], \quad 0 \leq u \leq 1 . \quad (6.1)$$

It is evident from (6.1) that a Lorenz curve will be continuous, non-decreasing convex function that is differentiable almost everywhere in $[0, 1]$. These are properties that we expect from a Lorenz curve following Lorenz’s original description of such a curve.

If $\underline{x} = (x_1, x_2, \ldots, x_n)$ has non-negative coordinates (not all zero), then we can consider two Lorenz curves associated with this vector. First we may use Lorenz’s original definition. Secondly we can consider a random variable $X$ associated with selecting a coordinate of $\underline{x}$ at random with all coordinates equally likely. For such a random variable $X$ we can use Definition 6.1 to determine
its corresponding Lorenz curve. The same curve is arrived at in both manners. Thus (6.1) can be
viewed as a legitimate extension of Lorenz’s original concept, including Lorenz’s bow shaped curve
for finite populations as a special case, that can be interpreted as one corresponding to a random
selection of a unit from the finite population.

Our Lorenz order is then naturally defined in the space $\mathcal{L}^+$ as follows.

Definition 6.2: For $X, Y \in \mathcal{L}^+$, we write $X \leq_L Y$ if and only if $L_X(u) \geq L_Y(u), \forall u \in [0, 1]$.

Note that this ordering actually relates equivalence classes of random variables in $\mathcal{L}^+$ where
two random variables are said to be equivalent if one is a constant multiple of the other.

By considering a non-negative random variable as a limit in distribution of random variables
taking on a finite number of possible values we can develop an analog to Theorem 3.4 (which gave
alternative necessary and sufficient conditions for majorization).

Theorem 6.3: Let $X, Y \in \mathcal{L}^+$ with $E(X) = E(Y)$. The following are equivalent.

(i) $X \leq_L Y$

(ii) $X$ is an averaging of $Y$ in the sense that there exist jointly distributed random variables $X', Z'$
such that $Y \overset{d}{=} Y'$ and $X \overset{d}{=} E(Y'|Z')$.

(iii) $E(h(X)) \leq E(h(Y))$ for every continuous convex function $h : \mathbb{R} \to \mathbb{R}$.

(iv) $E((X - c)^+) \leq E((Y - c)^+), \forall c \in \mathbb{R}^+$.
7 Extremal patterns

Among the set of vectors $\mathbf{x} \in \mathbb{R}^+$ with $\sum_{i=1}^{n} x_i = c > 0$ the extremal cases with respect to majorization are of the form $(c/n, c/n, \ldots, c/n)$ and $(0, 0, 0, \ldots, 0, c)$ corresponding, in income terms, to distributions in which respectively the wealth is evenly distributed or is all in the hands of one individual.

Many summary measures of inequality are defined in terms of the Lorenz curve. Three examples are:

(i) The Gini index, $G$, defined to be two times the area between the Lorenz curve and the egalitarian line (joining $(0,0)$ to $(1,1)$).

(ii) The Pietra index, $P$, defined to be the maximum vertical distance between the Lorenz curve and the egalitarian line.

(iii) The Kakwani index, $K$, defined to be the length of the Lorenz curve.

It is evident that all three measures, $G$, $P$ and $K$, respect the Lorenz order. Moreover, by considering the extremal cases we may immediately identify bounds on these summary measures of inequality, thus

\[ 0 \leq G \leq 1 \]
\[ 0 \leq P \leq 1 \]
\[ \sqrt{2} \leq K \leq 2 \]  \hspace{1cm} (7.1)

and

\[ \sqrt{2} \leq K \leq 2 \]

(in order to obtain a measure whose values range from 0 to 1, Kakwani (1980) actually proposed using $(K - \sqrt{2})/(2 - \sqrt{2})$ as an inequality measure).
But in fact any Schur convex function can be used as a summary measure of inequality by applying it to the vector \((x_1/\sum_{i=1}^{n} x_i, \ldots, x_n/\sum_{i=1}^{n} x_i)\). In particular, separable convex functions of the form \(\sum_{i=1}^{n} g(x_i)\) where \(g\) is convex are often utilized. One of the first of such measures to be used involved the choice \(g(x) = x^2\) and orders the inequality of populations in terms of their coefficients of variation. An alternative choice is \(g(x) = x \log x\), the negative of the entropy function.

However our interest in this section is not in the use of the Lorenz order or its surrogate majorization, on its home grounds (income inequality comparisons), but we wish to highlight its potential role in disparate situations seemingly far removed from the income context. The remarkable diversity of such situations and the growing number of them, assure an evergreen future for Lorenz’s curve and its associated partial order.

As a rule of thumb, if a problem deals with vectors in \(\mathbb{R}^n^+\) and if the extremal solution to that problem is a vector of the form \((c, c, \ldots, c)\), then we should look for the potential role of some Schur convex function and be on the lookout for a Lorenz ordering (or majorization) interpretation of the phenomenon in question. A more extensive collection of such examples may be found in Arnold (2005) (described specifically in terms of majorization), but it is hoped that the small selection presented below will serve to indicate some of the possibilities and will encourage readers to be on the lookout for more places for Max Lorenz’s ordering to play a fruitful role.

8 Some examples

8.1 Time to absorption in a continuous time Markov chain

Consider a continuous time Markov chain with \((n + 1)\) states, of which \(n\) states \((1, 2, \ldots, n)\) are transient and one state, \(n + 1\), is absorbing. Consider the time \(T\) until absorption in state \((n + 1)\). The family of distributions of such absorption times is indexed by \(\alpha = (\alpha_1, \ldots, \alpha_n)\) the initial
probability distribution vector, and the matrix $Q$ of intensities of transitions among the $n$ transient states. Following Neuts (1975) we will say that $T$ has a phase-type distribution with parameters $\alpha$ and $Q$ and we write $T \sim PH(\alpha, Q)$.

We will say that $T$ has a phase type distribution of order $n$ if $n$ is the smallest integer such that the distribution can be identified with the time to absorption in a chain with $n$ transient and 1 absorbing state. Perhaps the simplest example of a Phase type distribution of order $n$ is provided by a gamma distribution with shape parameter $n$ and intensity parameters $\lambda$, i.e. a sum of $n$ i.i.d. exponential ($\lambda$) random variables. It corresponds to a chain which begins in state 1 (i.e. we have $\alpha = (1, 0, \ldots, 0)$) and progresses sequentially through the states 2, 3, \ldots, $n + 1$ spending an independent exponential ($\lambda$) time in each transient state. Denote such a simple Phase type random variable by $T^*$.

It is evident that Phase type distribution of order $n$ with the same mean can exhibit considerable differences in their variability. For example if in our simple example instead of spending and $\exp(\lambda)$ time in each state the process spends an $\exp(\lambda_i)$ time in state $i$, $i = 1, 2, \ldots, n$, it is well known that the choice $\lambda_i = \lambda$, $i = 1, 2, \ldots, n$ results in minimal variability in the sense of Lorenz ordering. In fact, O’Cinneide (1991) has verified that among all phase type distributions of order $n$, the simple gamma random variable $T^* \sim \text{gamma}(n, \lambda)$ exhibits least variability as measured by the Lorenz order.

8.2 Connected components of a random graph

Following Ross (1981) consider a random graph with nodes 1, 2, \ldots, $n$. We construct a random graph by drawing $n$ arcs, each one emanating from one of the nodes. The arc beginning at node $i$ ends at node $X(i)$ where the $X(i)$’s are independent identically distributed random variables with

$$P(X(i) = j) = p_j, \quad j = 1, 2, \ldots, n$$  \hspace{1cm} (8.1)
where $p_j \geq 0$ and $\sum_{j=1}^{n} p_j = 1$. Note that an arc might link a node to itself and note that several arcs might end at the same node (though they begin from separate nodes). The number of connected components of such a graph will be denoted by $M$. The possible values of $M$ are clearly $1, 2, \ldots, n$ and the distribution of $M$ will be governed by $\underline{p}$, the vector of probabilities appearing in (8.1). An extreme case, would involve $\underline{p} = (1, 0, 0, \ldots, 0)$. In such a case all arcs will terminate at node 1 and we will have all nodes connected and $M = 1$ with probability 1. Thus $M$ is stochastically smallest when all the probability in $\underline{p}$ is concentrated at one point (in income terms, when one individual has all the money). We might conjecture that for each $m \in \{1, 2, \ldots, n\}$ the function $P(M \geq m)$ will be a Schur concave function of $\underline{p}$. The exact distribution of $M$ is not easy to evaluate. However, Ross (1981) is able to provide an attractively simple expression for the expected values of $M$:

$$E(M) = \sum_{S} (|S| - 1)! \prod_{j \in S} p_j$$

where the summation is taken over every non-empty subset $S$ of $\{1, 2, \ldots, n\}$. Using (8.2) he is able to verify that $E(M)$ is a Schur concave function of $\underline{p}$. Thus to maximize the expected number of connected components of the graph we should set $p_j = 1/n, \ j = 1, 2, \ldots, n$.

Note that in this example it was not a priori obvious that $(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n})$ was an extremal choice for $\underline{p}$ but it was obvious that $(1, 0, \ldots, 0)$ was extremal and this motivated a search for a Schur convex or concave function.

8.3 Catchability

A closed ecological community such as an isolated island will contain an unknown number $\nu$ of species of butterflies (for example). Butterflies are hunted until $n$ individuals have been trapped. Let $R$ denote the random number of different species represented among the $n$ trapped butterflies. It is common practice to use the observed value of $R$ to estimate the unknown parameter $\nu$. We may write

$$R = \sum_{i=1}^{\nu} I(X_i > 0)$$
where \( \mathbf{X} = (X_1, \ldots, X_\nu) \) is random vector with a multinomial distribution, i.e. \( \mathbf{X} \sim \text{multinomial}(n, \mathbf{p}) \). Here \( \mathbf{p} = (p_1, p_2, \ldots, p_\nu) \) in which \( p_i \) denotes the probability that a particular trapped butterfly will be from species \( i, \; i = 1, 2, \ldots, \nu \). Intuition suggests strongly that the distribution of \( R \) will depend on the variability in \( \mathbf{p} \) (or on the inequality between the components of \( \mathbf{p} \)). \( p_i \) is often described as the catchability of species \( i \). Under an equal catchability assumption (i.e. \( p_i = 1/\nu, \; i = 1, 2, \ldots, \nu \)), we may expect many species to be represented among the \( n \) trapped butterflies, i.e. \( R \) will be large. At another extreme, if one species is “trap happy” then we find most, if not all, of our trapped butterflies to be of that species, i.e. \( R \) will be small. Nayak and Christman (1992) show that for any \( r = 1, 2, \ldots, n \) the quantity

\[
P_p(R \leq r)
\]

is indeed a Schur convex function of \( \mathbf{p} \). From this it follows that the expected number of species represented in the sample, \( E_p(R) \), is a Schur concave function of \( \mathbf{p} \). Consequently most of the estimates of \( \nu \) that are quite sensible assuming equal catchability will be negatively biased with the bias increasing as the variability in catchability increases.

It is possible to reinterpret these computations in the context of a simple random sample taken with replacement from a population broken into \( \nu \) subpopulations of various sizes. Then \( R \) will represent the number of subpopulations represented in the sample. The model can also be invoked in a numismatic context where ancient coins come from an unknown number of minting locations.

9 To higher dimensions

It is natural to seek versions of inequality measures and orderings that can be used for multivariate income or wealth distributions. Thus we may wish to model incomes of related individuals, incomes of individuals at different time points or incomes of one individual in various currencies. Given the attractiveness of Lorenz’s partial order in the univariate setting, we are led to seek an appropriate \( k \) dimensional extension of Lorenz’s curve. Some rather unsuccessful attempts at such a generalization
can be found in the literature (e.g. Taguchi (1972), Arnold (1983)). The perceived difficulty was due perhaps to an over emphasis on the Gastwirth definition of the Lorenz curve, (6.1), in terms of the quantile function $F_X^{-1}$. Such a definition involves ordering observations and it appeared that some multivariate version of either order statistics or the quantile function would be required in order to take Lorenz’s curve successfully to higher dimensions. Instead, the key lay in redefining the Lorenz curve in a manner that did not involve ordering the data.

As Lorenz did in the one dimensional case, we begin by considering a finite population of $n$ individuals. However now we assume that associated with the $j$’th individual is an $m$-dimensional wealth vector $\mathbf{x}(j) = (x_{j1}, \ldots, x_{jm})$ where, for example, $x_{j\ell}$ represents the wealth of individual $j$ in currency $\ell$ ($\ell = 1, 2, \ldots, m$). Now rather than consider only subpopulations consisting of relatively poor individuals, we will associate a point in $\mathbb{R}^{m+1}$ with each subpopulation of the full set of $n$ individuals in the population. Associated with a particular subpopulation $G$ including say $t$ individuals we associate the vector

$$
\left( \frac{t}{n}, \frac{\sum_{j \in G} x_{j1}}{\sum_{j=1}^{n} x_{j1}}, \ldots, \frac{\sum_{j \in G} x_{jm}}{\sum_{j=1}^{n} x_{jm}} \right).
$$

(9.1)

The first coordinate of this vector represents the proportion of the total populations included in the group $G$, while coordinates $2, \ldots, m + 1$ represent the proportion of the total wealth of the population held by individuals in group $G$ for each of the $m$ currencies under consideration.

The convex hull of the points of the form (9.1) constitutes the Lorenz zonoid of the population, introduced by Koshevoy (1995). Comparison of inequality between populations is then based on whether their corresponding Lorenz zonoids are nested. In a subsequent series of papers Koshevoy and Mosler provided a natural extension of this concept to deal with quite general $m$-dimensional random variables (not just those taking on a finite number of values). Inter alia they provided links to other variability orderings and in some cases natural extensions of various univariate characterizations of the Lorenz order, such as those listed in Theorem 6.3 above. An excellent survey which includes discussion of this body of research may be found in Mosler (2002).
The Lorenz zonoid order appears to be the compelling choice as a suitable \( m \)-dimensional analog of Lorenz’s order. Though it has a few perhaps unforeseen surprises for us. Intuitively the volume of the zonoid does provide an alternative summary measure of inequality (it actually coincides with the classical Gini index when wealth is one dimensional). Moreover a zero value for a measure of inequality is expected to be associated with an egalitarian distribution. When we deal with \( m \) currencies, the zonoid can have zero volume for some non-degenerate distributions. Mosler (2002) provides a variant definition to rectify this anomaly.

A second surprise is waiting for us if we consider the possibility of changing all the wealth of the individuals into one selected currency. An \( m \)-dimensional wealth random variable \( \mathbf{X} \) will then have associated with it a one dimensional random variable \( c' \mathbf{X} \) where the \( c_i \)'s are non-negative quantities representing exchange rates. One idea for defining inequality among non-negative integrable \( m \)-dimensional random variables is based on such a currency exchange scenario. We will say that \( \mathbf{X} \) is exchange rate Lorenz ordered with respect to \( \mathbf{Y} \) if \( c' \mathbf{X} \leq_L c' \mathbf{Y} \) in the usual univariate Lorenz order sense for every \( c \in \mathbb{R}^+ \).

Some candidate definitions for a Lorenz order among \( m \)-dimensional non-negative random vectors (with positive finite marginal expectations) are listed next.

(i) \( \mathbf{X} \leq_L \mathbf{Y} \) if \( L(X) \subseteq L(Y) \) (where \( L(X) \) denotes the Lorenz zonoid of \( X \)).

(ii) \( \mathbf{X} \leq_L \mathbf{Y} \) if \( E(g(\frac{X_1}{E(X_1)}, \ldots, \frac{X_m}{E(X_m)})) \leq E(g(\frac{Y_1}{E(Y_1)}, \ldots, \frac{Y_m}{E(Y_m)})) \) for every continuous convex function \( g : \mathbb{R}^+ \to \mathbb{R} \) whose expectations exist.

(iii) \( \mathbf{X} \leq_{L_2} \mathbf{Y} \) if \( a' \mathbf{X} \leq_L a' \mathbf{Y} \), \( \forall a \in \mathbb{R}^m \)

(iv) \( \mathbf{X} \leq_{L_3} \mathbf{Y} \) if \( c' \mathbf{X} \leq c' \mathbf{Y} \), \( \forall c \in \mathbb{R}^{+m} \).

(v) \( \mathbf{X} \leq_{L_4} \mathbf{Y} \) if \( X_i \leq_L Y_i \), \( i = 1, 2, \ldots, m \).

Evidently the last three of these orderings are related in the sense that (iii) \( \Rightarrow \) (iv) \( \Rightarrow \) (v).
Definition (v) could be labelled marginal Lorenz ordering. Our Lorenz zonoid order, (i), is actually equivalent to ordering (iii). Thus zonoid ordering implies but is not implied by exchange rate Lorenz ordering (or positive combinations Lorenz ordering in the terminology of Joe and Verducci(1992)). Adding to its list of aliases, Mosler (2002) calls it the price Lorenz order. It would appear that in order to give the zonoid ordering an interpretation in terms of exchange rates we must consider some negative(!) exchange rates.

10 An evergreen future

On its one hundredth birthday, Lorenz’s curve continues to generate interest and to be usefully applied in an amazingly broad spectrum of research areas. Open questions still abound, especially in the $m$-dimensional setting. And, of course, since in many instances Lorenz curves are not nested but cross once, twice or many times, considerable scope exists for consideration of Lorenz ordering over subsets of populations (e.g. studies of the poor, or the rich). The insights in that brief paper in 1905, coupled with inputs from mathematical discussions of majorization and its more abstract versions, continue to enable us to usefully grasp inequality, dispersion and variability concepts here, there and everywhere.

REFERENCES


