Lorenz curves and partial orders

T. Kämpke, F.J. Radermacher

Forschungsinstitut für Anwendungsorientierte Wissensverarbeitung/n FAW/n

Lise-Meitner-Str. 9, 89081 Ulm, Germany

{kaempke,radermacher}@faw-neu-ulm.de

Abstract

We compile classical relations between the Lorenz order and majorization for finite discrete distributions and between the Lorenz order and the convex order for all other distributions. Simple transfers of the Pigou-Dalton type are extended as far as possible including approximations of continuous distributions.

Key words: convex stochastic order, Atkinson theorem, majorization, Lorenz order.

1 Introduction

Out of the wide body of literature on partial order relations over Lorenz curves only those are treated here which are close to redistribution of income, wealth and possibly other goods which come in non-negative magnitudes. We follow the classical extension of majorization for vectors to the convex stochastic order for distributions. Thus, eventually, the Atkinson theorem will be extended to other than finite distributions; the Lorenz order will be seen to be equivalent to the convex or concave stochastic order for all distributions with identical, finite means.

When confusions with real numbers are avoided, vectors will be denoted by x, y etc. and they are understood as column vectors. Matrices are denoted by A, B etc. and with transposes of vectors and matrices are denoted by x^T, A^T etc.

2 Partial orders for Lorenz curves

The pivotal order for Lorenz curves is that for Lorenz curves without intersection point.

Definition 1 *Random variable X with Lorenz curves L_X is defined to be smaller in the Lorenz order than random variable Y with Lorenz curve L_Y if and only if $L_X(u) \geq L_Y(u)$ for all $u \in [0,1]$. Notation $X \leq_L Y$.*

The idea of the Lorenz order is unevenness since the one of two Lorenz curves, if any, which is consistently closer to the diagonal (representing the egalitarian distribution) is the smaller. Obviously, two Lorenz curves may be incomparable in the Lorenz order and, even worse, distributions need not have Lorenz curves. The Lorenz order is identical to the stochastic order for Lorenz curves considered as distribution functions over the unit interval.

The Lorenz order may refer to Lorenz curves directly which means no explicit reference is made to random variables or distributions. Two Lorenz curves are ordered as $L_1 \leq_L L_2$ if and only if $L_1(u) \geq L_2(u)$ for all $u \in [0,1]$. $L_1 \leq_L L_2$ implies the same order for the Gini indices $G_1 \leq G_2$. The Gini index of a Lorenz curve equals twice the area between the Lorenz curve and the diagonal $G = 2 \cdot \int_0^1 u - L(u)\, du$.

When $L_1 \leq_L L_2$ and when the Gini indices of the two Lorenz curves are identical, then $L_1(u) = L_2(u)$ for all $u \in [0,1]$. This follows from the continuity of Lorenz curves or by applying a general result for distribution functions [MS, theorem 1.2.9, p. 5] to the special case of Lorenz curves.

Lorenz curve that violate the Lorenz order may have an arbitrarily large number of intersection points. This can be seen from the following construction. Consider an auxiliary polygon with $k \geq 4$ nodes which itself is a Lorenz curve. The slopes of any two linear segments are taken to be different. Then one Lorenz curve is constructed from joining the by straight lines the first and third node, the third and fifth node etc. and the other Lorenz curve is constructed from straight line segments from second to fourth node, from fourth to sixth node etc. At the boundaries, direct neighboring nodes instead of second-neighbor nodes may have to be connected, comp. figure 1. The resulting two Lorenz curves have $k-3$ intersection points. The same result can be obtained from a strictly convex Lorenz curve serving as auxiliary function. All nodes are then drawn from this curve.

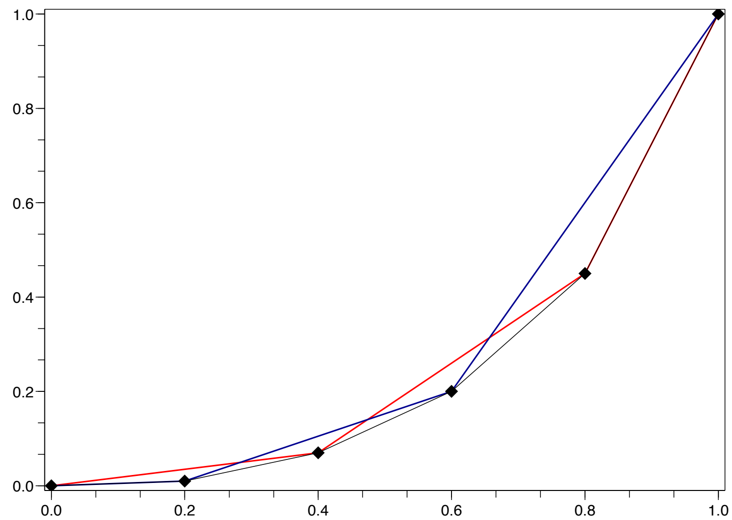

Figure 1: Two piecewise linear Lorenz curves with $k = 6$ polygonal nodes and three intersection points.

2.1 Lorenz order and majorization

A concept for unevenness of vectors from an n-dimensional Euclidean space is majorization. Any vector from $\mathbb{R}^n$, $n \geq 2$, can be considered as a finite income sample with all equal probabilities $1/n$. These probabilities are ignored. When the sum over all incomes of two vectors is identical, unevenness can be formulated coordinate-wise beginning with the largest entries.

Definition 2 *Vector y majorizes vector x when they have the same number of coordinates and when the decreasingly sorted coordinates $x_{[1]} \geq \ldots \geq x_{[n]}$ and $y_{[1]} \geq \ldots \geq y_{[n]}$ satisfy the partial sum conditions*

$$
\begin{aligned}
&y_{[1]} \geq x_{[1]} \\
&y_{[1]} + y_{[2]} \geq x_{[1]} + x_{[2]} \\
&\vdots \\
&y_{[1]} + \ldots + y_{[n-1]} \geq x_{[1]} + \ldots + x_{[n-1]} \\
&y_{[1]} + \ldots + y_{[n]} = x_{[1]} + \ldots + x_{[n]}.
\end{aligned}
$$

Notation $x \leq_m y$.

An example is $x = (3, 3.5, 2.5)^T \leq_m (2, 3, 4)^T = y$. Majorization is verified by $4 \geq 3.5$, $4+3 \geq 3.5+3$ and $4+3+2 = 3.5+3+2.5$. The vector of all equal entries $(c, c, \ldots, c)$ is majorized by every other vector whose coordinates sum to $n \cdot c$.

A relaxed notion is weak majorization which amounts to only requiring in the last partial sum condition that $y_{[1]} + \ldots + y_{[n]} \geq x_{[1]} + \ldots + x_{[n]}$. In addition, the concept of majorization applies to vectors with some or all coordinates being negative. None of these variations is opted for. The reason for the latter is that incomes are assumed to be non-negative.

Though majorization is reflexive ($x \leq_m x$ for all x) and transitive ($x \leq_m y$ and $y \leq_m z$ implies $x \leq_m z$), it is not a partial order since it is not antisymmetric; it is possible that $x \leq_m y$ and $y \leq_m x$ but still $x \neq y$. Thus, majorization is not a partial order but a pre-(partial) oder. Vectors that majorize each other are related to permutations.

More precisely, all vectors that are majorized by a particular vector form a convex set. This set is the convex hull spanned by the vectors that result from all coordinate permutations of the given vector. The geometry of majorization is quite simple in two dimensions as sketched in figure 2 and slightly more complex in three dimensions as sketched in figure 3.

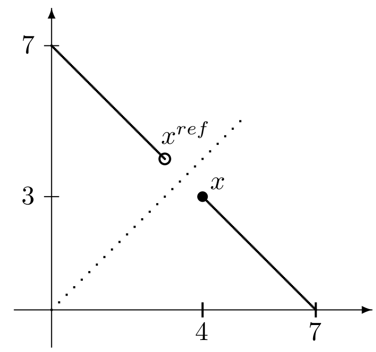

Figure 2: The points which have non-negative coordinates and which majorize $x = (4, 3)^T$ lie on the two line segments having slope -1 with $x^{ref} = (3, 4)^T$. Majorization is thus seen to pull away from the diagonal.

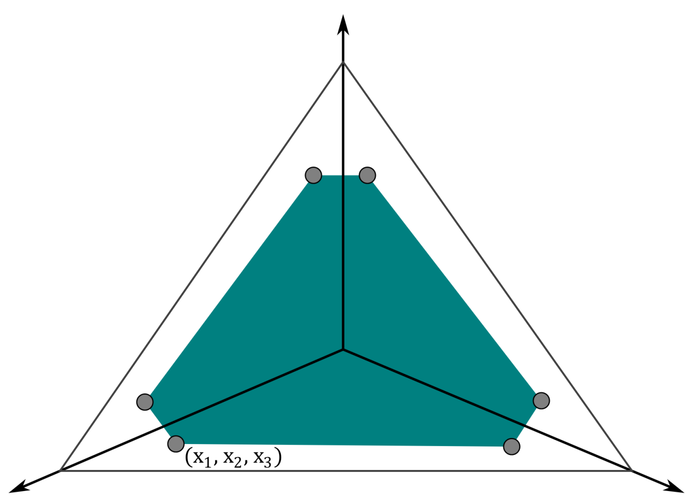

Figure 3: The vector $x = (4, 1, 0.5)^T$ leads to six vectors that lie in the plane $H = \{(x_1, x_2, x_3)^T | x_1 + x_2 + x_3 = 5.5\}$ when its coordinates are permuted in all possible ways (dark dots). The set majorized by x is the convex hull spanned by the six vectors (shaded area) and the set of vectors that majorize x is the complement within the plane (white). All vectors outside the plane and incomparable to x with respect to majorization.

Majorization can be obtained from redistribution. A given vector x with decreasingly sorted coordinates is compared to a vector x' that differs only in two coordinates: $x'_i = x_i - \varepsilon$ and $x'_j = x_j + \varepsilon$ with $i < j$ and $\varepsilon > 0$. The given vector majorizes the new vector, which means $x' \leq_m x$, if the redistribution amount ε is small enough to preserve sorting. Vector x' is said to be obtained by a simple redistribution step or a Pigou-Dalton transfer from x. Since the decreased coordinate is larger than the increased ($x_i > x_j$), this transfer step can be considered as a redistribution "from rich to poor".

Lemma 1 *("Equivalence of majorization and successive Pigou-Daltor transfers")*
For $x \leq_m y$ there is a finite sequence $x = x^n \leq_m \ldots \leq_m x^2 \leq_m x^1 = y$ such that x^{k+1} results from a Pigou-Dalton transfer from $x^k + 1$, $k = n-1, \ldots, 1$.

Proof. Let both vectors have decreasingly sorted coordinates. The interim sequence is then constructed by successive Pigou-Dalton transfers of adjacent coordinates beginning with the first transfer amount $\varepsilon = y_1 - x_1$ as indicated

$$x^2 = \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \\ \vdots \\ x_n^2 \end{pmatrix} = \begin{pmatrix} y_1 - \varepsilon \\ y_2 + \varepsilon \\ y_3 \\ \vdots \\ y_n \end{pmatrix} \leq_m \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix} = y.$$

The next transfer is from the second to the third coordinate so that the new second coordinate equals x_2. These steps are continued until the new vector equals x. $\diamond$

Majorization enables a certain kind of function monotonicity. A real–valued function F defined on vectors is understood to be Schur-convex if $x \leq_m y$ implies $F(x) \leq F(y)$ and it is Schur-concave if $x \leq_m y$ implies $F(x) \geq F(y)$. A special case of Schur-convexity is obtained from ordinary convexity by the next result.

Theorem 1 *$x \leq_m y$ if and only if $x_1 + \ldots + x_n = y_1 + \ldots + y_n$ and $f(x_1) + \ldots + f(x_n) \leq f(y_1) + \ldots + f(y_n)$ for any real-valued function f which is convex and contnuous.*

Proof. See [MS, corollary 1.5.7, p. 34]. $\diamond$

When the function f is concave, the theorem implies that $x \leq_m y$ entails $f(x_1) + \ldots + f(x_n) \geq f(y_1) + \ldots + f(y_n)$. A characterization of majorization by multiplication with a doubly stochastic matrix is given, for example, in [BB, p. 30]. A square matrix is doubly stochastic if all its entries are non-negative and all row sums as well as all column sums are equal to one.

Theorem 2 *$x \leq_m y$ is equivalent to the existence of a doubly stochastic matrix A with $x = Ay$.*

The sample vectors $x = (3, 3.5, 2.5)^T \leq_m (2, 3, 4)^T = y$ allow the mapping of the larger vector to the smaller vector as

$$\begin{pmatrix} 3 \\ 3.5 \\ 2.5 \end{pmatrix} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \\ 2/3 & 1/6 & 1/6 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}.$$

The doubly stochastic matrix which maps the larger vector to the smaller need not be unique. In the foregoing case, even a symmetric doubly stochastic matrix applies

$$\begin{pmatrix} 3 \\ 3.5 \\ 2.5 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}.$$

A symmetric matrix between majorized vectors may be difficult to obtain and it is a non-trivial issue to decide if this is possible. Transitivity of majorization does not help for the symmetry issue though transitivity can be expressed by $x = Ay$ and $y = Bz$ for doubly stochastic matrices A, B implying $x = ABz$. Note that the product of two doubly stochastic matrices is again doubly stochastic. But symmetry is not guaranteed since the product of two symmetric matrices, even if doubly stochastic, need not be symmetric. An example of two symmetric, doubly stochastic matrices with asymmetric product is

$$\begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{pmatrix} = \begin{pmatrix} * & 1/2 & * \\ 1/4 & * & * \\ * & * & * \end{pmatrix}.$$

A straightforward consequence from the definition of majorization is that it entails the Lorenz order. The Lorenz order therefore refers to finite distributions which attain the values $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$, respectively. All probabilities are equal so that $P(X = x_i) = P(Y = y_i) = 1/n$ for all i when all events are pairwise different. When multiple events x_i are equal, probabilities add up according to event multiplicity. The same applies to multiple events y_i being equal. These distributions associated with vectors will henceforth be denoted as natural distributions.

Lemma 2 *$x \leq_m y$ if and only if $X \leq_L Y$.*

Proof. "$\Longrightarrow$". Let the coordinates of x and y be sorted decreasingly. Then $x \leq_m y$ implies

$$\begin{aligned} x_n &\geq y_n \\ x_n + x_{n-1} &\geq y_n + y_{n-1} \\ x_n + \ldots + x_2 &\geq y_n + \ldots + y_2. \end{aligned}$$

This can be seen from $x_n + \ldots + x_{n-k} = x_1 + \ldots + x_n - (x_1 + \ldots + x_{n-k-1}) \geq y_1 + \ldots + y_n - (y_1 + \ldots + y_{n-k-1}) = y_n + \ldots + y_{n-k}$ for all $k = 0, \ldots, n-2$. The generalized inverses of the two distribution functions are

$$F_X^{-1}(u) = \begin{cases} x_n, & \text{if } \ 0 \leq u \leq 1/n \\ x_{n-1}, & \text{if } \ 1/n < u \leq 2/n \\ \vdots & \vdots \\ x_1, & \text{if } \ (n-1)/n < u \leq 1. \end{cases} \quad \text{and } F_Y^{-1}(u) = \begin{cases} y_n, & \text{if } \ 0 \leq u \leq 1/n \\ y_{n-1}, & \text{if } \ 1/n < u \leq 2/n \\ \vdots & \vdots \\ y_1, & \text{if } \ (n-1)/n < u \leq 1. \end{cases}$$

Thus, the generalized inverses have identical jump points $0, \frac{1}{n}, \frac{2}{n}, \ldots, 1$. There, the Lorenz curve of x has the values $0, \frac{x_n}{x_1+\ldots+x_n}, \frac{x_n+x_{n-1}}{x_1+\ldots+x_n}, \ldots, 1$ and the Lorenz curve of y has the values $0, \frac{y_n}{y_1+\ldots+y_n}, \frac{y_n+y_{n-1}}{y_1+\ldots+y_n}, \ldots, 1$. This means that the intended inequality is valid at all jump points of the generalized inverses: $L_X(\frac{k}{n}) \geq L_Y(\frac{k}{n})$ for $k = 0, \ldots, n$. Since both Lorenz curves are linear between the jump points, the desired inequality $L_X(u) \geq L_Y(u)$ follows for all $u \in [0, 1]$.

'$\Longleftarrow$". The foregoing steps are reversible. $\diamond$

Next, majorization will be extended from vectors to distributions. The idea of Schur-convexity is used to yield order relations that are defined by integrals over certain classes of functions.

2.2 Lorenz order and integral orders

When a vector x is majorized by a vector y and when a function f is convex then the inequality from theorem 1 can be written as $\frac{1}{n} f(x_1) + \ldots + \frac{1}{n} f(x_n) \leq \frac{1}{n} f(y_1) + \ldots + \frac{1}{n} f(y_n)$. Using the probabilistic interpretation of vectors from the end of section 2.1 allows to rewrite the inequality in the more abstract form $Ef(X) \leq Ef(Y)$. This inequality can be valid for general probability distributions.

Definition 3 *Random variable X is defined to be convex (stochastically) smaller than random variable Y if and only if $Ef(X) \leq Ef(Y)$ for all convex and continuous functions f. Notation $X \leq_{cx} Y$.*

The integration functions f are neither required to be increasing nor decreasing and the requirement of being continuous only matters for points that do not lie in the domain interior of an integration function. The reason is that every convex function is continuous in its interior. A concave (stochastic) order $\leq_{cv}$ which were defined in the same way as the convex order except that integration functions were required to be concave would essentially result in the same relation. Only an order reversal must be accepted:

$$X \leq_{cx} Y \Longleftrightarrow Ef(X) \leq Ef(Y) \Longleftrightarrow E - f(X) \geq E - f(Y) \Longleftrightarrow Y \leq_{cv} X.$$

The equivalences are true for all convex functions f and for all concave functions $-f$; every concave function g can be written as $g = -f$ with convex function f. Thus $X \leq_{cx}$ implies $EX = EY$. More specialized order relations result from increasing integration functions. Since replacing a function f by $-f$ will alter the monotonicity direction, the following order relations are not reversals of each other.

Definition 4 *Random variable X is defined to be*

1. *increasingly convex (stochastically) smaller than random variable Y if and only if $Ef(X) \leq Ef(Y)$ for all increasing, convex and continuous functions f. Notation $X \leq_{icx} Y$.*
2. *increasingly concave (stochastically) smaller than random variable Y if and only if $Ef(X) \leq Ef(Y)$ for all increasing, concave and continuous functions f. Notation $X \leq_{icv} Y$.*

Both increasing order relations are related by $X \leq_{icv} Y \Longleftrightarrow -Y \leq_{icx} -X$ and the increasing convex order allows to be "tested" by the special convex functions

$$f_a(x) = (x-a)_+ = \begin{cases} 0, & \text{if } x \leq a \\ x - a, & \text{if } x \geq a. \end{cases}$$

These functions are sketched in figure 4 and their importance for the increasing convex order is given in the next result whose proof is based on the dominated convergence theorem, see [MS, theorem 1.5.7, p. 18]. The expectations used in this result are illustrated in figure 5.

Lemma 3 *$X \leq_{icx} Y$ f and only if $Ef_a(X) \leq Ef_a(Y)$ for all $a \in \mathbb{R}$.*

Interestingly, equal expectations entail that the increasing convex order is identical to the convex order.

Lemma 4 *When $EX = EY$ then $X \leq_{cx} Y$ is equivalent to $X \leq_{icx} Y$.*

Proof. [Mö] or [S, p. 9]. ◇

Equality of expectations has an even deeper consequence, namely the equivalence of the Lorenz order and the convex order [MOA, theorem C.8, p. 719]:

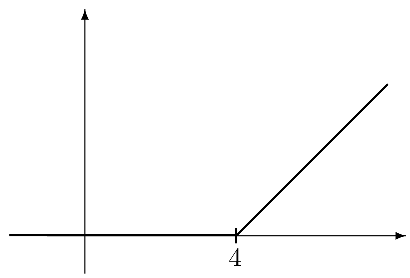

Figure 4: Function $f_4(x)$.

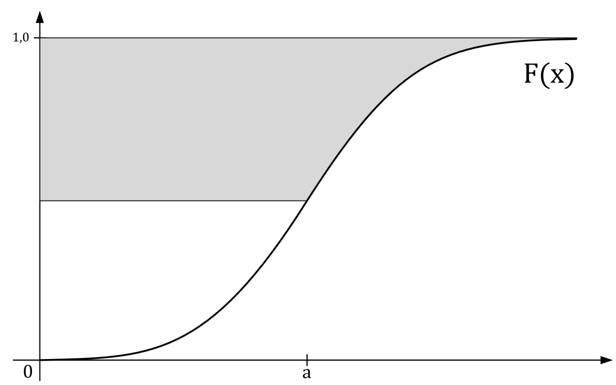

Figure 5: The shaded area equals $Ef_a(X)$ for $F(x) = P(X \leq x)$.

Theorem 3 *("Equivalence of Lorenz order and convex stochastic order for equal means")*
Let X, Y be random variables with Lorenz curves and $EX = EY$. Then $X \leq_L Y$ if and only if $X \leq_{cx} Y$.

A consequence from the last theorem is that $EX = EY$ and the Lorenz order $X \leq_L Y$ imply the variance inequality $VarX \leq VarY$. This follows from applying the convex order to the convex function $f(x) = (x - EX)^2$.

Anoter consequence from threorem 3 is that the single crossing condition by Karlin and Novikoff together with identical mean values is a simple and sufficient condition for Lorenz order. The single crossing condition for two distribution functions $F_X(x)$ and $F_Y(x)$ states that there is at least one point $x_0 \in \mathbb{R}$ such that $F_X(x) \leq F_Y(x)$ for all $x < x_0$ and $F_X(x) \geq F_Y(x)$ for all $x \geq x_0$. The single crossing condition together with $EX = EY$ implies $X \leq_{icx} Y$ [MS]; $EX \leq EY$ together with the single crossing condition suffices. Equality of the expectations further ensures $X \leq_{cx} Y$ and $X \leq_L Y$.

The single crossing condition is not necessary for the increasing convex order, not even for identical expectations and not even for natural distributions. An example is $x = (5, 4, 3, 2)^T$ and $y = (4.5, 4.5, 2.5, 2.5)^T$. Then $y \leq_m x$ so that the natural distributions satisfy $F_Y \leq_{icx} F_X$; see theorem 1 and lemma 4. But the distribution functions have more than one crossing as shown in figure 6.

2.3 An extension of ordinary Pigou-Dalton transfers to more complex distributions

As majorization corresponds to a finite sequence of Pigou-Dalton transfers, see lemma 1, a possibly infinite process of Pigou-Dalton transfers will now be shown to lead from "any" probability distribution with finite expectation to "any" other which is smaller in the sense of increasing convex order and which has the same expectation.

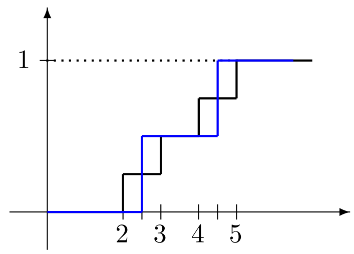

Figure 6: Natural distribution functions of x (four jumps) and of y (two jumps) without single crossing property.

When vectors are cleverly constructed, they allow to represent all finite distributions with rational probabilities as natural distributions. To see this, let a distribution have the probabilities $P(X = x_i) = p_i = r_i/s_i$ with $r_i, s_i \in I\!N$. With the least common multiple of the denominators $S = lcm(s_1, \ldots, s_n)$, the probabilities can be rewritten as $p_i = R_i/S$ with $R_i \in I\!N$. Note that $R_1 + \ldots + R_n = S$. This allows to consider the "natural" vector $x(X)$ with S coordinates having R_1 repetitions of x_1, R_2 repetitions of x_2 etc.

An example is $P(X = 17) = 1/3$, $P(X = 19) = 1/6$, $P(X = 22) = 3/8$, $P(X = 25) = 1/8$. Then $S = lcm(3, 6, 8) = 24$ and $P(X = 17) = 8/24$, $P(X = 19) = 4/24$, $P(X = 22) = 9/24$, $P(X = 25) = 3/24$. The natural vector of the distribution has 24 coordinates. Each can be thought of having probability 1/24.

$$x(X) = (\underbrace{17, \ldots, 17}_{8\,times}, 19, 19, 19, 19, \underbrace{22, \ldots, 22}_{9\,times}, 25, 25, 25)^T.$$

Instead of choosing S to be the least common multiple of all denominators, a multiple of S can be chosen. This will increase the length of the vector and the number of coordinate repetitions. Using larger vector lengths becomes important when two distributions are to be compared. The distributions may have a different number of support points. Let $P(X = x_i) = r_i/s_i$ with $r_i, s_i \in I\!N$ for $i = 1, \ldots, n$ and $P(Y = y_j) = u_j/v_j$ with $u_j, v_j \in I\!N$ for $j = 1, \ldots, m$. Then the length of the common natural vectors of the two distributions can be chosen as $S_+ = lcm(s_1, \ldots, s_n, v_1, \ldots, v_m)$. The probabilities are $r_i/s_i = R_i/S_+$ and $u_j/v_j = U_j/S_+$. The common natural vectors are denoted as $x(X, Y)$ and $y(X, Y)$.

An example is $P(X = x_1) = 2/10$, $P(X = x_2) = 4/10$, $P(X = x_3) = 3/10$, $P(X = x_4) = 1/10$ and $P(Y = y_1) = 1/3$, $P(Y = y_2) = 1/3$, $P(Y = y_3) = 1/3$. Then $S_+ = lcm(10, 3) = 30$ so that common natural vectors can be chosen as

$$\begin{aligned} x(X,Y) &= (\underbrace{x_1, \ldots, x_1}_{6\,times}, \underbrace{x_2, \ldots, x_2}_{12\,times}, \underbrace{x_3, \ldots, x_3}_{9\,times}, \underbrace{x_4, x_4, x_4}_{3\,times})^T \\ y(X,Y) &= (\underbrace{y_1, \ldots, y_1}_{10\,times}, \underbrace{y_2, \ldots, y_2}_{10\,times}, \underbrace{y_3, \ldots, y_3}_{10\,times})^T. \end{aligned}$$

Lemma 5 *Let $EX = EY$ and let the two distributions have rational probabilities $P(X = x_i) = p_i > 0$, $i = 1, \ldots, n$ and $P(Y = y_i) = q_i > 0$, $j = 1, \ldots, m$. Then $P^X \leq_{icx} P^Y$ if and only if $x(X, Y) \leq_m y(X, Y)$.*

Proof. The distributions are the natural distributions of $x(X, Y)$ and $y(X, Y)$, respectively. Thus $x(X, Y) \leq_m y(X, Y)$ is equivalent to $P^X \leq_{cx} P^Y$ according to theorem 1 and $P^X \leq_{cx} P^Y$ is equivalent to $P^X \leq_{icx} P^Y$ according to lemma 4. $\diamond$

A consequence of the foregoing lemma is that a finite sequence of Pigou-Dalton transfers leads from the larger to the smaller distribution, comp. lemma 1. All these transfers refer to natural vectors with same number of coordinates. Extension to more general distributions involve approximations. A core result towards these approximations allows to handle irrational probabilities.

Lemma 6 *("Replacement with rational probabilities")*
For any finite distribution with arbitrary probabilities there exist two finite distributions which (1) have the same expected value as the original distribution, (2) have rational probabilities only and are (3a) larger resp. (3b) smaller than the original distribution in increasing convex order.

Both finite distributions with rational probabilities have only one support point more than the original distribution. The construction is tedious but elementary and transfers all irrationalities from the probabilities to the additional support point. A slight modification of the construction allows all but one support point to become rational. The latter, however, will not be needed here.

Proof of lemma 6. Let the given probabilities be $P(X = x_i) = p_i$ for sorted support points $x_1 < \ldots < x_n$. All probabilities except for the largest support point are approximated from below by rational values $q_1, \ldots, q_{n-1}$, see figure 7. This will affect the contributions to the expected value by

$$\begin{aligned}
A_1 &= (x_2 - x_1) \cdot (q_1 - p_1) \\
A_2 &= (x_3 - x_2) \cdot (q_1 + q_2 - (p_1 + p_2)) \\
&\vdots \\
A_{n-2} &= (x_{n-1} - x_{n-2}) \cdot (q_1 + \ldots + q_{n-2} - (p_1 + \ldots + p_{n-2}))
\end{aligned}$$

Now, another rational probability value $q_n < 1$ is chosen such that

1. $p_1 + \ldots + p_{n-1} < q_1 + \ldots q_{n-2} + q_n$ and
2. $A_1 + \ldots + A_{n-2} < (x_n - x_{n-1}) \cdot (q_1 + \ldots + q_{n-2} + q_n - (p_1 + \ldots + p_{n-1}))$

If necessary for the two last inequalities to be valid, the formerly picked rational values are set closer to the original probability values. An additional support point $x_0 \in (x_{n-1}, x_n)$ is now picked such that

$$A_1+\ldots+A_{n-2}+(x_0-x_{n-1})\cdot(p_1+\ldots+p_{n-1}-(q_1+\ldots+q_{n-1})) = (x_n-x_0)\cdot(q_1+\ldots+q_n-(p_1+\ldots+p_{n-1}))$$

for a suitable rational probability q_{n-1}. Such an additional support point and a suitable rational probability always exists due to continuity. The last equation ensures equality of the expectations and the single crossing conditions ensures that the approximating distribution is increasingly convex smaller than the original distribution.

An increasingly convex larger distribution is obtained by approximating the first $n-2$ constant segments from above and the final segment from above and from below. ◇

The approximation of irrational probabilities by rational probabilities can be made arbitrarily fine so that rational approximations can be made to converge towards the given distribution with finite support.

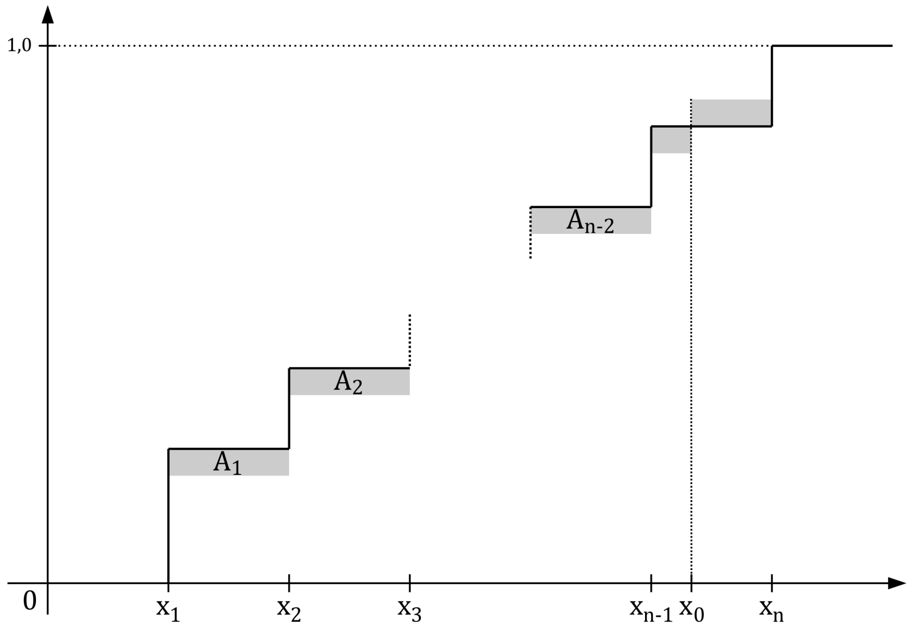

Figure 7: Original distribution function and increasingly convex smaller distribution function with identical expectation and rational levels only.

Lemma 7 *("Converging replacements with rational probabilities")*
For any finite distribution with arbitrary probabilities there exist two sequences of finite distributions which (1) all have the same expected value as the original distribution, (2) have rational probabilities only, are (3a) larger resp. (3b) smaller than the original distribution in increasing convex order and (4) converge in distribution towards the given distribution.

Proof (sketch only). Starting with a distribution as given in the prof of lemma 6, the approximation of all given probabilites can be made arbitrarily fine as, for example, by the settings

$$\begin{aligned} q_i^{(N)} - p_i &= \frac{1}{10^N} \cdot (q_i - p_i),\ i = 1, \dots, n-2, \\ q_{n-1}^{(N)} - p_{n-1} &= \frac{1}{10^N} \cdot (q_{n-1} - p_{n-1}), \\ q_n^{(N)} - p_{n-1} &= \frac{1}{10^N} \cdot (q_n - p_{n-1}). \end{aligned}$$

The sequence of these finite distributions converges towards the given distribution in all continuity points of the given distribution function. In particular, the additional support point x_0 is the same for all approximations and its limiting probability is zero. $\diamond$

An important convergence result which indicates "approximate applicability" of Pigou-Dalton transfers to continuous distributions is the following theorem, see [MS, theorem 1.5.30, p. 30].

Theorem 4 *Let $X \leq_{cx} Y$. Then there exists sequences of random variables X_n and Y_n such that (1) $X_n \leq_{cx} Y_n$ for all n, (2) $EX_n = EX$, $EY_n = EY$ for all n, (3) $X_n \to X$ and $Y_n \to Y$ in distribution and (4) all X_n and all Y_n have finite support.*

The convergence result is now considered for the special case $EX = EY$ so that $\leq_{icx}$ and $\leq_{cx}$ are equivalent. Asymptotically, $X \leq_{icx} Y$ are then replaced by $X_n \leq_{icx} Y_n$ as of theorem 4. Lemma 7 then implies the existence of distributions X_n^* and Y_n^* with $X_n^* \leq_{icx} X_n \leq_{icx} Y_n \leq_{icx} Y_n^*$. All four distributions have the same expectations and the outer have finite support with rational probabilities only. Thus, Pigou-Dalton transfers can be applied to one of the outer distributions to result in the other. This is understood as applying Pigou-Dalton transfers to the arbitrary distributions X and Y.

The drawback of the convergence result from theorem 4 is that convergence in distribution only guarantees pointwise convergence of the approximating distribution functions at continuity points of the limit function. When values at discontinuity points matter, the approximations and, hence the Pigou-Dalton transfers may not refer to the proper income values. This is remedied by strengthening the convergence result to hold, also, in an arbitrary finite collection of discontinuity points of the limit distribution function. Therefore, the orignal construction from the proof of theorem 4 is modified. To keep that modification as simple as possible, only random variables with non-negative values are considered.

Theorem 5 *Let X and Y attain non-negative values only with $X \leq_{cx} Y$ and let $p^* \in (0,1)$. Then there exists sequences of random variables X_n and Y_n such that (1) $X_n \leq_{cx} Y_n$ for all n, (2) $EX_n = EX$, $EY_n = EY$ for all n, (3) $X_n \to X$ and $Y_n \to Y$ in distribution and (4) all X_n and all Y_n have finite support. The convergence holds, also, pointwise for the distribution functions in all finite many discontinuity points of $F_X(x)$ and $F_Y(x)$ with possible exception of all jumps whose sizes sum to a value below p^*.*

Proof (Modification of the proof for theorem 1.5.30 [MS, p. 30-31]). For the survival functions $P(X > x) = 1 - F_X(x)$ and $P(Y > x) = 1 - F_Y(x)$ the integrated survival functions $\pi_X(x) = \int_x^\infty P(X > u)\,du$ and $\pi_Y(x) = \int_x^\infty P(Y > u)\,du$ are decreasing and convex. Moreover, they denote the expected values by $\pi_X(0) = EX$ and $\pi_Y(0) = EY$, see figure 8. By (left-side and right-side) derivation, the distribution function can be retained from the integrated survival function. A jump of a distribution function or a survival function (which is equivalent) corresponds to a jump of the same size in slope of the integrated survival function, see figure 9.

As $EX = EY$, $X \leq_{cx} Y$ is equivalent to $\pi_X(x) \leq \pi_Y(x)$ for all $x \geq 0$. Approximations of the integrated survival functions which maintain the pointwise order will then result in random variables which are ordered in the convex sense.

Now

$$\begin{aligned}\pi_{X_n}(x) &= \max\{0, EX - x, \phi_1(x), \ldots, \phi_n(x)\} \\ \pi_{Y_n}(x) &= \max\{0, EY - x, \psi_1(x), \ldots, \psi_n(x)\}\end{aligned}$$

where $\{\phi_1(x), \phi_2(x), \ldots\}$ is a countable set of linearly decreasing support functions of $\pi_X(x)$ which contains two functions for each discontinuity point of $F_X(x)$ with jump size exceeding p^*. The two functions for each such discontinuity point have the extreme slopes as in figure 9. Similarly, $\{\psi_1(x), \psi_2(x), \ldots\}$ is a countable set of linearly decreasing support functions of $\pi_Y(x)$ which contains two functions for each discontinuity point of $F_Y(x)$ with jump size exceeding p^*.

Convexity of the integrated survival functions of X and Y ensures that all supporting linear functions lie below them. Hence $\pi_{X_n}(0) = \max\{0, EX, \phi_1(0), \ldots, \phi_n(0)\} = EX = \pi_X(0)$ so that $EX_n = EX$. Similarly, $EY_n = EY$. The desired convergence and order properties follow from dominated convergence as in [MS]. $\diamond$

It appears that theorem 5 can be strengthened such that all approximating random distributions have rational jump heights only. This issue has not been investigated here.

2.4 A probabilistic version of Pigou-Dalton transfers

The single crossing condition also enables probabilistic versions of Pigou-Dalton transfers. Such transfers will refer to probability of certain incomes rather than to incomes directly. Conceptually, probabilistic Pigou-Dalton transfers are more complicated than their ordinary counterparts but working with them in a probabilistic context is much simpler. As a preparatory step, an ordinary Pigou-Dalton transfer is expressed in terms of distributions. Consider the sample transfer

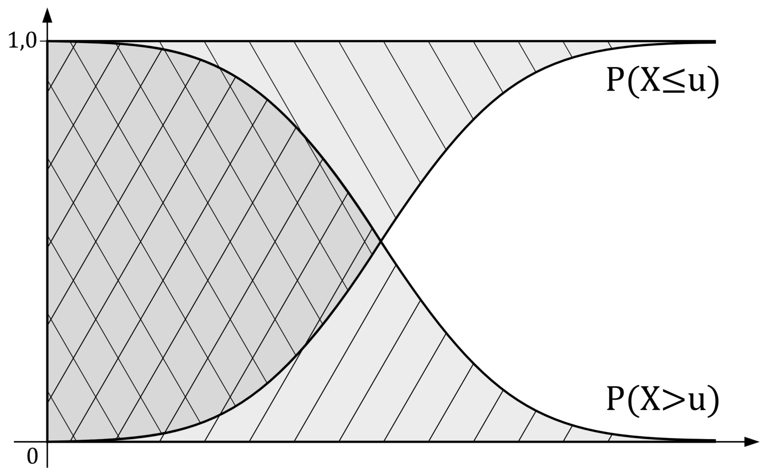

Figure 8: The expected value (shaded areas) is computable by integration over the distribution function as well as over the survival function.

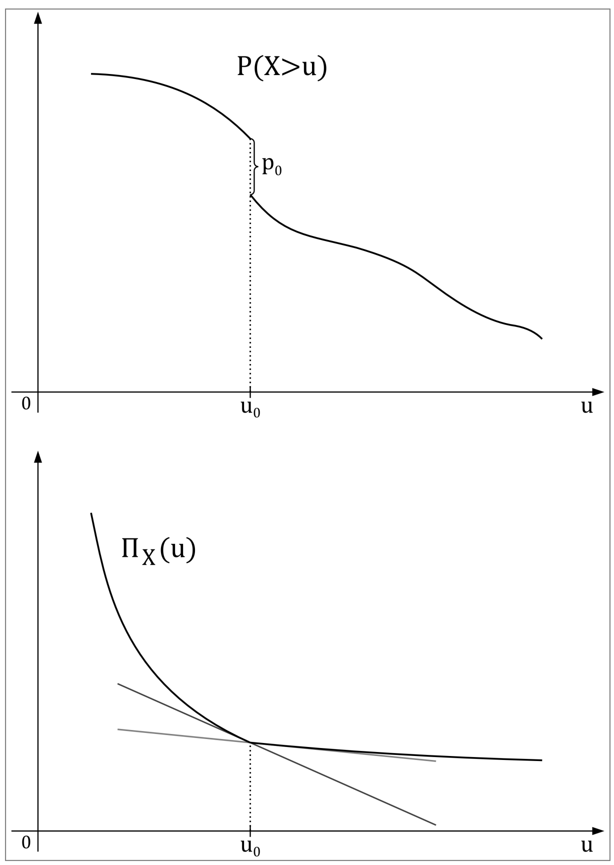

Figure 9: A jump of size p_0 of the distribution function or survival function corresponds to a jump in slope of size p_0 of the integrated survival function. The jump in slope is expressed by two linear functions which support the convex function at the same location with right and left slopes differing by p_0.

$$\begin{pmatrix} 10 \\ 8 \\ 7 \\ 4 \\ 3 \\ 2 \end{pmatrix} \geq_m \begin{pmatrix} 10 \\ 7.5 \\ 7.5 \\ 4 \\ 3 \\ 2 \end{pmatrix}.$$

The effect of this transfer on the natural distribution functions is sketched in figure 10. A value transfer becomes a transfer of probability. Moreover, the single crossing condition implies that the distribution after transfer is the increasing convex smaller of the two. This perspective allows a distributional view that is no longer attached to vectors.

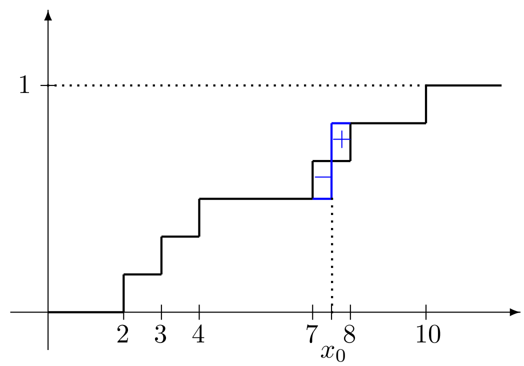

Figure 10: Natural distribution functions of an ordinary Pigou-Dalton transfer and crossing point x_0 showing that the majorized vector belongs to the smaller distribution in the increasing convex order.

When some amount of probability of an arbitrary distribution function is shifted towards smaller values such that expectations remain identical, then the distribution after transfer is increasingly convex smaller than the original distribution, see figure 11. Though transfers may "enclose" the expectation, this need not be so and suitable transfers result in the one-point distribution at the expectation being the smallest distribution in increasing convex order which can be obtained by probabilistic Pigou-Dalton transfers.

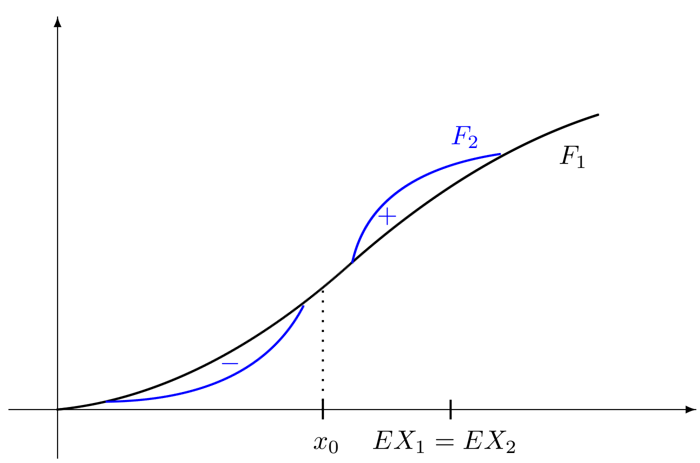

Figure 11: Distribution F_2 is smaller than distribution F_1 in increasing convex order according to the single crossing condition. The point x_0 required for the condition may be any point between the transfer sections. The upper transfer section "encloses" the common expected value.

2.5 Further issue

An equivalence to the Lorenz order without equality of expectations can be obtained from the harmonic new better than used in expectation order from reliability analysis.

Definition 5 *Let $0 < EX, EY < \infty$. Then the* <u>*harmonic new better than used in expectation*</u> *HNBUE order is defined by the condition* $\frac{1}{EX}\int_x^\infty P(X > t)\,dt \le \frac{1}{EY}\int_x^\infty P(Y > t)\,dt$ *for all $x > 0$. Notation $X \le_{HNBUE} Y$.*

Theorem 6 *("Equivalence of Lorenz order and HNBUE order even for unequal means ")*
$X \le_{HNBUE} Y$ if and only if $X \le_{cx} Y$.

Proof. See [BoBe, theorem 5]. $\diamond$

3 Partial orders, Lorenz curves and income distributions

The relation between the Lorenz order and majorization for income distribution has been formulated by the Atkinson theorem. It refers to finite distributions but can easily be extended to other distributions when applying several of the foregoing results. For clarity, the Atkinson theorem is stated for finite distributions and is then extended. The proof of the extended case covers the original formulation

Theorem 7 *("Atkinson theorem")*
Let X and Y be finite distributions with equal expectations. Then $X \le_L Y$ if and only if $X \ge_{icv} Y$.

Theorem 8 *("Extended Atkinson theorem")*
Let X and Y be distributions with finite, equal expectations. Then $X \le_L Y$ if and only if $X \ge_{icv} Y$.

Proof. Equality of the expectations admits the following equivalences

$$\begin{aligned} X \le_L Y &\iff X \le_{cx} Y \\ &\iff Y \le_{cv} X \\ &\iff Y \le_{icv} X \end{aligned}$$

The first equivalence follows from theorem 3, the second from the general relation between the convex and concave orderings and the last from lemma 3.2. $\diamond$

As a consequence of Atkinson's theorem and its extension, a distribution being less uneven than another in the sense $X \le_L Y$ but with same mean value is equivalent to $Ef(X) \ge Ef(Y)$ for all increasing and concave functions. The integration functions are sometimes called welfare functions in relation to Atkinson's theorem. It should be noted that the Lorenz order is equivalent to the concave order so that the welfare functions need not be increasing but concave only. Yet, this relaxation may have a limited economical meaning.

References

[BB] Beckenbach, E.F., Bellman, R., "Inequalities", 3rd ed., Springer, Berlin, 1971.

[BoBe] Borzadaran, G.R.M., Behdani, Z., "Notes due to Lorenz curves and Lorenz ordering in view of weighted distributions", Univ. of Isfahan, 2008.

[MOA] Marshall, A.W., Olkin, I., Arnold, B.C., "Inequalities: theory of majorization and its applications", 2nd ed., Springer, New York, 2010.

[Mö] Möller, T., "Stochastic orders in dynamic reinsurance markets", Finance and Stochastics 8, 2004, p. 479-499.

[MS] Müller, A., Stoyan, D., "Comparison methods for stochastic models and risks", Wiley, New York, 2002.

[S] Stoyan, D., "Qualitative Eigenschaften und Abschätzungen stochastischer Modelle", Oldenbourg, München, 1977.