Lorenz curves — history, state and some recent results

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

February 15, 2012

Abstract

Lorenz curves are established under general conditions following an integral transform approach aging approximately 40 years. The derivation of this approach is given together with quite a few explanations, well known and widely-spread (yet well hidden?) facts from the literature, and some fresh views. The approach is based on relating the underlying distributions, typically income distributions, to probability distributions.

The equity calculus, which is inspired by a notion of relative poverty, is shown to deliver a variety of differential equations for Lorenz curves. Most admit one-parametric closed form solutions while solvability of others is still unsettled.

Key words: generalized function inversion, integral transforms, absolute continuity, Pareto distribution, differential equations, symbolic solutions.

1 Introduction

The idea of Lorenz curves, which are more than one hundred years old, is to describe unevenness of data samples independent from the number of data and their size – and their origin. Different data sets and, slightly more general, different distributions of real values become comparable beyond their absolute magnitudes.

Lorenz curves for any finite income distribution, which, basically, is a finite collection of income values, can be introduced in two ways. One is an ad-hoc approach consisting of one operation of sorting and two operations of normalization. More complicated income distributions require another, more rigorous introduction. One benefit of the latter is that the unevenness of large data sets can receive approximate descriptions by few parameters only.

Here, Lorenz curves will be derived and analyzed for three kinds of distributions and every mixture of these kinds: distributions with countable support set, distributions with Lebesgue density and distributions with uncountable support sets of Lebesgue measure zero, so-called singular distributions. Attaining this generality entails, among others, a detailed consideration of the generalized inverse of probability distribution functions.

A considerable portion of the work is dedicated to the equity calculus. This calculus is governed by proportionality laws that relate individual incomes to averages of certain income ranges. Proportionality laws exhibit certain systematic patterns. Various proportionality laws can be established and their Lorenz curves can be derived in closed form for some of them. A one-parametric version of the Pareto distribution is the lead. Up to date symbolic solution techniques are used. But not all proportionality laws lend to closed form solutions. It is firmly believed that many proportionality laws and corresponding Lorenz curves have not been published before.

Though there is an enormous body of literature on various aspects of the subject, only selected references are given. These cover aspects which are considered as rather fundamental or interesting for their own sake. Quite a few visualizations and most numerical computations were enabled by the free platform Scilab [Sci].

2 Preliminaries

For every real-valued random variable X its distribution function is defined as $F(x) = P(X \leq x)$ for all real numbers x. The values that a random variable may actually attain are often denoted as events. A distribution function accumulates the probability of all events up to level x. Every distribution function has the four properties of being increasing, right-continuous, $\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$. The converse is also true which means that every function with the foregoing four properties entails a real-valued random variable – even if the function is obtained from outside any probabilistic context.

For example, a random variable U which is uniformly distributed between zero and one has the distribution function

$$F_U(x) = \begin{cases} 1, & \text{if } \ x \geq 1 \\ x, & \text{if } \ 0 < x < 1 \\ 0, & \text{if } \ x \leq 0. \end{cases}$$

Two random variables X and Y having the same distribution is denoted by $P^X = P^Y$ and $X \sim Y$. Two distributions are known to be equal if their distribution functions are equal at all arguments. Two or more distributions being stochastically independent and identically distributed is abbreviated by i.i.d. The end of a formal argument will always be denoted by a $\diamond$.

3 A "gentle" derivation of the generalized inverse for distribution functions

A distribution function may be invertible in the usual sense of function inversion or it may be not. If not, the properties of being right-continuous and increasing – though not strictly increasing – allow to establish a surrogate inverse which behaves like the ordinary inverse in many ways. This generalized inverse is often introduced without derivation in the literature. In the following, such a derivation is given in terms of plain probability.

Any real-valued random variable can be inserted into its own distribution function to yield another random variable which attains only values between zero and one. Formally, a random variable X with distribution function F always leads to a new random variable $F(X)$. When the distribution function is invertible in the ordinary sense, the new random variable is uniformly distributed between zero and one; $F(X) \sim U$. This can be seen in a straightforward manner by computing the distribution function of the new random variable for all arguments $u \in (0, 1)$:

$$\begin{aligned} P(F(X) \leq u) &= P(F^{-1}(F(X)) \leq F^{-1}(u)) \\ &= P(X \leq F^{-1}(u)) \\ &= F(F^{-1}(u)) \\ &= u. \end{aligned}$$

Also, the converse is true for invertible distribution functions meaning that inserting a zero-one uniform random variable U into the inverse distribution function reproduces the original distribution; $F^{-1}(U) \sim X$. This can be seen by computing the distribution function of $F^{-1}(U)$ for all arguments $x \in \mathbb{R}$:

$$\begin{aligned} P(F^{-1}(U) \leq x) &= P(F(F^{-1}(U)) \leq F(x)) \\ &= P(U \leq F(x)) \\ &= F_U(F(x)) \\ &= F(x). \end{aligned}$$

Interestingly, this reproduction of the original random variable can also be obtained for non-invertible distribution functions. It is therefore observed – still assuming invertibility – that the first two of the foregoing equations hold not only in probability but as equality of sets:

$$\{u|\, F^{-1}(u) \leq x\} = \{u|\, u \leq F(x)\}.$$

The set equality can be written as the equivalence

$$F^{-1}(u) \leq x \Longleftrightarrow F(x) \geq u.$$

Suppose now that the inverse does not exist. Searching a function G with the reproduction property $P(G(U) \leq x) = P(U \leq F(x))$ then motivates the equivalence

$$G(u) \leq x \Longleftrightarrow F(x) \geq u.$$

Thus, if such a function G exists at all, it must satisfy $G(u) \leq x$ for all x with $F(x) \geq u$. This implies the lower bound inequality

$$G(u) \leq \inf\{x|\, F(x) \geq u\}$$

for all arguments $u \in (0,1)$. The infimum itself is a candidate for function G, namely the largest. It actually has the desired reproduction property, see theorem 1 below, and can hence be defined as the generalized inverse.

Definition 1 *For any increasing and right-continuous function* $F : I\!R \to [0,1]$ *with* $lim_{x\to-\infty}F(x) = 0$ *and* $lim_{x\to\infty}F(x) = 1$ *its generalized inverse is defined for arguments between zero and one by*

$$F^{-1}(u) = \inf\{x|\, F(x) \geq u\}.$$

The generalized inverse distribution function, also denoted as quantile function, indicates the minimum size of events such that the probabilities of all events up to that size accumulate to a given probability value u; when the given probability value cannot be met exactly – due to "discretization effects" – it will be exceeded by the smallest possible amount. Typical mappings of the generalized inverse are indicated in figure 1. Ordinary and generalized inverse are identical if both exist, see below. Thus they will both be denoted the same, unless a distinction is needed. Then the ordinary and the generalized inverse are abbreviated by $F_{ord}^{-1}(u)$ and $F_{gen}^{-1}(u)$.

A finite distribution with probabilities $P(X = x_i) = p_i > 0$, $i = 1, \ldots, n$ and $p_1 + \ldots + p_n = 1$ over sorted events $x_1 < \ldots < x_n$ has the distribution function and generalized inverse distribution function with respective arguments $x \in I\!R$ and $u \in (0,1)$

$$F(x) = \begin{cases} 0, & \text{if } x < x_1 \\ p_1 + \ldots + p_i, & \text{if } x_i \leq x < x_{i+1} \text{ for } i = 1, \ldots, n-1 \\ 1, & \text{if } x_n \leq x \end{cases}$$

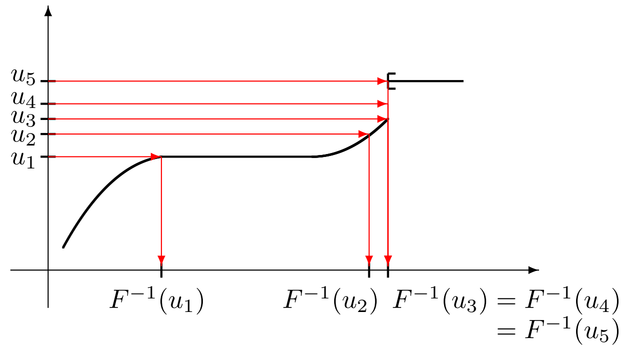

Figure 1: Values of the generalized inverse for five points of its domain.

$$F^{-1}(u) = \begin{cases} x_1, & \text{if} \quad u \le p_1 \\ x_2, & \text{if} \quad p_1 < u \le p_1 + p_2 \\ \vdots & \vdots \\ x_{n-1}, & \text{if} \quad p_1 + \ldots + p_{n-2} < u \le p_1 + \ldots p_{n-1} \\ x_n, & \text{if} \quad p_1 + \ldots + p_{n-1} < u. \end{cases}$$

4 Properties of the generalized inverse distribution function

Theorem 1 *("Reproduction property")*
Let X be a real-valued random variable with distribution function F and let U be uniformly distributed between zero and one. Then $F^{-1}(U) \sim X$.

Proof. It will be shown that the distribution functions of $F^{-1}(U)$ and of X are identical.

$$\begin{aligned} P(F^{-1}(U) \le x) &= P(U \le F(x) \\ &= F_U(F(x) \\ &= F(x). \end{aligned}$$

The second equation is nothing but the definition of the distribution function for U and the third equation follows from $F_U(v) = v$ for all $v \in (0,1)$ so that the first equation remains to be shown. As in the derivation section 2, the equality follows from the equivalence $F^{-1}(u) \le x \Longleftrightarrow F(x) \ge u$" which is verified next.

"$\Longleftarrow$" $F(x) \ge u$ implies $F^{-1}(u) = \inf\{w | F(w) \ge u\} \le x$, because x belongs to the set over which the infimum is formed.

"$\Longrightarrow$" $F^{-1}(u) \le x$ implies $x \ge F^{-1}(u) = inf(A) = x_0$ for $A = \{w | F(w) \ge u\}$.

Case 1: $x_0 \in A$. Then $F(x_0) \ge u$ and, by monotonicity of the distribution function $F(x) \ge F(x_0) \ge u$.

Case 2: $x_0 \notin A$. Then exists a sequence $(x_i)_{i=1}^{\infty} \subseteq A$ with $\lim_{i\to\infty} x_i = x_0$ and $x_i > x_0$ for all i. Right-continuity of F then implies $F(x_0) = \lim_{i\to\infty} F(x_i) \ge u$. Monotonicity of the distribution function again implies $F(x) \ge F(x_0) \ge u$. ◇

The reproduction property enables the inversion method in Monte-Carlo simulation to generate random numbers of an arbitrary distribution from uniformly distributed random numbers. The method will work whenever the generalized inverse can be computed explicitly. Note that the distributional equality $F(X) \sim U$, which was the first considered in section 3, need no longer hold when the distribution function is not invertible in the ordinary sense.

Lemma 1 *("Basic properties of the generalized inverse")*

1. $F^{-1}(F(x)) \leq x$ *for all* $x \in \mathbb{R}$.
2. $F(F^{-1}(u)) \geq u$ *for all* $u \in (0,1)$.
3. *Function* $F^{-1}(u)$ *is increasing (though not necessarily strictly increasing).*
4. *Function* $F^{-1}(u)$ *is left-continuous for all* $u \in (0,1)$.
5. $\sup\{x|\, F(x) < u\} = \inf\{x|\, F(x) \geq u\}$ *for all* $u \in (0,1)$.
6. $\min\{x|\, F(x) \geq u\} = \inf\{x|\, F(x) \geq u\}$ *for all* $u \in (0,1)$.

Proof. Part 1. Let x_0 be arbitrary and fixed with $F(x_0) = u_0$. Then

$$\begin{aligned} F^{-1}(F(x_0)) = F^{-1}(u_0) &= \inf\{x|\, F(x) \geq u_0\} \\ &= \inf\{x|\, F(x) \geq F(x_0)\} \\ &\leq x_0, \text{ since } x_0 \in \{x|\, F(x) \geq F(x_0)\}. \end{aligned}$$

Part 2. Let u_0 be arbitrary and fixed between zero and one with $F^{-1}(u_0) = x_0 = \inf(A)$ for $A = \{x|\, F(x) \geq u_0\}$. If $x_0 \in A$ (the infimum belongs to the set) then $F(x_0) \geq u_0$.

If $x_0 \notin A$ (the infimum does not belong to the set) then exists a sequence $(x_i)_{i=1}^{\infty} \subseteq A$ with $\lim_{i\to\infty} x_i = x_0$ and $x_i > x_0$ for all i. Right-continuity of F then implies $F(x_0) = \lim_{i\to\infty} F(x_i) \geq u_0$.

So, in both cases $F(F^{-1}(u_0)) = F(x_0) \geq u_0$.

Part 3. When the argument increases, the set over which the infimum is formed reduces to a subset or remains the same. Hence the infimum will either increase or remain steady.

Part 4. It suffices to show that $\lim_{n\to\infty} F^{-1}(u_n) = F^{-1}(u_0)$ for an increasing sequence $(u_n)_{n=1}^{\infty}$ with $\lim_{n\to\infty} u_n = u_0 < 1$. Since u_n is increasing and bounded from above by u_0, $F^{-1}(u_n)$ is increasing and bounded from above by $F^{-1}(u_0)$ so that it converges to α_0 with $\alpha_0 \leq F^{-1}(u_0)$. Assume $\alpha_0 < F^{-1}(u_0)$.

Then $F^{-1}(u_n) = \inf\{x|\, F(x) \geq u_n\} \leq \alpha_0$ for all n. Thus $u_n \leq F(F^{-1})(u_n) \leq F(\alpha_0)$ for all n according to part 2. This, in turn, implies $F(\alpha_0) \geq u_0$; otherwise there were n_0 with $F(\alpha_0) < u_{n_0} \leq u_0$.

All in all this results in $F^{-1}(u_0) = \inf\{x|\, F(x) \geq u_0\} \leq \alpha_0$ which contradicts the assumption.

Part 5. Let $u_0 \in (0,1)$. Each element of the set $B = \{x|\, F(x) < u_0\}$ is a lower bound of the set A (from the proof of part 2) and each element of A is an upper bound of B so that $\sup(B) \leq \inf(A)$. Assuming $\sup(B) < \inf(A)$ yields the existence of some x_0 with $\sup(B) < x_0 < \inf(A)$. If $F(x_0) < u_0$ then $x_0 \in B$ which leads to the contradiction $x_0 \leq \sup(B) < x_0$ and if $F(x_0) \geq u_0$ then $x_0 \in A$ leads to the contradiction $x_0 < inf(A) \leq x_0$. Thus $\sup(B) = \inf(A)$.

Part 6. As in the proof of part 2, right-continuity of F implies $F(x_0) \geq u_0$ for $x_0 = \inf\{x|\, F(x) \geq u_0\}$. Thus $x_0 \in A$ so that the infimum is attained. ◇

In the light of lemma 1.6 it seems odd to use the less handy infimum rather than the minimum to express the generalized inverse. Yet the arguments $u = 0$ and $u = 1$ may cause trouble for the minimum. In neither case the minimum need to be attained so that the infimum notation is retained together with the conventions $\inf(\emptyset) = \infty$ and $\inf(\mathbb{R}) = -\infty$. The generalized inverse can thus either be defined as a function $F^{-1} : [0,1] \to \mathbb{R} \cup \{-\infty, \infty\}$ via the infimum or as $F^{-1} : (0,1) \to \mathbb{R}$ via the minimum.

More minimum formulations may be possible but that depends on the distribution function F actually attaining the values zero and one.

The situation of lemma 1.1 and 1.2 with genuine inequalities is illustrated by figure 2. The left part of the figure also shows that a continuous distribution function may have a discontinuous generalized invers. Every region of constant values strictly between zero and one causes the generalized inverse to have jump; the jump height equals the width of the level region.

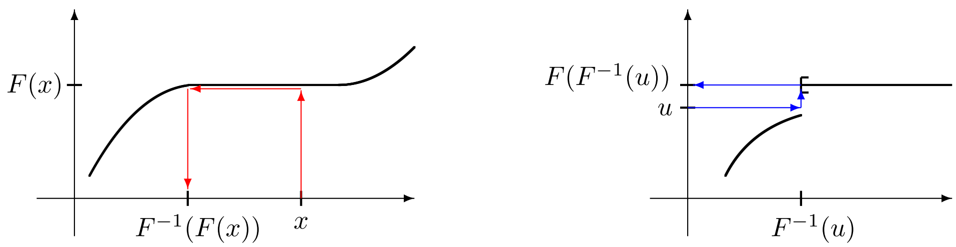

Figure 2: The generalized inverse maps every point of a constant region of the distribution function to the infimum of that region (left, lemma 1.1). At a jump, an enclosed probability value is lifted to the upper jump level (right, lemma 1.2).

Distributions functions and generalized inverses being increasing and continuous from one side (though from different sides) makes them share a certain distinction property.

Lemma 2 *("Clear distinction property")*
If two distribution functions differ in only one point, they differ on a set of strictly positive Lebesgue measure and if the generalized inverses of two distribution functions differ in only one point, they differ on a set of strictly positive Lebesgue measure.

Proof. The proof will only be given for generalized inverses as it is similar for distribution functions but the result for inverses will be used below. Assume that $F_1^{-1}(u_0) \neq F_2^{-1}(u_0)$ for some $u_0 \in (0,1)$. Then the two inverses are different on an interval of values below u_0 which can be seen as follows.

Without loss of generality one value is the larger: $F_1^{-1}(u_0) > F_2^{-1}(u_0)$. This leads to a value $\varepsilon = F_1^{-1}(u_0) - F_2^{-1}(u_0) > 0$, see figure 3. Monotonicity and continuity of $F_1^{-1}(u)$ from the left imply that for $\varepsilon/2$ some $\delta > 0$ exists so that all u with $0 < u_0 - u < \delta$ satisfy $0 < F_1^{-1}(u_0) - F_1^{-1}(u) < \varepsilon/2$. Then

$$\begin{aligned} F_1^{-1}(u) &> F_1^{-1}(u_0) - \frac{\varepsilon}{2} \\ &= \frac{F_1^{-1}(u_0)}{2} + \frac{F_2^{-1}(u_0)}{2} \\ &> F_2^{-1}(u_0) \\ &\geq F_2^{-1}(u) \text{ for all } u < u_0. \end{aligned}$$

The last inequality follows from the monotonicity of the second generalized inverse. Altogether, the generalized inverses differ on the interval $(u_0 - \delta, u_0]$. ⋄

Lemma 3 *("Relation between generalized and ordinary inverse")*
When the distribution function F is continuous and invertible in the ordinary sense, then the ordinary inverse and the generalized inverse are equal over $(0,1)$.

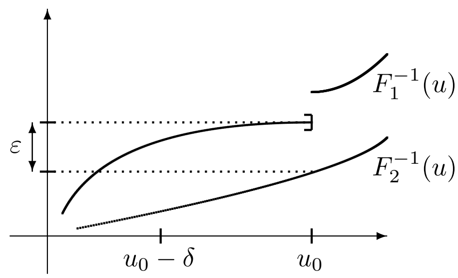

Figure 3: Visualization of clear distinction property.

Proof. With $A = \{x | F(x) \geq u\}$, as in the proof of theorem 1, the generalized inverse $F_{gen}^{-1}(u) = x_0 = inf(A)$ satisfies $F(x_0) \geq u$. Each $x < x_0$ (value below infimum) then fulfills $x \notin A$ so that $F(x) < u$. A sequence of values $x_n < x_0$ converging to x_0 then satisfies $u \leq F(x_0) = \lim_{n\to\infty} F(x_n) \leq u$. The equality follows from the continuity of the distribution function.

Thus $u = F(x_0)$ so that ordinary invertibility implies $x_0 = F_{ord}^{-1}(x_0)$. $\diamond$

When the distribution function is not continuous, it has jumps but may still be invertible over its range. The ordinary inverse then is a partially defined function only in contrast to the generalized inverse, see figure 4.

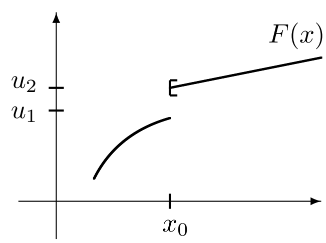

Figure 4: Distribution function with ordinary inverse which is not defined over $[u_1, u_2)$ so that it differs from the generalized inverse; $F_{gen}^{-1}(u_1) = F_{gen}^{-1}(u_2) = F_{ord}^{-1}(u_2) = x_0$ and $F_{ord}^{-1}(u_1)$ is not defined.

Though x and $F^{-1}(F(x))$ may be different when, for example, F is constant around x, yet another application of the distribution function flattens out the inequality to equality. The same applies to the inequality between u and $F(F^{-1}(u))$. This clarifies the interplay between a distribution function and its generalized inverse when the latter is not the ordinary inverse.

Lemma 4 *("Pointwise inversion' formulas')*

1. $F(x) = F(F^{-1}(F(x)))$.
2. $F^{-1}(F(F^{-1}(u))) = F^{-1}(u)$.

Proof. Part 1. It can be observed that (1) $F^{-1}(F(x)) \leq x$ because of lemma 1.1 so that monotonicity of the distribution function implies $F(F^{-1}(F(x))) \leq F(x)$. (2) On the other hand, $F(F^{-1}(u)) \geq u$ according to lemma 1.2, so hat inserting $u = F(x)$ results in $F(F^{-1}(F(x))) \geq F(x)$.

Part 2. Analogous. $\diamond$

Lemma 5 *("Probabilistic inversion formulas")*

1. $P(F^{-1}(F(X)) = X) = 1$.
2. $P(X \in F^{-1}(0,1)) = 1$.

Proof. Part 1. $P(F^{-1}(F(X)) = X) = P(F^{-1}(F(F^{-1}(U))) = F^{-1}(U)) = 1$. The first equation holds because of theorem 1 and the second because the equality in the set under consideration holds pointwise according to lemma 4.2.

Part 2. $P(X \in F^{-1}(0,1)) = P(F^{-1}(U) \in F^{-1}(0,1)) = 1$. The first equality follows, again, from theorem 1 and the second holds pointwise for all $u \in (0,1)$. ◇

Reversing the order of the generalized inverse and the distribution function may destroy their almost sure opposing effect, since $P(F(F^{-1}(U)) = U) < 1$ holds for discrete distributions.

5 Generalized inverse and order relations

A consequence of lemma 5 is that the distribution functions of X and $F^{-1}(F(X))$ are always equal. But when F is not invertible in the ordinary sense, the distribution functions of U and $F(F^{-1}(U))$ need not be equal. Yet, the latter distribution function consistently lies below the former. This is denoted as stochastic order.

Definition 2 *Let F and $G(x)$ be distribution functions with $F(x) \geq G(x)$ for all $x \in \mathbb{R}$. Then $F(x)$ is* **stochastically smaller** *than $G(x)$ which is abbreviated by $F \leq_{ST} G$.*

The stochastic order between distribution functions is, also, understood as order between distributions and random variables.

Lemma 6 $U \leq_{ST} F(F^{-1}(U)) \sim F(X)$.

Proof. The stochastic inequality follows from the pointwise inequality of lemma 1.2 and the distributional equality follows from theorem 1. ◇

The relation between the distribution considered can now be completely summarized in table 1. The stochastic order decreases to identity whenever the distribution function is changed to become invertible in the ordinary sense. Thus, the order can be considered as an indicator of "non-invertibility"; it does not indicate magnitude of events. In particular, the difference $EF(X) - EU = EF(X) - 1/2$ can be used as a crude measure of non-invertibility.

F has ordinary inverse	F has generalized inverse only
$U \sim F(X)$	$U \leq_{ST} F(X)$ with $U \not\sim F(X)$
$F^{-1}(U) \sim X$	$F^{-1}(U) \sim X$

Table 1: Relation between random variable X with distribution function F and random variable U uniformly distributed between zero and one.

For an example, X is taken as a finite distribution with probabilities $P(X = 3) = 0.2$, $P(X = 5) = 0.5$ and $P(X = 9) = 0.3$. Then $F(X)$ also has a finite distribution with probabilities $P(F(X) = 0.2) = 0.2$, $P(F(X) = 0.7) = 0.5$ and $P(F(X) = 1) = 0.3$. Thus $EF(X) - EU = 0.2 \cdot 0.2 + 0.5 \cdot 0.7 + 0.3 \cdot 1 - 0.5 = 0.19$. The distribution function of $F(X)$ is shown in figure 5.

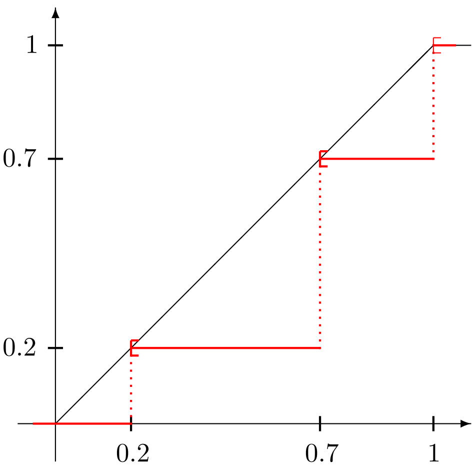

Figure 5: The distribution function of $F(X)$ lies below that of U except at finite many points over the relevant domain $[0, 1]$.

Lemma 7 *("Inverted order for the inverses")*
Let F and G be distribution functions with $F(x) \leq G(x)$ for all $x \in \mathbb{R}$. Then their generalized inverses are inversely ordered meaning that $G^{-1}(u) \leq F^{-1}(u)$ for all $u \in (0, 1)$.

Proof. The sets $A = \{x | F(x) \geq u\}$ and $C = \{x | G(x) \geq u\}$ are related by $A \subseteq C$ since $F(x) \geq u$ implies $G(x) \geq u$. Thus $G^{-1}(u) = \inf(C) \leq \inf(A) = F^{-1}(u)$. ⋄

Theorem 2 *("Approximation of generalized inverse from above")*
Let $(F_n(x))_{n=1}^{\infty}$ be a sequence of continuous distribution functions with ordinary inverses that converges for all real values x decreasingly ("from above") to a distribution function $F(x)$ with generalized inverse only. Then $\lim_{n\to\infty} F_n^{-1}(u) = F^{-1}(u)$ for all $u \in (0, 1)$.

Proof. $F_n(x) \geq F_{n+1}(x) \geq F(x)$ for all n and all real values x implies $F_n^{-1}(u) \leq F_{n+1}^{-1}(u) \leq F^{-1}(u)$ for all n and all u between zero and one. Thus, the sequence $(F_n^{-1}(u))_{n=1}^{\infty}$ is increasing and bounded for each u so hat it converges with $\lim_{n\to\infty} F_n^{-1}(u) \leq F^{-1}(u)$. Assuming that $\alpha_0 = \lim_{n\to\infty} F_n^{-1}(u) < F^{-1}(u)$ implies that $F_n^{-1}(u) \leq \alpha_0$ so that, exploiting ordinary invertibility and monotonicity of each approximating distribution function leads to $u \leq F_n(\alpha_0)$ for all n. This implies $\lim_{n\to\infty} F_n(\alpha_0) \geq u$.

Now $u > F(\alpha_0)$ which can be seen by setting $F^{-1}(u) = inf(A) = \alpha$ with $A = \{x | F(x) \geq u\}$. Hence $\alpha_0 < \alpha$ (α_0 below infimum of A) implies $\alpha_0 \notin A$ so that $F(\alpha_0) < u$.

All in all, this leads to $\lim_{n\to\infty} F_n(\alpha_0) \geq u > F(\alpha_0)$ which contradicts pointwise convergence of the distribution functions towards their limit function at α_0. ⋄

The limiting process from theorem 2 can be used for an alternative definition of the generalized inverse provided that an approximating sequence from above by continuous and invertible distribution functions always exists. This will be discussed below and one form of the alternative definition of generalized inverses is

$$F_{gen}^{-1}(u) = \sup\{G_{ord}^{-1}(u) | G \text{ continuous distribution function, ordinarily invertible and } G(x) \leq_{ST} F(x)\}.$$

Without approaching from above, a sequence of distribution functions with ordinary inverses converging to a distribution function with generalized inverse only, need not converge to the generalized inverse for every argument. This even so, when all approximating distribution functions lie below the limiting distribution function. Formally, it is possible to have

1. $F_n(x) \leq F(x)$ for all n and all x,
2. $\lim_{n\to\infty} F_n(x) = F(x)$ for all x
3. but $\lim_{n\to\infty} F_n^{-1}(u) \neq F^{-1}(u)$ for some arguments $u \in (0,1)$.

A situation with all three properties is sketched in figure 6.

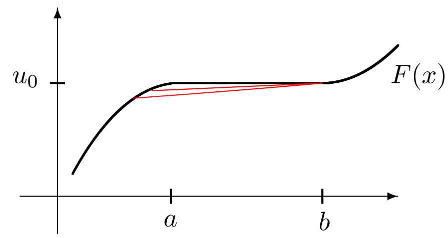

Figure 6: The approximating functions replace the distribution function $F(x)$ over its level region $[a,b]$ and slightly to the left by a strictly increasing line segment. When the left boundary of the replacement converges to a, then $\lim_{n\to\infty} F_n(x) = F(x)$ for all x but $\lim_{n\to\infty} F_n^{-1}(u_0) = b > a = F^{-1}(u_0)$.

6 Generalized inverse and computations

Though the generalized inverse distribution function may appear to be unhandy and static at first sight, it allows some computations.

6.1 Generalized inverses from generalized inverses

Transformations of variables may lead from distribution functions to other distribution functions or from their generalized inverses to other generalized inverses. Variable transformations may apply, in the first place, to distributions functions and inverses.

A strictly increasing and continuous function $S : I\!R \to I\!R$ with $S(I\!R) = I\!R$ leads to a distribution function $F_2(x) = F_1 \circ S(x)$ when $F_1(x)$ is a distribution function. The generalized inverse is $F_2^{-1}(u) = S^{-1} \circ F_1^{-1}(u)$. A prominent example is positive affine scaling with $S(x) = \alpha \cdot x$ for some constant $\alpha > 0$. The generalized inverses then are related by $F_2^{-1}(u) = 1/\alpha \cdot F_1^{-1}(u)$.

Whenever the function $T : [0,1] \to [0,1]$ is increasing and continuous with $T(0) = 0$ and $T(1) = 1$, the function $F^{-1} \circ T(u)$ is a generalized inverse provided that $F^{-1}(u)$ is. A noticeable example is $T(u) = \sqrt{u}$. Function $F^{-1}(\sqrt{u})$ then is the generalized inverse distribution function for the maximum of two i.i.d. random variables.

$$\begin{aligned} F_{max}(x) &= P(\max\{X_1, X_2\} \leq x) \\ &= P(X_1 \leq x, X_2 \leq x) \\ &= P(X_1 \leq x) \cdot P(X_2 \leq x) \\ &= F^2(x). \end{aligned}$$

Now it can be seen that operations on a distribution function may find their counterparts as inverse operations on the argument of the generalized inverse.

$$\begin{aligned} F_{max}^{-1}(u) &= \inf\{x|\, F_{max}(x) \geq u\} \\ &= \inf\{x|\, F^2(x) \geq u\} \\ &= \inf\{x|\, F(x) \geq \sqrt{u}\} \\ &= F^{-1}(\sqrt{u}). \end{aligned}$$

6.2 Generalized inverse of the generalized inverse

In analogy to the generalized inverse of distribution functions, generalized inverses themselves can be inverted leading to distribution functions. This however, requires to define the generalized inverse for generalized inverses in a slightly different manner as generalized inverses themselves. As the construction will, later, be needed only for non-negative random variables, only generalized inverses with non-negative values are considered.

Definition 3 *("Generalized inverse of generalized inverses")*
For any increasing and left-continuous function $\varphi : (0,1) \to I\!R_{\geq}$ its generalized inverse is defined for real arguments by

$$\varphi^{-1}(x) = \begin{cases} \sup\{u|\, \varphi(u) \leq x \text{ and } u \in (0,1)\}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases}$$

The definition is adopted from appendix B of [IK] with the convention $\sup(\emptyset) = 0$. Proofs for the generalized inverse of the generalized inverse are restricted to essentials since they work by merely interchanging supremum with infimum and reversing inequalities in proofs for the 'ordinary' generalized inverse.

Lemma 8 *("Generalized inverse of a generalized inverse is a distribution function")*
Let $\varphi : (0,1) \to I\!R_{\geq}$ be increasing and left-continuous. Then $\varphi^{-1}(x)$ has the following properties and, in particular, is a distribution function according to properties 2, 3 and 4:

1. $\varphi(\varphi^{-1}(x)) \leq x$.
2. $\lim_{x\to-\infty} \varphi^{-1}(x) = 0$ *and* $\lim_{x\to\infty} \varphi^{-1}(x) = 1$.
3. $\varphi^{-1}(x)$ *is increasing.*
4. $\varphi^{-1}(x)$ *is right-continuous.*

Proof. Part 1. For fixed x consider $\varphi^{-1}(x) = \sup\{u|\, \varphi(u) \leq x\} = u_0$ with $C = \{u|\, \varphi(u) \leq x\}$. If $u_0 \in C$ then $\varphi(u_0) \leq x$. If $u_0 \notin C$ then there is an increasing sequence $(u_n)_{n=1}^{\infty} \subseteq C$ with $\lim_{n\to\infty} u_n = u_0$. Left-continuity of the given function then implies $\varphi(u_0) = \lim_{n\to\infty} \varphi(u_n) \leq x$.

So, in both cases, $\varphi(\varphi^{-1}(x)) = \varphi(u_0) \leq x$.

Parts 2 and 3. Obvious.

Part 4. It suffices to show that $\lim_{n\to\infty} \varphi^{-1}(x_n) = \varphi^{-1}(x_0)$ for a decreasing sequence $(x_n)_{n=1}^{\infty}$ with $\lim_{n\to\infty} x_n = x_0$. Since x_n is decreasing and bounded from below by x_0, $\varphi^{-1}(x_n)$ is decreasing and bounded from below by $\varphi^{-1}(x_0)$ so that it converges to β_0 with $\beta_0 \geq \varphi^{-1}(x_0)$. Assume $\varphi^{-1}(x_0) < \beta_0$.

Then $\beta_0 \le \varphi^{-1}(x_n) = \sup\{u|\,\varphi(u) \le x_n\}$ for all n. Thus $\varphi(\beta_0) \le \varphi(\varphi^{-1}(x_n)) \le x_n$ for all n by part 1. This, in turn, implies $\varphi(\beta_0) \le x_0$; otherwise there were n_0 with $\varphi(\beta_0) > x_{n_0} \ge x_0$.

All in all this results in $\varphi^{-1}(x_0) = \sup\{u|\,\varphi(u) \le x_0\} \ge \beta_0$ since $u = \beta_0$ is a feasible setting in the supremum. This contradicts the assumption. $\diamond$

The inequality from lemma 8.1 corresponds to $F^{-1}(F(x)) \le x$ as stated in lemma 1.1.

Theorem 3 *("The generalized inverse of a generalized inverse is the original function")*

1. *When $F(x)$ is a distribution function over $[0, \infty)$ with $F(0) = 0$ and when $\varphi(u) = F^{-1}(u)$, then*

$$\varphi^{-1}(x) = F(x)$$

   for all $x \ge 0$.

2. *When $\varphi : (0, 1) \to I\!R_{\ge}$ is a left-continuous and increasing function with inverse $F(x) = \varphi^{-1}(x)$, then*

$$F^{-1}(u) = \varphi(u)$$

   for all $u \in (0, 1)$.

Proof. Part 1. Case 1. $F(x) = u_0$ is attained once. Then $F^{-1}(u_0) = x$ and monotonicity of the generalized inverse distribution function implies $\sup\{u|\,F^{-1}(u) \le x\} = u_0 = F(x)$.

Case 2. $F(x)$ is attained more than once. Monotonicity of the distribution function implies that $u_0 = F(x)$ is attained over some interval and right-continuity of the distribution function implies that the lower interval boundary belongs to that interval. The lower interval boundary equals $F^{-1}(u_0)$. Then $\sup\{u|\,F^{-1}(u) \le x\} = u_0 = F(x)$.

Part 2. Analogous to part 1. $\diamond$

It is important to note that the two inversions in theorem 3 are defined differently. Still, with all due caution, the foregoing reproduction results allow the mnemonic notations $\left(F^{-1}\right)^{-1}(x) = F(x)$ and $\left(\varphi^{-1}\right)^{-1}(u) = \varphi(u)$.

Infimum and supremum are suited to reverse one-sided continuity for monotone functions without inversion. This means that whenever $F(x)$ is increasing and right-continuous, $F^{+}(x) = \inf\{F(y)|\,y > x\}$ is increasing and left-continuous. And when $F(x)$ is increasing and left-continuous, $F^{-}(x) = \sup\{F(y)|\,y < x\}$ is increasing and right-continuous. $F^{-}(x) = F(x) = F^{+}(x)$ for all continuity points of $F(x)$. When $F(x)$ is increasing then $F^{-}(x) \le F(x) = F^{+}(x)$ whenever the function is right-continuous in x and $F^{-}(x) = F(x) < F^{+}(x)$ whenever the function is left-continuous in x. The infimum leaves an increasing and right-continuous function $F(x)$ unchanged: $F^{+}(x) = F(x)$ for all x. And the supremum leaves an increasing and left-continuous function $F(x)$ unchanged: $F^{-}(x) = F(x)$ for all x.

These relations illustrate that generalized inversion of distribution functions and the "opposing" function inversion are based on infimum and supremum rather than both being based on the same. When distribution functions are defined as being increasing and left-continuous, see [Ba], generalized inversion is still possible. All relations between increasing left-continuous and right-continuous functions and their inverses are depicted in the commutating diagram figure 7.

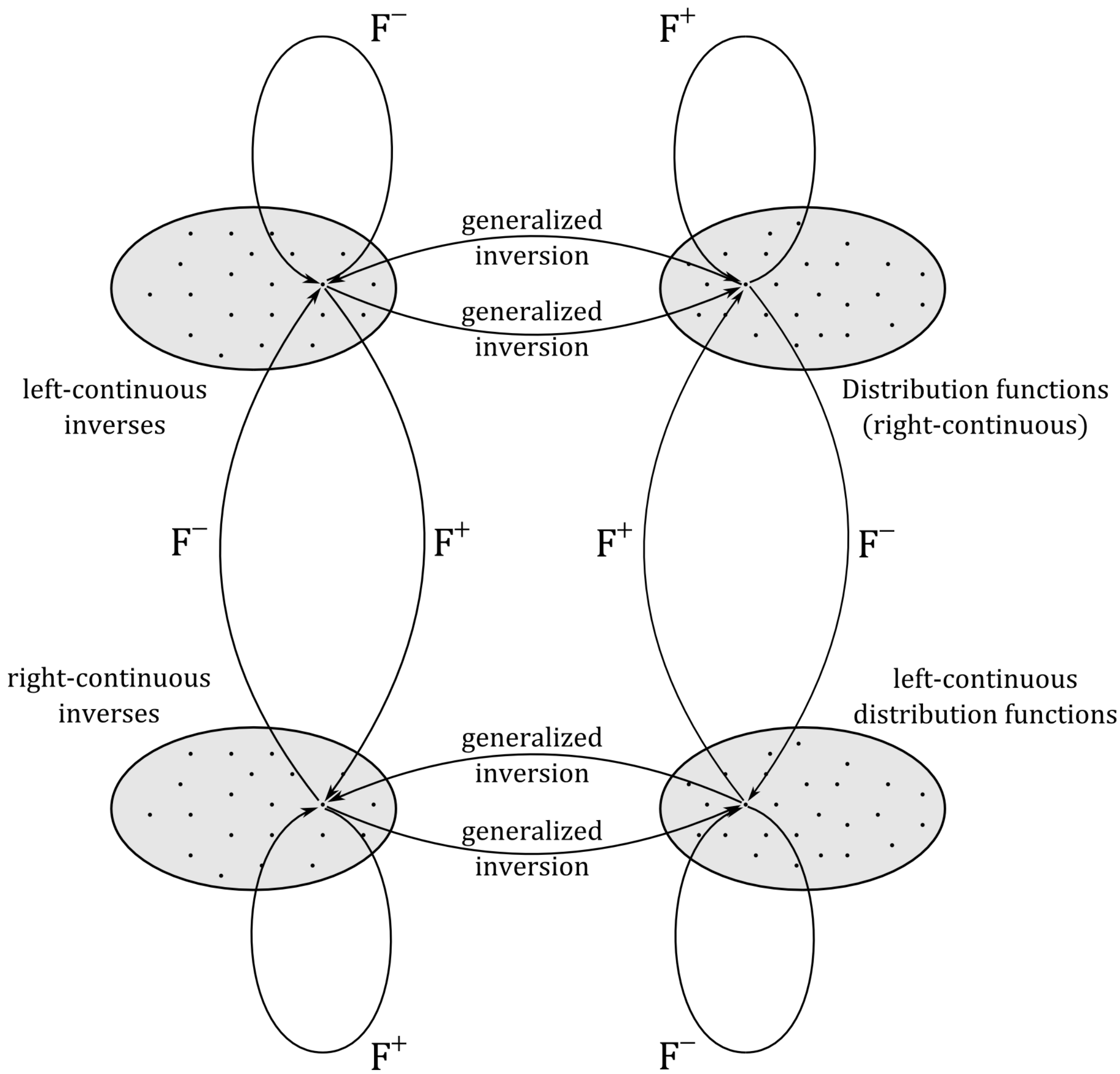

Figure 7: Relation between generalized inversion and exchange of infimum and supremum.

6.3 Generalized inverse and expectation values

The expected value of a discrete finite distribution over the support points x_i with probabilities $P(X = x_i) = p_i > 0$, $i = 1, \ldots, n$ admits the two complementary notations

$$\begin{aligned} EX &= \sum_{i=1}^{n} x_i \cdot P(X = x_i) \\ &= \sum_{i=1}^{n} F^{-1}(p_i) \cdot p_i. \end{aligned}$$

The geometry of both formulas is the same as illustrated in figure 8. A similar formula applies to arbitrary real-valued distributions as long as expectations are finite. That formula is illustrated in figure 9 and stated as the next result.

Theorem 4 *("Expectations from the generalized inverse distribution function")*

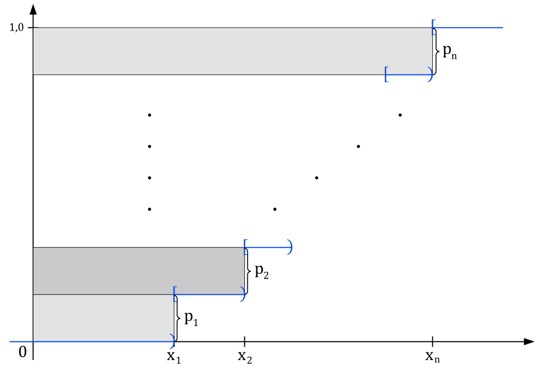

Figure 8: Expected value of a finite discrete distribution computed from stripes of the distribution function.

1. $EX = \int_0^1 F^{-1}(v)\,dv$.
2. $E(X|\,X \leq F^{-1}(u)) = \int_0^u F^{-1}(v)\,dv$ *for all* $u \in (0,1)$.

Proof. Part 1. The proof is based on the general transformation theorem for measures, see [Ba, p. 95]

$$\int_\Omega f \circ T\,d\mu = \int_{T(\Omega)} f\;d\,(T \circ \mu).$$

The particular settings

$$\begin{aligned} T &= F^{-1} \\ \mu &= P^U \\ f(x) &= x \\ \Omega &= (0,1) \end{aligned}$$

result in

$$\begin{aligned} \int_{(0,1)} F^{-1}\,dP^U &= \int_{F^{-1}(0,1)} x\;d\,(P^U)^{F^{-1}}(x) \\ &= \int_{F^{-1}(0,1)} x\;d\,P^X(x) \\ &= \int_{\mathbb{R}} x\;d\,P^X(x). \end{aligned}$$

The second equation follows from theorem 1 and the third from lemma 5.2, since $\mathbb{R} - F^{-1}(0,1)$ is a set of P^X-measure zero and, thus, does not contribute to the integral. This allows to verify the intended equation via

$$\begin{aligned} EX &= \int X\,dP = \int_{\mathbb{R}} x\,dP^X(x) \\ &= \int_{(0,1)} F^{-1}\,dP^U \\ &= \int_{(0,1)} F^{-1}(v)\,d\,\lambda(v) \\ &= \int_{(0,1)} F^{-1}(v)\,d\,v. \end{aligned}$$

◇

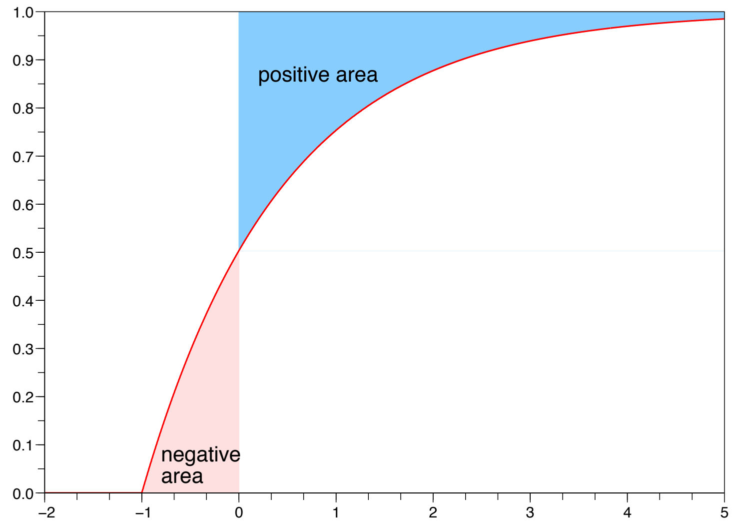

Figure 9: Exponential distribution function $F(x) = 1 - exp(-\lambda \cdot (x+1))$ for $x \geq -1$ and $F(x) = 0$ for $x < -1$ with parameter $\lambda = 0.7$. The expected value equals the sum of the shaded areas with positive (negative) contribution to the right (left) of the ordinate axis.

7 Income distributions

7.1 Introduction of Lorenz densities and Lorenz curves

When X denotes incomes, the equality $P(X \leq x) = P(X \leq F^{-1}(F(x)))$ of lemma 4.1 means that the accumulated probability of all incomes up to level x is equal to the accumulated probability of all incomes up to level only $F^{-1}(F(x))$. Intuitively, the generalized inverse indicates the minimum income level such that the probabilities of all incomes up to that level accumulate to a given probability value u. Averaging the minimum income levels for all probabilities up to value u is the basic idea of the Lorenz curve. A tentative, proportional law is

$$L(u) \sim \int_0^u F^{-1}(v)\, dv = E(X|\, X \leq F^{-1}(u)).$$

Normalization by the expected income which is assumed, from now on, to be finite and strictly positive, makes the minimum income values independent from absolute income levels. Integration will then yield the Lorenz curve [G] and this proceeds along a certain density function.

Definition 4 *The* **Lorenz density** *(of a distribution) is the generalized inverse distribution function normalized by the expected value* $f(u) = \frac{F^{-1}(u)}{\int_0^1 F^{-1}(v)\, dv}$.

Definition 5 *Every distribution (function) induces, via the Lorenz density, a* **Lorenz curve** *which is the function*

$$L(u) = \int_0^u f(v)\, dv.$$

Obviously, the Lorenz curve can be written as

$$L(u) = \frac{\int_0^u F^{-1}(v)\,dv}{\int_0^1 F^{-1}(v)\,dv} = \frac{E(X|\,X \le F^{-1}(u))}{EX}.$$

So far, the generalized inverse was neither defined at zero nor at one. Any finite values by which the generalized inverse might be continued in the boundary points will not alter the integrals so that the Lorenz curve can be considered as a function defined on $[0,1]$ with $L(0)=0$ and $L(1)=1$. To avoid trivial complications, the extra assumption will be made that all incomes are non-negative (almost surely). As these assumptions on underlying incomes are quite important, they are stated again and referred to as "general assumptions":

1. $0 < EX < \infty$.
2. $X \ge 0$.

The second condition is equivalent to $F(x)=0$ for all $x<0$. Under these assumptions, each value of a Lorenz curve admits the interpretation of being the ratio of two areas above the distribution function as indicated in figure 10.

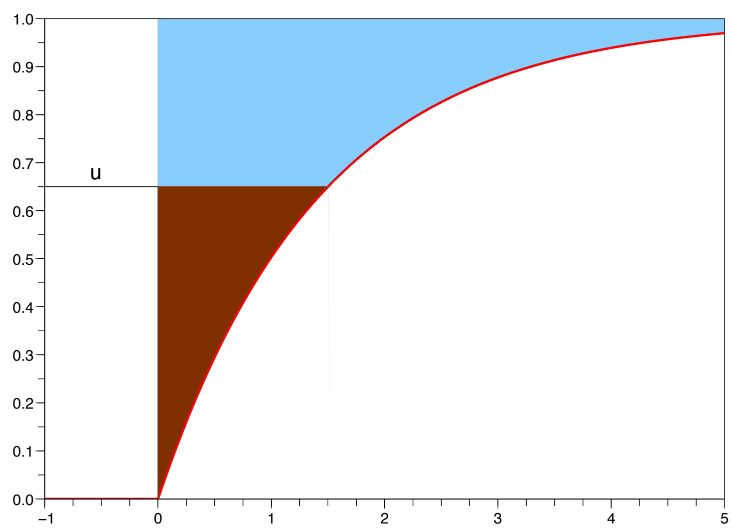

Figure 10: Exponential distribution function $F(x) = 1 - exp(-\lambda \cdot x)$ for $x \ge 0$ and $F(x) = 0$ for $x < 0$ with parameter $\lambda = 0.7$. The value of the Lorenz curve for $u = 0.65$ is the ratio of the dark area over the sum of the dark and the light area.

As a simple example, all incomes are supposed to be equal so that the income distribution is a one-point distribution. The Lorenz curve is then linear which is now computed from the definition. Thus, the definition turns out to be operational even if the generalized inverse consists only of jumps and constant segments. A one-point distribution with single income level x_0 has the distribution function and the generalized inverse

$$F(x) = \begin{cases} 0, & \text{if } x < x_0 \\ 1, & \text{if } x_0 \le x \end{cases}, \quad F^{-1}(u) = x_0 \text{ for all } u \in (0,1).$$

The Lorenz density and the Lorenz curve are now easily computable as, respectively, a constant function and the main diagonal of the unit square

$$\begin{aligned} f(u) &= \frac{F^{-1}(u)}{\int_0^1 F^{-1}(v)\,dv} = \frac{x_0}{1 \cdot x_0} = 1 \\ L(u) &= \int_0^u f(v)\,dv = \int_0^u 1\,dv = u. \end{aligned}$$

For a two-point distribution with strictly positive income levels $x_1 < x_2$ and probabilities $P(X = x_1) = p_1 > 0$ and $P(X = x_2) = p_2 > 0$ the generalized inverse is

$$F^{-1}(u) = \begin{cases} x_1, & \text{if } u \le p_1 \\ x_2, & \text{if } p_1 < u. \end{cases}$$

Lorenz density and Lorenz curve are

$$\begin{aligned} f(u) &= \frac{1}{p_1 x_1 + p_2 x_2} \cdot \begin{cases} x_1, & \text{if } u \le p_1 \\ x_2, & \text{if } p_1 < u \end{cases} \\ L(u) &= \frac{1}{p_1 x_1 + p_2 x_2} \cdot \begin{cases} u\,x_1, & \text{if } u \le p_1 \\ p_1\,x_1 + (u - p_1)\,x_2, & \text{if } p_1 < u \end{cases} \end{aligned}$$

The Lorenz curve is piecewise linear with increasing slope since $x_1 < x_2$, see figure 11.

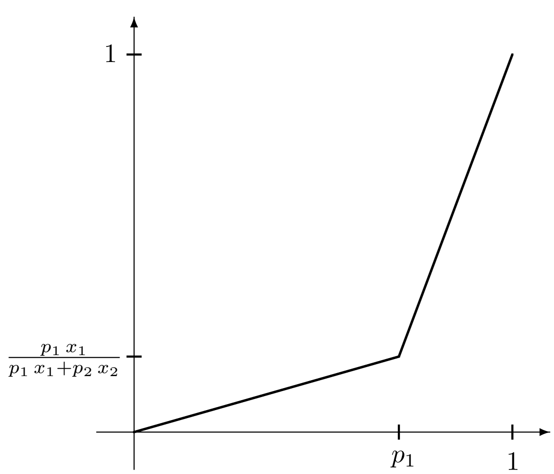

Figure 11: Lorenz curve induced by a two-point distribution.

For a uniform distribution over the interval $[a, b]$ with $0 \le a < b$ the distribution function and the inverse are given as

$$F(x) = \begin{cases} 0, & \text{if } x < a \\ (x-a)/(b-a), & \text{if } a \le x \le b \\ 1, & \text{if } b < x \end{cases}, \quad F^{-1}(u) = a + (b-1) \cdot u \text{ for all } u \in (0, 1).$$

Now, Lorenz density and Lorenz curve are

$$\begin{aligned} f(u) &= \frac{a+(b-a)\cdot u}{\int_0^1 a+(b-a)\cdot v\,dv} = \frac{a+(b-a)\cdot u}{\frac{a+b}{2}} \\ L(u) &= \int_0^u f(v)\,dv = \frac{\int_0^u a+(b-a)\cdot v\,dv}{\frac{a+b}{2}} = \frac{a\,u+\frac{b-a}{2}\cdot u^2}{\frac{a+b}{2}}. \end{aligned}$$

Beginning with a distribution function, the Lorenz curve is always, in principle, derived along the following steps. Whether the Lorenz curve can be obtained in closed form depends on the generalized inverse and the integrals over it being explicitly computable.

Distribution function $F(x)$

↓

Generalized inverse $F^{-1}(u)$

↓

Lorenz density $\frac{F^{-1}(u)}{\int_0^1 F^{-1}(v)\,dv}$

↓

Lorenz curve $L(u) = \frac{\int_0^u F^{-1}(v)\,dv}{\int_0^1 F^{-1}(v)\,dv}$

7.2 Some properties of Lorenz densities and Lorenz curves

Lemma 9 *("Elementary properties of Lorenz curves")*

1. *$L(0) = 0$ and $L(1) = 1$.*
2. *The Lorenz curve is continuous over $[0, 1]$.*
3. *The Lorenz curve is convex over $[0, 1]$.*

Proof. Part 1. trivial.

Part 2. Case 1: $x \in [0, 1)$. When $(x_n)_{n=1}^{\infty}$ converges to x then, eventually, $x_n < 1/2 \cdot (x + 1) = z < 1$ and

$$\begin{aligned} |\int_0^{x_n} f(u)\,du - \int_0^x f(u)\,du| &\leq |x - x_n| \cdot f(\max\{x_n, x\}) \\ &\leq |x - x_n| \cdot f(z) \to 0 (x_n \to x). \end{aligned}$$

Both inequalities follow from every Lorenz density being increasing, according to lemma 1.3 and non-negative according to the general assumptions.

Case 2: $x = 1$. When $(x_n)_{n=1}^{\infty}$ converges to 1 with all $x_n < 1$, the functions $f_n(x) = f(x) \cdot 1_{[0,x_n]}(x)$ converge to $f(x)$ pointwise and, thus, almost surely on $(0, 1)$. They also satisfy the majorization condition $f_n(x) \leq f(x)$ for all arguments. Thus, the dominated convergence theorem [Ba, p. 77] implies

$$\int_0^{x_n} f(x)\,dx = \int_0^1 f_n(x)\,dx \longrightarrow \int_0^1 f(x)\,dx \ (n \to \infty).$$

Part 3. Since L is continuous, it suffices to verify $L(\frac{u+v}{2}) \leq \frac{1}{2}\,L(u)+\frac{1}{2}\,L(v)$ for all $u, v \in [0,1]$. (Belastbare Quelle hierfuer). The function increments $A = L(v)-L(\frac{u+v}{2})$ and $B = L(\frac{u+v}{2})-L(u)$ are now considered for $0 \leq u < v \leq 1$. They increase, see figure 12, since

$$\begin{aligned}
B \quad &= \quad L(\frac{u+v}{2}) - L(u) \quad = \quad \int_0^{\frac{u+v}{2}} f(w)\,dw - \int_0^u f(w)\,dw \\
&= \quad \int_u^{\frac{u+v}{2}} f(w)\,dw \\
&= \quad \int_{\frac{u+v}{2}}^{v} f(w - \frac{v-u}{2})\,dw \\
&\leq \quad \int_{\frac{u+v}{2}}^{v} f(w)\,dw \\
&= \quad \int_0^v f(w)\,dw - \int_0^{\frac{u+v}{2}} f(w)\,dw \quad = \quad L(v) - L(\frac{u+v}{2}) \quad = \quad A.
\end{aligned}$$

The inequality follows from the Lorenz density being increasing. This allows the inequalities

$$\begin{aligned}
B \quad &\leq \quad A \\
\Longrightarrow L(\frac{u+v}{2}) - L(u) \quad &\leq \quad L(v) - L(\frac{u+v}{2}) \\
\Longrightarrow 2 \cdot L(\frac{u+v}{2}) \quad &\leq \quad L(u) + L(v) \\
\Longrightarrow L(\frac{u+v}{2}) \quad &\leq \quad \frac{1}{2} \cdot L(u) + \frac{1}{2} \cdot L(v).
\end{aligned}$$

◇

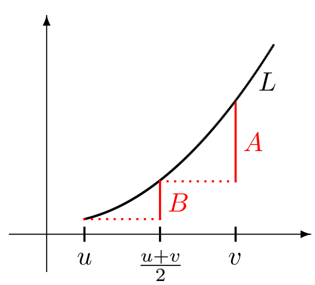

Figure 12: Increasing function increments of Lorenz curve.

The elementary properties of Lorenz curves have a few simple and some deep consequences. A simple consequence of properties 1 and 3 is that each Lorenz curve lies in the lower right triangle of the unit square. Another simple consequence is that each Lorenz curve is increasing though not necessarily strictly increasing as the Lorenz curve may attain value zero on some interval. All values different from zero are attained exactly once.

A less known, yet immediate consequence of its monotonicity is that each Lorenz curve is differentiable with the exception of, at most, a set of measure zero. This is due to a theorem by Lebesgue. A deep consequence of its convexity is that every Lorenz curve is absolutely continuous which ensures a certain compatibility between integration and differentiation even if the Lorenz curve is not everywhere differentiable.

Definition 6 *A function $\varphi : I \to I\!R$ with $I \subseteq I\!R$ is* **absolutely continuous** *if for every $\varepsilon > 0$ exists $\delta > 0$ such that for any finite collection of pairwise disjoint intervals $(a_i, b_i) \subseteq I$ with $i \in K$ and $\sum_{i\in K} b_i - a_i < \delta$ it is true that*

$$\sum_{i\in K} |\varphi(b_i) - \varphi(a_i)| < \varepsilon.$$

As a preparatory step, a result for convex functions applies to Lorenz curves

Lemma 10 *A Lorenz curve is absolutely continuous on every closed subinterval of $(0,1)$, its derivative exists almost everywhere and allows the representation $L(u) - L(w) = \int_w^u L'(v)\,dv$ for all $0 < w < u < 1$.*

Proof. Continuity of the Lorenz curve implies absolute continuity on every closed subinterval of $(0,1)$ according to [Roy, p. 109]. The derivative then exists almost everywhere [Roy, p. 105] and allows the integral representation according to [Roy, p. 106]. $\diamond$

Continuity of a Lorenz curve in zero and one imply that the integral representation also holds there so that $L(u) = \int_0^u L'(v)\,dv$ for all $0 \le u \le 1$ which can be verified by dominated convergence as in the proof of lemma 9.2.

7.3 Characterizations of Lorenz densities and Lorenz curves

Theorem 5 *("Characterization of Lorenz densities")*
Lorenz densities are exactly the left-continuous and increasing functions $f : (0,1) \to I\!R_{\ge}$ with normalization condition $\int_0^1 f(u)\,du = 1$.

Proof. "$\Rightarrow$". Lorenz densities are increasing according to lemma 1.3, left-continuous according to lemma 1.4 and the normalization holds by definition. Also, a Lorenz density attains only non-negative values. If strictly negative at some point u_0, monotonicity would lead to strictly negative values over the whole interval $(0, u_0)$ which implies $L(u_0) < 0$.

"$\Leftarrow$". The generalized inverse function f^{-1} in the sense of definition 3 can be considered as a distribution function F of a non-negative random variable according to lemma 8. The generalized inverse of this distribution function is the original function according to theorem 3.2 so that $f(u) = F^{-1}(u)$ for all $u \in (0,1)$. The Lorenz curve of X thus is

$$L(u) = \frac{\int_0^u F^{-1}(v)\,dv}{\int_0^1 F^{-1}(v)\,dv} = \frac{\int_0^u f^{-1}(v)\,dv}{\int_0^1 f^{-1}(v)\,dv} = \int_0^u f^{-1}(v)\,dv.$$

Thus, f is a Lorenz density. $\diamond$

Theorem 6 *("Characterization of Lorenz curves")*
Lorenz curves are exactly the convex and continuous functions $L : [0,1] \to [0,1]$ with boundary values $L(0) = 0$ and $L(1) = 1$.

Proof. "⇒". See lemma 9.

"⇐". Similar to the appendix B of [IK] it is argued that convexity suffices for L having a left derivative L'_- on $(0,1]$ which is increasing and left-continuous there, see [Roc, theorem 24.1]. Though, also, the right derivative exists on $[0,1)$ and is right-continuous there, the right derivative is not considered.

Note that the derivative of L exists almost everywhere and allows the integral representation as in lemma 10 and the remark following it. This is essentially due to the continuity of L. This ensures the existence of a left-continuous and increasing function f such that $f(u) = L'_-(u)$ almost everywhere and $L(u) = \int_0^u f(v)\,dv$. In particular, continuity of function L at one results in the normalization condition $1 = L(1) = \int_0^1 f(v)\,dv$.

Thus, theorem 6 becomes applicable guaranteeing the existence of a distribution function $F(x) = f^{-1}(x)$ so that $F^{-1}(u) = f(u)$ and

$$L(u) = \frac{\int_0^u F^{-1}(v)\,dv}{\int_0^u F^{-1}(v)\,dv}.$$

◇

Theorem 6 is sometimes denoted as duality between Lorenz curves and distribution functions. Theorem 5 may be considered so as well and when Lorenz curves are different, then their distribution functions are different as well. Yet the converse is not be true, but allows only little "slack".

Theorem 7 *("Affine uniqueness")*
Two random variables X_1, X_2 induce the same Lorenz curve if and only if they are positive affinely scaled which means that a real value $\alpha > 0$ exists such that $X_2 = \alpha \cdot X_1$.

Proof. "⇐". Positive affine scaling of random variables carries over to positive affine scaling of the values of generalized inverses, see section 6.1. Thus

$$L_2(u) = \frac{\int_0^u F_2^{-1}(v)\,dv}{\int_0^1 F_2^{-1}(v)\,dv} = \frac{\int_0^u 1/\alpha \cdot F_1^{-1}(v)\,dv}{\int_0^1 1/\alpha \cdot F_1^{-1}(v)\,dv} = L_1(u).$$

"⇒". It will be shown that when $EX_1 = EX_2$ and $L_1(u) = L_2(u)$ for all $u \in (0,1)$ then $F_1(x) = F_2(x)$ for all x.

Assume $F_1(x_0) \neq F_2(x_0)$ for some $x_0 \geq 0$. Then there exist $u_0 \in (0,1)$ such as $u_0 = 1/2 \cdot (F_1(x_0) + F_2(x_0))$ so that $F_1^{-1}(u_0) \neq F_2^{-1}(u_0)$. Thus, without loss of generality $F_1^{-1}(u_0) > F_2^{-1}(u_0)$. The clear distinction property of lemma 2 implies the existence of some $\delta > 0$ so that $F_1^{-1}(u) > F_2^{-1}(u)$ for all $u \in (u_0 - \delta, u_0]$, see figure 3. Integration over the given interval of positive length and equality of the expectations results in

$$\frac{\int_{u_0-\delta}^{u_0} F_1^{-1}(v)\,dv}{EX_1} > \frac{\int_{u_0-\delta}^{u_0} F_2^{-1}(v)\,dv}{EX_2}.$$

Equality of the Lorenz curves at the lower integration bound results in

$$L_1(u_0 - \delta) = \frac{\int_0^{u_0-\delta} F_1^{-1}(v)\,dv}{EX_1} = \frac{\int_0^{u_0-\delta} F_2^{-1}(v)\,dv}{EX_2} = L_2(u_0 - \delta).$$

Adding the integrals results in the strict inequality

$$L_1(u_0) = \frac{\int_0^{u_0} F_1^{-1}(v)\,dv}{EX_1} > \frac{\int_0^{u_0} F_2^{-1}(v)\,dv}{EX_2} = L_2(u_0).$$

This inequality contradicts the pointwise equality of the two Lorenz curves. ◇

The correspondences between Lorenz curves, Lorenz densities and distribution functions are wrapped up in theorem 8 and visualized in figure 13.

Theorem 8 *("Correspondences")*
The class of Lorenz curves is related to the class of Lorenz densities by a one-to-one correspondence and, also, there is a one-to-one correspondence between the class of Lorenz densities and the class of distribution functions of non-negative random variables with expectation one.

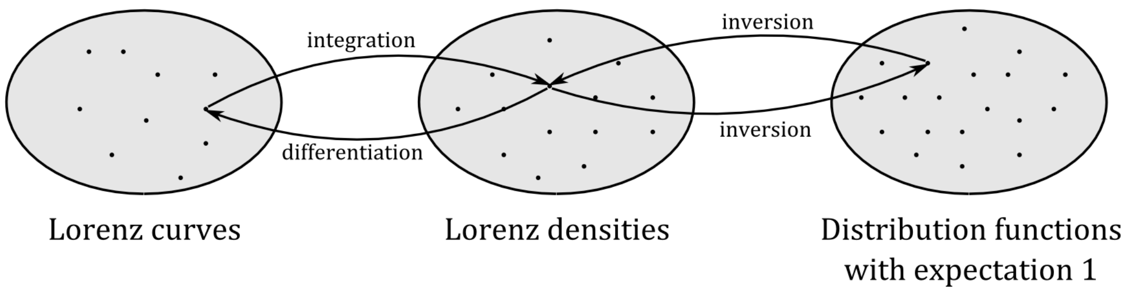

Figure 13: One-to-one correspondences between Lorenz curves, Lorenz densities and distribution functions.

7.4 Lorenz curves for all distributions

According to the Lebesgue decomposition theorem [E], each probability distribution can be expanded as a convex combination of a discrete distribution, a distribution with Lebesgue density and a distribution assigning all its mass to an uncountable set of Lebesgue measure zero; $P = \alpha \cdot P_d + \beta \cdot P_c + \gamma \cdot P_{nc}$ for $0 \leq \alpha, \beta, \gamma$ with $\alpha + \beta + \gamma = 1$. The discrete distributions and those on an uncountable set of Lebesgue measure zero are sometimes summarized as singular (with respect to the Lebesgue measure).

A Lorenz curve for a distribution over an uncountable set of Lebesgue measure zero will be discussed for the Cantor distribution. The Cantor distribution attains values in the Cantor set which is iteratively defined from the closed unit interval. The closed unit interval is denoted as $C_0 = [0, 1]$ in this context and the open middle section is removed to result in the set $C_1 = [0, 1/3] \cup [2/3, 1]$. The open middle interval of both intervals are removed to result in a union of four closed intervals $C_2 = [0, 1/9] \cup [2/9, 3/9] \cup [6/9, 7/9] \cup [8/9, 1]$. The process of removing open middle intervals is repeated to result in a doubling number of closed intervals that are retained in each iteration. The union of these intervals in each iteration n admits the iterative description

$$C_n = \frac{C_{n-1}}{3} \cup \left(\frac{2}{3} + \frac{C_{n-1}}{3}\right); \, n = 1, \ldots .$$

These sets are denoted as pre-Cantor sets and the Cantor set is defined as the intersection of all its pre-Cantor sets

$$C = \bigcap_{n=0}^{\infty} C_n.$$

As a third of each set if removed in each iteration, the contributing sets have Lebesge measure $\lambda(C_n) = (\frac{2}{3})^n$, $n = 0, 1, \ldots$ so that the Cantor set indeed has measure zero $\lambda(C) = \lim_{n\to\infty}(\frac{2}{3})^n = 0$. Every element of the Cantor has a ternary representation using only digits zero and two: $x = \sum_{i=1}^{n} \frac{x_i}{3^i} \in C$ if and only if $x_i \in \{0, 2\}$ for all i. Note that the maximum value of the Cantor set admits such a representation since $1 = (0.222\ldots)_3$. The ternary representation of numbers allows to verify that C is uncountable.

A distribution function can be assigned to every pre-Cantor set by $F_n(x) = \lambda([0, x] \cap C_n)$. These pre-Cantor distribution functions are continuous everywhere and solely consist of linear segments which have either slope zero or $(\frac{3}{2})^n$. All slanted segments have the same length while constant segments may have different lengths. A pre-Cantor distribution function assigns equal probability to all equally sized subsets of the underlying pre-Cantor set.

The limit over the pre-Cantor distribution functions is the Cantor distribution function

$$F(x) = \lim_{n\to\infty} F_n(x) = \lambda([0, x] \cap C).$$

The Cantor distribution function is continuous everywhere, differentiable almost everywhere with derivative zero almost everywhere but not absolutely continuous. One pre-Cantor distribution function and the Cantor distribution function, also called devil's staircase, are sketched in figure 14.

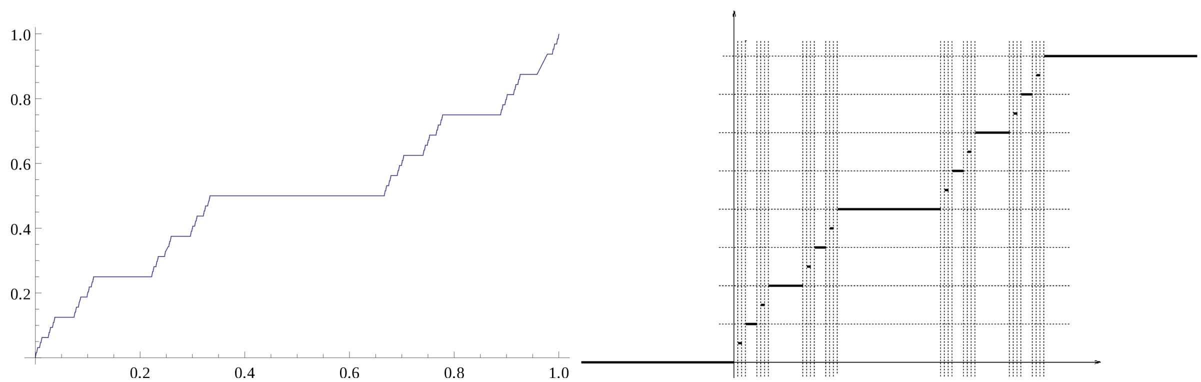

Figure 14: pre-Cantor distribution function for $n = 10$ (left) and constant segments of the Cantor distribution function (right).

Due to symmetry, the random variables X_n and X with pre-Cantor distributions and Cantor distribution, respectively, all have the same expectation $EX_n = EX = 1/2$, $n = 0, 1, \ldots$. The Lorenz curves of the first three pre-Cantor distributions are

$$\begin{aligned}
L_0(u) &= u^2 \\
L_1(u) &= \begin{cases} \frac{2}{3}u^2, & \text{if } 0 \le u \le 0.5 \\ \frac{2}{3}u^2 + \frac{2}{3}u - \frac{1}{3}, & \text{if } 0.5 \le u \le 1 \end{cases} \\
L_2(u) &= \begin{cases} \frac{4}{9}u^2, & \text{if } 0 \le u \le 0.25 \\ \frac{4}{9}u^2 + \frac{2}{9}u - \frac{1}{18}, & \text{if } 0.25 \le u \le 0.5 \\ \frac{4}{9}u^2 + \frac{8}{9}u - \frac{7}{18}, & \text{if } 0.5 \le u \le 0.75 \\ \frac{4}{9}u^2 + \frac{10}{9}u - \frac{5}{9}, & \text{if } 0.75 \le u \le 1 \end{cases}.
\end{aligned}$$

These Lorenz curves are shown in figure 15. It is conjectured that the pre-Cantor distribution are increasingly unequal in the sense $L_{n+1}(u) \le L_n(u)$ for all $u \in [0, 1]$ and $n = 0, 1, \ldots$.

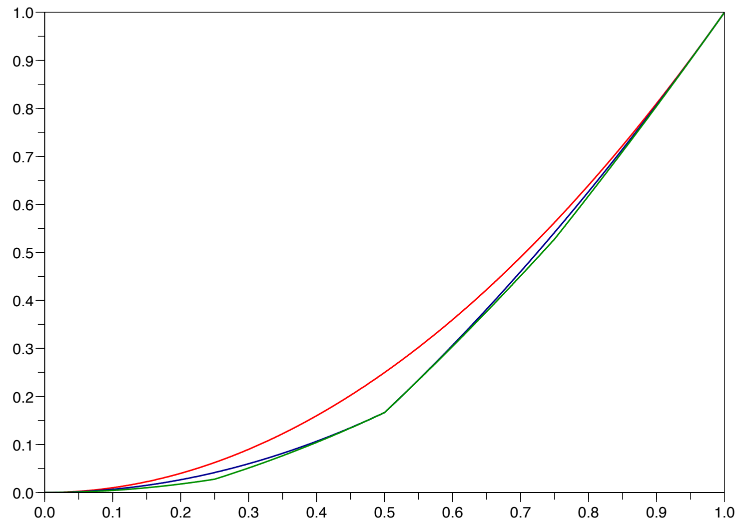

Figure 15: Lorenz curves of the pre-Cantor distributions for $n = 0$ (top), $n = 1$ (middle) and $n = 2$ (bottom).

When a pre-Cantor distribution function has a constant segment at level $k/2^n$ for some $k \in \{1, \ldots, 2^n - 1\}$, then the pre-Cantor Lorenz curve is not smooth at this point. More important, the pre-Cantor Lorenz curves do not change values at this point any more when the approximation index increases.

Lemma 11 *When $F_n(x) = k/2^n$ for different values x, some n and some $k \in \{1, \ldots, 2^n - 1\}$, then $L_n(\frac{k}{2^n}) = L_{n+1}(\frac{k}{2^n}) = \ldots = L(\frac{k}{2^n})$.*

Proof (Sketch). The result follows from $\int_0^{k/2^n} F_n^{-1}(u)\, du = \int_0^{k/2^n} F_{n+1}^{-1}(u)\, du$ which can be seen from figure 16. $\diamond$

In principle, as a consequence of lemma 11, the Lorenz curve of the Cantor distribution can be computed explicitly at inverse powers of two and multiples thereof. Sample values are

$$L(\frac{1}{4}) = \frac{1}{36},\ L(\frac{1}{2}) = \frac{1}{6},\ L(\frac{3}{4}) = \frac{19}{36}.$$

7.5 Lorenz curves from Lorenz curves

The product of any two Lorenz curves is a Lorenz curve so that, in particular, $L(u) \cdot u$ is a Lorenz curve whenever $L(u)$ is. Also, every convex combination of two or more Lorenz curves is a Lorenz curve. In addition, reflection along the minor diagonal $w = 1 - u$ for $0 \le u \le 1$ leads from a Lorenz curve to its reflected Lorenz curve $L^{ref}(u) = 1 - L^{-1}(1 - u)$.

A more subtle construction is to consider a Lorenz curve as a distribution function and to construct its own induced Lorenz curve. This process is repeatable and can be shown to have a fixed point which is the polynomial Lorenz curve $L(u) = u^g$ with golden section parameter $g = \frac{1+\sqrt{5}}{2}$.

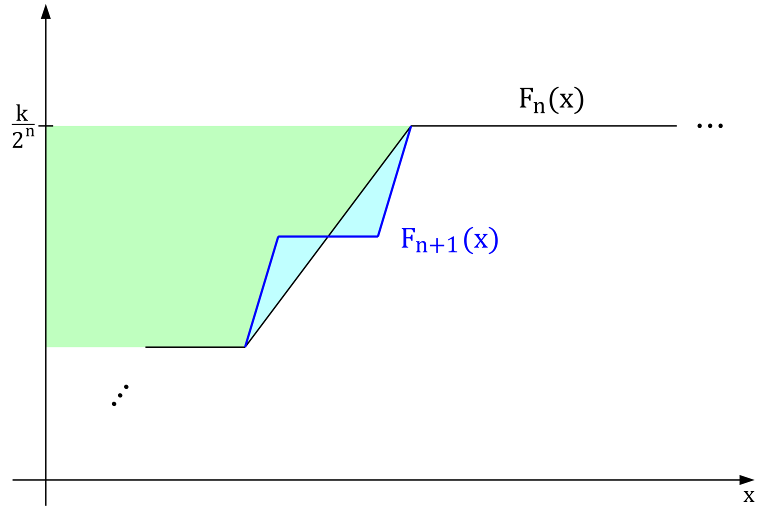

Figure 16: The integral values over the interval $[0, k/2^n]$ are equal for both pre-Cantor distribution functions because of symmetry. (The meaning of the integral in terms of Lorenz curves is illustrated by figure 10).

7.6 Lorenz curves induced by finite and infinite variance distributions

Least unevenness of an income distribution is obtained by the Lorenz curve lying on the main diagonal which is obtained for the one-point distribution: all incomes are equal. This distribution has zero variance so that, trivially, it has finite variance. On the other hand, Lorenz curves of finite variance distributions can approximate the most uneven limiting case. Even more, the variances of such approximations may be uniformly bounded which means that parametric distributions exist with

1. $Var_\epsilon X \leq M$ for some constant $M > 0$ and for all $\varepsilon \in (0, 1)$ and
2. $\lim_{\epsilon \to 0} L_\epsilon(u) = 0$ for all $u \in [0, 1)$.

This means that variability as expressed by Lorenz curves is unrelated to inducing distributions having finite or infinite variance. Two-point distributions with uniformly bounded variances suffice for approximating the most uneven limiting case. Therefore, support points and probabilities are considered as

$$x_1 = x_1(\varepsilon) = \varepsilon^2 \text{ with } p_1 = p_1(\varepsilon) = 1 - \varepsilon$$
$$x_2 = x_2(\varepsilon) = \sqrt{\varepsilon} \text{ with } p_2 = p_2(\varepsilon) = \varepsilon$$

for $\varepsilon \in (0, 1)$. Then $0 < x_1(\varepsilon) < x_2(\varepsilon) < 1$ for all considered ε. So, all realizations of a random variable X with these support point pairs lie between zero and one and hence, all realizations of X^2 do as well. Thus $Var_\epsilon X = E_\epsilon X^2 - (E_\epsilon X)^2 \leq E_\epsilon X^2 \leq 1$ for all $\varepsilon \in (0, 1)$.

Also, the points where the Lorenz curves of the two-point distributions change slope converge to the right lower corner of the unit interval. The change points have the coordinates, see figure 11

$$P(\varepsilon) = \left(\begin{array}{c} p_1(\varepsilon) \\ \frac{p_1(\varepsilon)\, x_1(\varepsilon)}{p_1(\varepsilon)\, x_1(\varepsilon) + p_2(\varepsilon)\, x_2(\varepsilon)} \end{array} \right).$$

Now, $p_1(\varepsilon) = 1 - \varepsilon \to 1$ $(\epsilon \to 0)$ and

$$\frac{p_1(\varepsilon)\,x_1(\varepsilon)}{p_1(\varepsilon)\,x_1(\varepsilon)+p_2(\varepsilon)\,x_2(\varepsilon)} = \frac{(1-\varepsilon)\,\varepsilon^2}{(1-\varepsilon)\,\varepsilon^2+\varepsilon\,\sqrt{\varepsilon}}$$
$$= \frac{(1-\varepsilon)\,\sqrt{\varepsilon}}{(1-\varepsilon)\,\sqrt{\varepsilon}+1} \to 0 \;\; (\epsilon \to 0).$$

This completes the example, see figure 17. It is slightly remarkable that he Lorenz curves indicate ever larger unevenness though both support points converge to zero so that the two-point distributions converge to a one-point distribution that has no unevenness at all. Ever increasing unevenness of the Lorenz curves is partially explained by the ratio of the larger income level over the smaller tending to infinity: $x_2(\varepsilon)/x_1(\varepsilon) = \sqrt{\varepsilon}/\varepsilon^2 = 1/(\sqrt{\varepsilon}\cdot\varepsilon) \to \infty$ $(\epsilon \to 0)$. This drags the change point downwards. Additionally, the probability of the smaller income level tends to one which drags the change point to the right boundary.

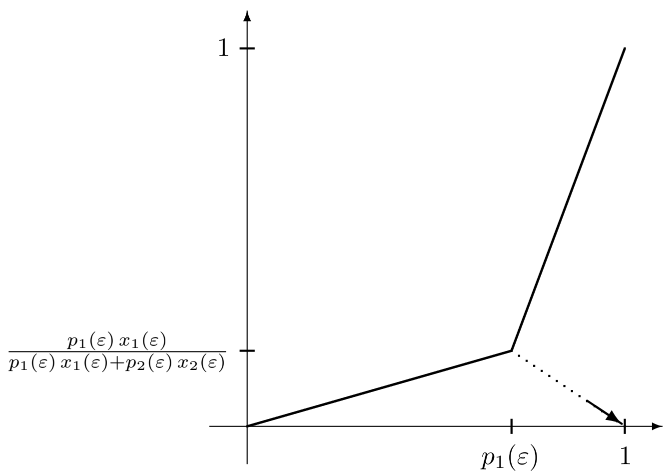

Figure 17: Lorenz curves induced by two-point distributions so that the change points $P(\varepsilon)$ converge to the lower right corner of the unit square for $\varepsilon \to 0$.

Is the converse also true which means that the main diagonal can be approximated by Lorenz curves of infinite variance distributions?

8 Equity calculus

A notion of relative poverty considers an individual to be poor if his income falls short of some given fraction of the average of all incomes. The fraction may by 50% or else. This perspective is now extended to all income levels: every income is supposed to be proportional to the average of all larger incomes. The proportionality factor is denoted as equity parameter ε and it ranges throughout $(0, 1]$. The following calculus is suited for normative and descriptive views alike and it is governed by the given proportionality assumption and variations thereof.

As individual incomes are proportional to the Lorenz density, the desired proportionality law can be stated in terms of a Lorenz curve and its derivative. More precisely, differentiable Lorenz curves with the desired proportionality law satisfy the linear inhomogenous differential equation

$$L'(u) = \varepsilon \cdot \frac{1-L(u)}{1-u}.$$

The solution is

$$L(u) = 1 - (1-u)^{\varepsilon}.$$

This Lorenz curve belongs to a Pareto distribution with single parameter ε. In general, the Pareto distribution is two-parametric with shape parameter $\alpha > 0$ and range parameter $x_m > 0$. Density and distribution function are

$$\begin{aligned} f(x) &= \begin{cases} \alpha \cdot (1/x_m)^{-\alpha} \cdot x^{-(\alpha+1)}, & \text{if } x \geq x_m \\ 0, & \text{if } x < x_m \end{cases} \\ F(x) &= \begin{cases} 1 - (\frac{x}{x_m})^{-\alpha}, & \text{if } x \geq x_m \\ 0, & \text{if } x < x_m \end{cases} . \end{aligned}$$

The Pareto distribution is a so-called heavy-tail distribution so that expectations, variances and higher moments may not be finite. The expectation is only finite for shape parameters $\alpha > 1$ and then has the value $x_m\alpha/(\alpha-1)$. The variance is only finite for larger shape parameters $\alpha > 2$ and then has the value $x_m^2\alpha/((\alpha-1)^2(\alpha-2))$. Interestingly, the Lorenz curve of the two-parametric Pareto distribution is independent from the range parameter and only depends on the shape parameter $L(u) = 1-(1-u)^{-1/\alpha+1}$ with $\alpha > 1$. All two-parametric Pareto distributions with same ratio of minimum income over mean income have the same Lorenz curve which is in accordance with theorem 7. The ratio equals $\frac{x_m}{x_m\,\alpha}(\alpha-1) = -1/\alpha + 1$.

Within the equity calculus, the two parameters of the Pareto distribution are always coupled in the following way: the range parameter is set equal to the equity parameter which indicates the minimum income; $x_m = \varepsilon$. The shape parameter is then adjusted so that the expected income becomes one; $\frac{\varepsilon\cdot\alpha}{\alpha-1} = 1 \iff \alpha = 1/(1-\varepsilon) \iff \epsilon = -1/\alpha + 1$. A one-parametric Pareto distribution is thus singled out from the two-parametric Pareto distributions by the ratio of minimum and mean income being equal to the equity parameter. Though the one-parametric Pareto distributions form a proper subset of the two-parametric Pareto distributions with finite mean, they have identical sets of Lorenz curves.

The two-parametric Pareto distribution with shape parameter greater one is the only so that each income is proportional to the average of all larger incomes. This has been shown by a differential equation for the distribution function [C, p. 155]. The same result is obtained by the differential equation for the one-parametric Lorenz curve only when the equity parameter does not exceed one. The reason is that these Lorenz curves are representative for all two-parametric Pareto distributions with shape parameter greater one.

From now on, the Pareto distributions are understood to be one-parametric. For equity parameters strictly smaller than one, the income density and the distribution function are, respectively

$$\begin{aligned} f(x) &= \begin{cases} \frac{1}{1-\varepsilon} \cdot \varepsilon^{1/(1-\varepsilon)} \cdot x^{1/(\varepsilon-1)-1}, & \text{if } x \geq \varepsilon \\ 0, & \text{if } x < \varepsilon \end{cases} \\ F(x) &= \begin{cases} 1 - (\frac{x}{\varepsilon})^{1/(\varepsilon-1)}, & \text{if } x \geq \varepsilon \\ 0, & \text{if } x < \varepsilon \end{cases} . \end{aligned}$$

This one-parametric Pareto distribution has finite mean value one for all equity parameters $0 < \varepsilon < 1$ and it has finite variance $(1-\varepsilon)^2/(2\varepsilon-1))$ only for equity parameters $0.5 < \varepsilon < 1$. The case $\varepsilon = 1$ is given separately by a single point distribution. A Pareto density and a distribution function are sketched in figure 18.

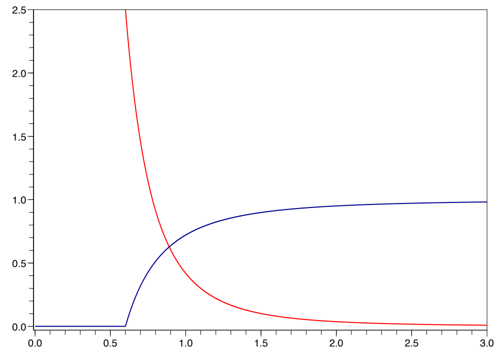

Figure 18: Density (decreasing) and distribution function (increasing) of the Pareto distribution with equity parameter $\varepsilon = 0.6$.

8.1 Geometrical interpretation

The differential equation has a geometrical interpretation in terms of tangent slopes of a Lorenz curve, see figure 19. The tangent slope is bounded by two secant slopes as

$$\frac{L(u)}{u} \leq L'(u) \leq \frac{1 - L(u)}{1 - u}.$$

Intending the upper inequality to become an equation motivates to decrease the upper bound. This is obtained by introducing a positive factor $\varepsilon \leq 1$ which results in the foregoing differential equation. Similarly, intending the lower inequality to become an equation motivates to increase the lower bound which is obtained by introducing a factor $M \geq 1$ which results in the differential equation

$$\frac{L(u)}{u} \cdot M = L'(u).$$

This linear homogenous differential equation has the polynomial solution $L(u) = u^M$. Without knowing this solution and without knowing the Pareto curve being a solution of the other differential equation, the solutions can be verified to be Lorenz curves just from a certain value property.

Lemma 12 *Every solution with $L(u) \leq 1$ of the upper differential equation is increasing and convex and every solution with $L(u) \geq 0$ of the lower differential equation is increasing and convex.*

Proof. The function $\varphi(u) = \frac{1-L(u)}{1-u}$ for $u \in [0, 1)$ is non-negative. It is also increasing which is equivalent to

$$\begin{aligned} \varphi'(u) &= \frac{-L'(u) \cdot (1-u) + 1 - L(u)}{(1-u)^2} \geq 0 \\ \iff \frac{1 - L(u)}{1 - u} &\geq L'(u). \end{aligned}$$

A solution of the upper differential equation with values not exceeding one satisfies $L'(u) = \varepsilon \cdot \frac{1-L(u)}{1-u} \leq \frac{1-L(u)}{1-u}$. Thus, $L'(u)$ is non-negative and increasing which implies that $L(u)$ is increasing and convex. The argument for the lower differential equation works similarly. ◇

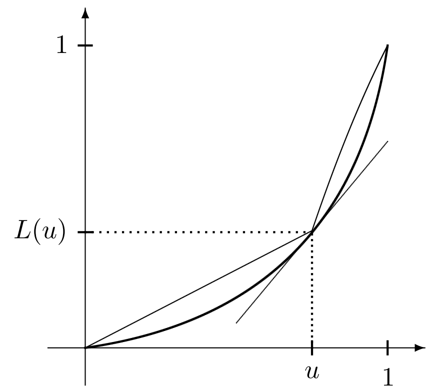

Figure 19: The tangent slope of any Lorenz curve at any interior point is sandwiched by secant slopes as indicated.

In principle, the tangent inequalities could be forced to become equations by addition rather than multiplication. This would lead to the differential equations $\frac{L(u)}{u} + A = L'(u)$ with $A \geq 0$ and $L'(u) = \frac{1-L(u)}{1-u} - a$ with $0 \leq a < 1$. However, these do not result in Lorenz curves with continuos derivatives except for the trivial case $a = A = 0$. This can be seen as follows for Lorenz curves which are continuously differentiable at zero: $L'(0) = \lim_{u\to 0} L'(u) = \lim_{u\to 0} \frac{L(u)}{u} + A = \lim_{u\to 0} \frac{L(u)-L(0)}{u-0} + A = L'(0) + A$. Thus, these differential equations are not pursued any further.

8.2 The main differential equations of the equity calculus

The equations $\frac{L(u)}{u} \cdot M = L'(u) = \varepsilon \cdot \frac{1-L(u)}{1-u}$ with $M \geq 1 \geq \varepsilon > 0$ have no common solution except the Egalitarian Lorenz curve $L(u) = u$. Yet we note that the given equations are three and that the polynomial curve and the Pareto curve result from breaking the symmetry to the left and to the right of the three equations, respectively. Crossing out the differential in the middle also results in an equation for Lorenz curves, namely in the functional equation

$$\frac{L(u)}{u} \cdot M = \varepsilon \cdot \frac{1-L(u)}{1-u}.$$

Few simple algebraic manipulations show that the solution is the one-parametric function $L(u) = \frac{u}{M/\varepsilon - (M/\varepsilon - 1)\,u} = \frac{u}{\bar{M} - (\bar{M}-1)\,u}$ with $\bar{M} = M/\varepsilon \geq 1$. It is verified by insertion that this function, which was not derived from differentials, satisfies even two homogenous differential equations

$$(\frac{L(u)}{u})^2 \cdot \bar{M} = L'(u) = \frac{1}{\bar{M}} \cdot (\frac{1-L(u)}{1-u})^2.$$

An income distribution with this Lorenz curve asserts that each income is proportional to the squared average of all larger incomes and, simultaneously, of all smaller incomes! Also, each income equals the average of all larger incomes with proportionality 'factor' being the average of all smaller incomes and vice versa. This can be seen from the Lorenz curve also satisfying the symmetric differential equation of the Bernoulli type

$$L'(u) = \frac{L(u)}{u} \cdot \frac{1-L(u)}{1-u}.$$

Generalizations of the differential equations with squared average incomes can be obtained by raising the average incomes to even higher powers. All in all, this results in the sets of lower and upper differential equations

$$(\frac{L(u)}{u})^n \cdot M = L'(u) \text{ and } L'(u) = m \cdot (\frac{1-L(u)}{1-u})^n$$

for $n = 1, 2, 3, \ldots$ and $M \geq 1 \geq m > 0$. All these differential equations are solvable by Lorenz curves in closed form.

8.3 Closed form solutions

The lower differential equations have the solutions

$$L_l(u) = \begin{cases} u^M, & \text{if } n = 1 \\ \frac{u}{M-(M-1)\,u}, & \text{if } n = 2 \\ \frac{u}{\sqrt[n-1]{M-(M-1)\,u^{n-1}}}, & \text{if } n = 3, 4, \ldots \end{cases}$$

These solutions can be found with the assistance of the on-line symbolic ODE solver of the Wolfram alpha [Wα]. The upper differential equations have the solutions

$$L_u(u) = \begin{cases} 1-(1-u)^m, & \text{if } n = 1 \\ 1-\frac{1-u}{m+(1-m)\,(1-u)}, & \text{if } n = 2 \\ 1-\sqrt[n-1]{1/m} \cdot \frac{1-u}{\sqrt[n-1]{1+(1/m-1)\,(1-u)^{n-1}}}, & \text{if } n = 3, 4, \ldots \end{cases}$$

Though solutions of the upper differential equations can be given 'directly', the solution formulas are quite intricate and their length increases in the power n. The stated expressions were obtained by the function substitution $K(u) = 1 - L(1-u)$. Then $L(u) = 1 - K(1-u)$ and $L'(u) = K'(1-u)$. This allows to rewrite the differential equations as

$$K'(1-u) = L'(u) = m \cdot (\frac{1-L(u)}{1-u})^n = m \cdot (\frac{K(1-u)}{1-u})^n .$$

The variable substitution $w = 1 - u$ results in the lower differential equations except that the proportionality parameter does not exceed one

$$K'(w) = m \cdot (\frac{K(w)}{w})^n .$$

Using again the symbolic ODE solver of the Wolfram alpha and backsubstitution lead to the upper solution functions as stated. Note that the intermediate function $K(u)$ need not and will not be a Lorenz curve.

The case $n = 2$ results from the formulas for $n \geq 3$ for both lower and upper Lorenz curves. Yet, stating this case separately better illustrates the function type. Interestingly, for $n = 2$ and $M = 1/m$ lower and upper Lorenz curves are identical since they satisfy the lower as well as the upper differential equation, see above. Sample curves are shown in figures 20 and 21.

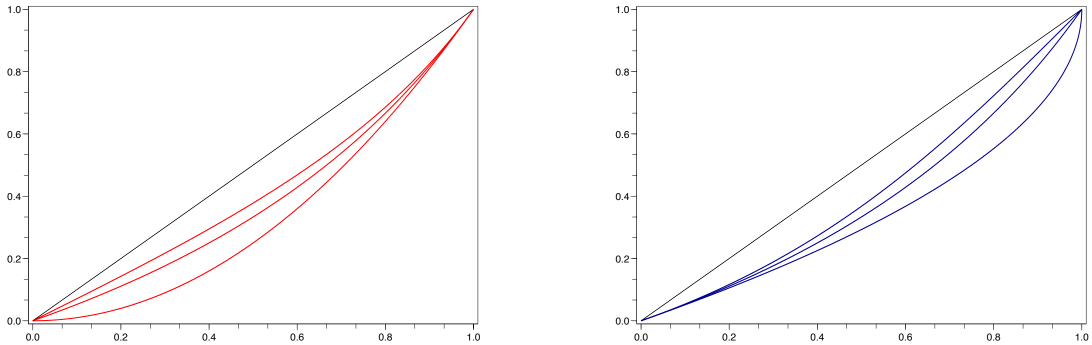

Figure 20: Lower Lorenz curves with $M = 2$ (left) and upper Lorenz curves with $m = 1/2$ (right) for powers $n = 1$ (bottom), $n = 2$ (middle) and $n = 3$ (top) in both cases. The upper Lorenz curve for $n = 1$ is a Pareto curve.

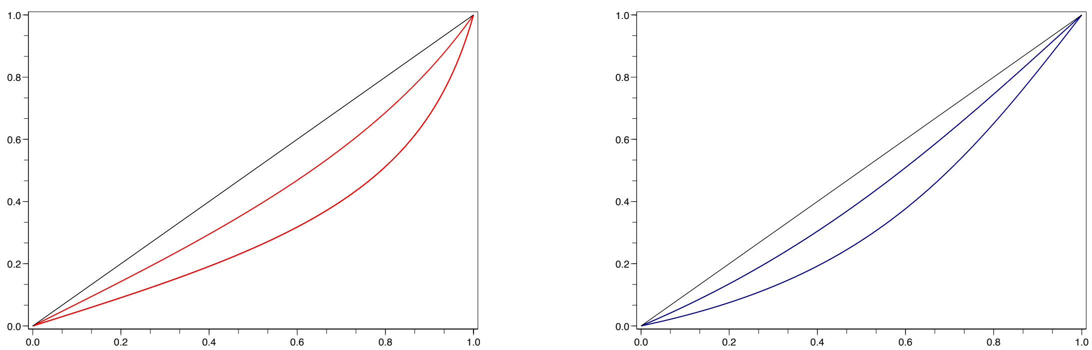

Figure 21: Lower Lorenz curves (left) with $M = 2$ (top) and $M = 5$ (bottom) and upper Lorenz curves (right) with $m = .6$ (top) and $m = .3$ (bottom). The power is set to $n = 3$ for all four curves.

In the purely algebraic approach, considering powers of average incomes does not lead to other types of Lorenz curves than just considering power one. This means that the functional equation $(\frac{L(u)}{u})^n = \mu \cdot (\frac{1-L(u)}{1-u})^n$ with $\mu \geq 1$ has no more solutions for $n \geq 2$ than for $n = 1$; only the parameter μ has to be replaced by $\sqrt[n]{\mu}$.

8.4 Empirics

Parametric Lorenz curves can be fitted to finite collections of support points by least squares minimization. For an upper Lorenz curve with given power n and given support point collection $(u_i, y_i)_{i \in C}$ this amounts to solving the problem

$$\min_{m \in (0,1]} \sum_{i \in C} \Big(L_u(u_i) - y_i \Big)^2.$$

Best least squares fits for the upper Lorenz curves for indices from one through ten have been computed for German income data, see figure 22. The curve for power two resulted in the best overall fit. The same observation can be made for other nations including the US. Noteworthy, the underlying income data span across the whole income range. The well known high fitting quality of the Pareto distribution for top levels of real income data [DS, p. 170] is outweighed by poor fitting elsewhere.

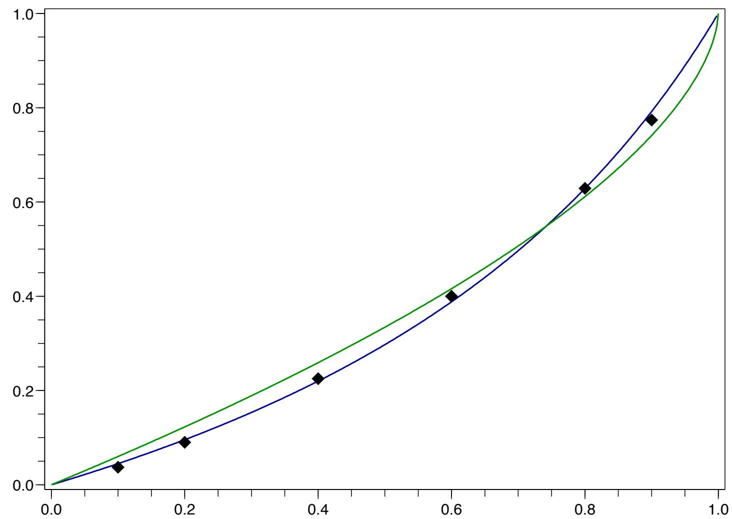

Figure 22: Best fit Pareto curve (green) and better fitting best fit upper curve for power $n = 2$ (blue) for German income data as provided by the Worldbank. Best fit proportionality parameters are $m = 0.587$ ('equity parameter') for the Pareto curve and $m = 0.423$ for the cure with power two.

8.5 Further differential equations of the equity calculus

8.5.1 Fractured exponents

The integer exponents of the income averages can be generalized to genuine rational and genuine real numbers. However, solutions are then intricate to obtain except in special cases. One such case is the lower fractured differential equation $L'(u) = M \cdot (L(u)/u)^{1.5}$ for $M \geq 1$ with Lorenz curve solutions

$$L(u) = \frac{u}{((M-1)\sqrt{u} - M)^2}.$$

Using function substitution and variable substitution as in section 8.3, these solutions carry over to solutions of the upper fractured differential equation $L'(u) = m \cdot ((1 - L(u))/(1-u))^{1.5}$ for $0 < m \leq 1$. Resulting Lorenz curves are

$$L(u) = 1 - \frac{1-u}{((m-1)\sqrt{1-u} - m)^2}.$$

8.5.2 Proportionality functions

All differential equations of the equity calculus, which were considered so far, have constant coefficients. This means that their proportionality factors are constants. These can be relaxed to functions and differential equation solutions can be given in closed form for certain proportionality functions. Only few explicit examples are given. The lower differential equations $L'(u) = (1+u^n) \cdot \frac{L(u)}{u}$ with $n = 1, 2, \ldots$ have the solutions

$$L(u) = u \cdot e^{\frac{u^n}{n} - \frac{1}{n}}$$

and the upper differential equations $L'(u) = u^n \cdot \frac{1-L(u)}{1-u}$ with $n = 1, 2, \ldots$ have the solutions

$$L(u) = 1 - (1-u) \cdot e^{\frac{u^n}{n} + \frac{u^{n-1}}{n-1} + \ldots + u}.$$

Note that polynomial Lorenz curves were used as proportionality functions for the upper differential equations. Even Pareto Lorenz curves can be inserted. An example is $L'(u) = \left(1 - (1-u)^{0.5}\right) \cdot \frac{1-L(u)}{1-u}$ with solution

$$L(u) = 1 - (1-u) \cdot e^{2 - 2\cdot\sqrt{1-u}}.$$

The curve is shown in figure 23. All examples fall into the pattern of generating Lorenz curves from given

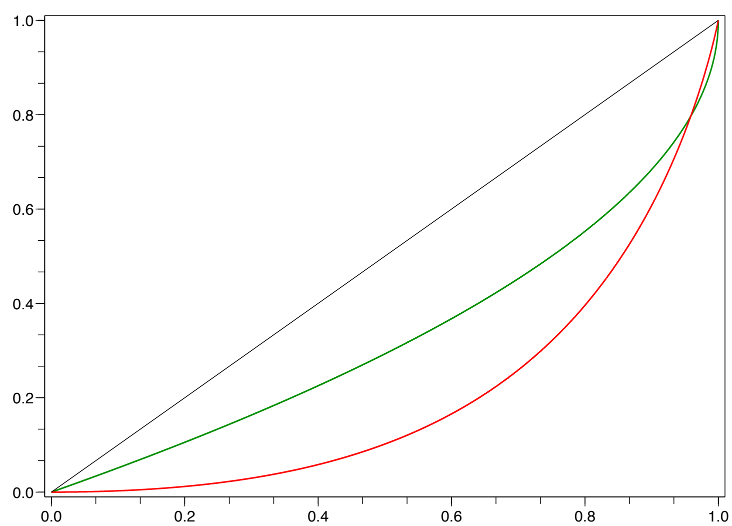

Figure 23: Pareto curve (green) with equity parameter 0.5 and Lorenz curve (red) using it as proportionality function.

Lorenz curves which serve as proportionality functions in differential equations.

Theorem 9 *Let $\bar{L}(u)$ be some differentiable Lorenz curve. Then*

1. *every solution of $L'(u) = \bar{L}(u) \cdot \frac{1-L(u)}{1-u}$ with $L(0) = 0$, $L(1) = 1$ and $L(u) \leq 1$ and*
2. *every solution of $L'(u) = (1 + \bar{L}(u)) \cdot \frac{L(u)}{u}$ with $L(0) = 0$, $L(1) = 1$ and $L(u) \geq 0$*

is a Lorenz curve.

Proof. Solution functions being increasing and convex is verified similar to lemma 12. ⋄

Proportionality functions need not be increasing. For example, the lower differential equation $L'(u) = (2 - 0.1\,u) \cdot L(u)/u$ has the Lorenz curve solution $L(u) = u^2 \cdot exp(0.1 - 0.1\,u)$. Verifying convexity of solution functions can be complicated for decreasing proportionality functions.

8.5.3 Slack functions

The slack between the secant slopes and the tangent slope can be filled-in by certain additive functions, called slack functions though not by constants as discussed towards the end of section 8.1. The lower differential equations with polynomial slack functions $L'(u) = L(u)/u + u^n$ for $n = 1, 2, \ldots$ have the Lorenz curve solutions

$$L(u) = \frac{n-1}{n} \cdot u + \frac{1}{n} \cdot u^{n+1}.$$

The upper differential equations with polynomial slack functions $L'(u) = (1 - L(u))/(1 - u) - (1 - u)^n$ for $n = 1, 2, \ldots$ have the Lorenz curve solutions

$$L(u) = 1 - \frac{n+1}{n} \cdot (1 - u) + \frac{1}{n} \cdot (1 - u)^{n+1}.$$

8.5.4 Averages over other income ranges

Individual incomes can be related to subsets of all larger incomes instead of *all* larger incomes and they can be related to subsets of all smaller incomes instead of *all* smaller incomes. For example, each of these ranges can be cut into half so that an income distribution is characterized by all incomes being proportional to the average over the upper 50% of all smaller incomes. This results in the differential equation

$$L'(u) = M \cdot \frac{L(u) - L(u/2)}{u/2}$$

for $M \geq 1$. The analog holds for all incomes being proportional to the average over the lower 50% of all larger incomes. This results in the differential equation

$$L'(u) = m \cdot \frac{L((1+u)/2) - L(u)}{(1-u)/2}$$

for $m \leq 1$. The geometry for the subset ranges is sketched in figure 24 and the corresponding differential equations are not of any standard type. At best, it is difficult to obtain solutions in closed form. They remain to be found or to be numerically approximated.

8.5.5 Limitations

It seems that a proportionality factor or a proportionality function is indispensable. The inequalities for the tangent slope of a Lorenz curve can be forced to become equations by transforming the secant slopes to larger or smaller values without using any kind of proportionalities. Since $0 \leq \frac{L(u)}{u} \leq 1$, any increasing transformation with $T(x) \geq x$ for $x \in [0, 1]$ can be used. Feasible examples are $T(x) = \sqrt[n]{x}$ for $n = 2, 3, \ldots$ leading to the Bernoulli differential equations $\sqrt[n]{L(u)/u} = L'(u)$. Similarly, the upper tangent inequality leads to the Bernoulli differential equations $L'(u) = \sqrt[n]{(1 - L(u))/(1 - u)}$. However, the only Lorenz curve that satisfies any of these differential equations is the Egalitarian curve.

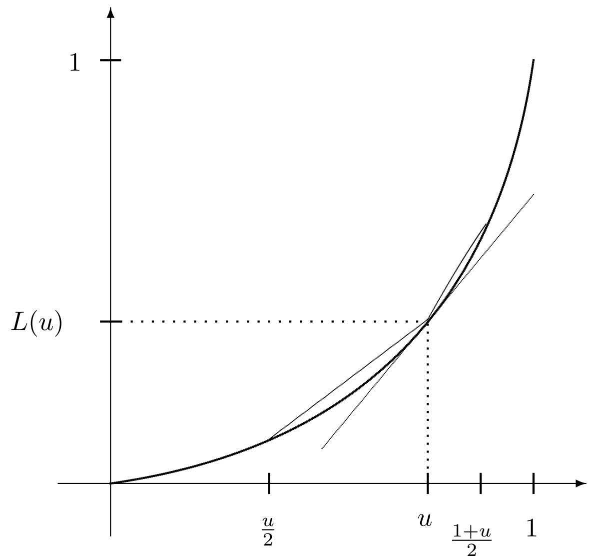

Figure 24: The tangent slope of a Lorenz curve at any interior point is tighter sandwiched by secant slopes than in figure 19.

8.6 System of proportionality laws

As a summary, a system of proportionality laws with their closed form solutions is given in table 2. The system contains all previously given cases but, ultimately, cannot be claimed to be complete in any reasonable sense. Though also based on differential equations, the given system differs from the Burr system [Ri] by directly referring to Lorenz curves so that all resulting functions are, in principle, suitable for describing income distributions.

References

[Ba] Bauer, H., "Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie", 2nd ed., deGruyter, Berlin, 1974.

[C] Cowell, F.A., "Measuring inequality", LSE Handbooks in Economics Series, London, 2009, darp.lse.ac.uk/papersDB/Cowell_measuringinequality3.pdf with additional data at darp.lse.ac.uk/ Frankweb/MI3.htm.

[DS] Diamond, P, Saez, E., "The case for a progresive tax: from basic research to policy recommendations", Journal of Economic Perspectives 25, 2011, p. 165-190.

[E] Elstrodt, J., "Maß- und Integrationstheorie", 6. Auflage, Springer, Berlin, 2009.

[G] Gastwirth, J.L., "A general definition of the Lorenz curve", Econometrica 39, 1971, p. 1037-39.

[IK] Iritani, J., Kuga, K., "Duality between the Lorenz curves and the income distribution functions", The Economic Studies Quarterly 34, 1983, p. 9-21.

[Ri] Rinne, H., "Taschenbuch der Statistik", Harri Deutsch Verlag, Frankfurt, 2008.

[Roc] Rockafeller, R.T., "Convex analysis", 2nd printing, Princeton University Press, Princeton, 1972.

[Roy] Royden, H.L., "Real analysis", 2nd ed., Macmillan, London, 1968.

[Th] Thistle, P.D., "Duality between generalized Lorenz curves and distribution functions", The Economic Studies Quarterly 40, 1989, p. 183-187.

Proportionality law	Parameter	Solution	Remark
$L'(u) = \varepsilon \cdot \frac{1-L(u)}{1-u}$	$0 < \varepsilon \le 1$	$L(u) = 1 - (1-u)^{\varepsilon}$	Pareto distribution
$L'(u) = m \cdot (\frac{1-L(u)}{1-u})^2$	$0 < m \le 1$	$L(u) = 1 - \frac{1-u}{m+(1-m)\,(1-u)}$	$L'(u) = \frac{L(u)}{u} \cdot \frac{1-L(u)}{1-u}$
$L'(u) = m \cdot (\frac{1-L(u)}{1-u})^n$	$0 < m \le 1$	$L(u) = 1 - \sqrt[n-1]{1/m} \cdot \frac{1-u}{\sqrt[n-1]{1+(1/m-1)\,(1-u)^{n-1}}}$	$n = 3, 4, \ldots$
$L'(u) = m \cdot (\frac{1-L(u)}{1-u})^{1.5}$	$0 < m \le 1$	$L(u) = 1 - \frac{1-u}{((m-1)\,\sqrt{1-u}-m)^2}$	fractured exponent
$L'(u) = M \cdot \frac{L(u)}{u}$	$M \ge 1$	$L(u) = u^M$	polynomial distr.
$L'(u) = M \cdot (\frac{L(u)}{u})^2$	$M \ge 1$	$L(u) = \frac{u}{M-(M-1)\,u}$	$L'(u) = \frac{L(u)}{u} \cdot \frac{1-L(u)}{1-u}$
$L'(u) = M \cdot (\frac{L(u)}{u})^n$	$M \ge 1$	$L(u) = \frac{u}{\sqrt[n-1]{M-(M-1)\,u^{n-1}}}$	$n = 3, 4, \ldots$
$L'(u) = M \cdot (\frac{L(u)}{u})^{1.5}$	$M \ge 1$	$L(u) = \frac{u}{((M-1)\,\sqrt{u}-M)^2}$	fractured exponent
$L'(u) = (1+u^n) \cdot \frac{L(u)}{u}$	–	$L(u) = u \cdot e^{\frac{u^n}{n} - \frac{1}{n}}$	$n = 1, 2, \ldots$
$L'(u) = u^n \cdot \frac{1-L(u)}{1-u}$	–	$L(u) = 1 - (1-u) \cdot e^{\frac{u^n}{n} + \frac{u^{n-1}}{n-1} + \ldots + u}$	$n = 1, 2, \ldots$
$L'(u) = \big(1-(1-u)^{0.5}\big) \cdot \frac{1-L(u)}{1-u}$	–	$L(u) = 1 - (1-u) \cdot e^{2-2\cdot\sqrt{1-u}}$	–
$L'(u) = (2-0.1\,u) \cdot L(u)/u$	–	$L(u) = u^2 \cdot exp(0.1-0.1\,u)$	–
$L'(u) = L(u)/u + u^n$	–	$L(u) = \frac{n-1}{n} \cdot u + \frac{1}{n} \cdot u^{n+1}$	$n = 1, 2, \ldots$
$L'(u) = (1-L(u))/(1-u) - (1-u)^n$	–	$L(u) = 1 - \frac{n+1}{n} \cdot (1-u) + \frac{1}{n} \cdot (1-u)^{n+1}$	$n = 1, 2, \ldots$

Table 2: System of selected, solvable proportionality laws and their explicit solutions.

[Wα] Wolfram alpha, "Computational knowledge engine", www.wolframalpha.com.

[Sci] Scilab, "The free platform for numerical computation", www.scilab.org.