

# Alternate definitions of Gini, Hoover and Lorenz measures of inequalities and convergence with respect to the Wasserstein $W_1$ metric

Valentin Melot\*

\* Inspection générale des finances (IGF) / General Inspectorate of Finance,  
French Ministry of Economics and Finance, Paris.

Sept. 2024

## Abstract

This article focuses on some properties of three tools used to measure economic inequalities with respect to a distribution of wealth  $\mu$ : Gini coefficient  $G$ , Hoover coefficient or Robin Hood coefficient  $H$ , and the Lorenz concentration curve  $L$ . To express the distributions of resources, we use the framework of random variables and abstract Borel measures, rather than discrete samples or probability densities. This allows us to consider arbitrary distributions of wealth, e.g. mixtures between discrete and continuous distributions.

In the first part (sections 1–4), we discuss alternate definitions of  $G$ ,  $H$  and  $L$  that can be found in economics literature. The Lorenz curve is defined as the normalized integral of the quantile function ([Gas71]), which is not the same as saying “ $L(p)$  is the share of wealth owned by the  $100p$  first centiles of the population” (proposition 1.4). The Gini and Hoover coefficients are introduced in terms of expectation of random variables. In section 3, we interpret Gini and Hoover as geometrical properties of the Lorenz curve (theorem 3.3 and corollary 3.10). In particular, we give a more general and straightforward proof of the main result of [Dor79]. Section 4 gives two direct applications. We *en route* prove the (not trivial) fact that the Lorenz curve fully characterizes a distribution, up to a rescaling (proposition 2.4).

The second part of the article (section 5–7) focuses on the consistency of  $G(\mu)$ ,  $H(\mu)$  and  $L_\mu$  as  $\mu$  is approximated or perturbed. The relevant tool to use is the Wasserstein metric  $W_1$ , i.e. the  $L^1$  metric between quantile functions.  $W_1(\mu_n, \mu_\infty) \rightarrow 0$  if and only if underlying random variables converge in distribution and the total amount of wealth converges. In theorem 5.2 and proposition 5.5, we show that if  $W_1(\mu_n, \mu_\infty) \rightarrow 0$ , then  $G(\mu_n) \rightarrow G(\mu_\infty)$ ,  $H(\mu_n) \rightarrow H(\mu_\infty)$  and  $L_{\mu_n} \rightarrow L_{\mu_\infty}$  uniformly. Subsection 5.4 discusses topological implications of this fact. Thus, applications 6.1, 6.3, 6.7, 6.8 and 6.12 justify that the empirical Gini, Hoover indexes and Lorenz curves computed on a sample or rebuilt with partial information converge to the real Gini, Hoover indexes and Lorenz curve as information increases. Eventually, in section 7, we discuss the situations where the  $W_1$  convergence is not ensured, but weaker assumptions can be made (convergence in distribution in 7.1, convergence of means in 7.2.1 or uniform integrability in 7.2.2)

**Introduction.** We discuss inequality indexes about a given resource, for instance income  $x$ , capital  $k$ , hours of work  $w$  or economic utility  $U(x, k, w)$ .

The repartition of the resource can be mod-

elized by a random variable  $X$  on a probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ . Most articles either modelize  $X$  as a discrete random variable, or as a density with respect to Lebesgue measure. (See for instance [Cho08] on different ways to modelize

incomes.) However, it is sensible to chose more complex modelizations, for instance a mixture of a Dirac mass in 0 (people with no gross income) and a random variable with density such as a gamma or a lognormal for the people with a nonzero income.

What interests us in order to study inequalities is not the variable  $X$  itself, which depends on the underlying probability space, but its distribution  $\mu$ , which is a Borel measure on  $\mathbf{R}_+$  — we restrict to nonnegative variables. In the following article, we will use the formalism of measure theory, as fully introduced in [Le 22]. The details on the notations we use can be found in appendix A. In particular, we note  $\mathfrak{B}$  the Borel  $\sigma$ -algebra on a topological space and  $\mathcal{L}$  the Lebesgue measure on a measurable subset of  $\mathbf{R}$ .  $\mathcal{M}_1(\mathcal{X}, \mathcal{F})$  is the set of measures on the measurable space  $(\mathcal{X}, \mathcal{F})$  with mass 1, i.e. of probability distributions on  $\mathcal{X}$ .

If  $\mu$  is a distribution,  $F_\mu$  is its cumulative distribution function (c.d.f.),  $m_\mu$  its mean (i.e.  $\mathbf{E}[X]$  where  $X \sim \mu$ ) and  $Q_\mu$  its quantile func-

tion. The definition and basic properties of quantile functions, which will be used very often in this article, are summed up in appendix B.1. In particular, we freely use the ‘‘Galois inequalities’’, i.e. the fact that for all  $p \in [0, 1)$  and  $q \in \mathbf{R}_+$ ,  $Q_\mu(p) \leq q \iff p \leq F_\mu(q)$ .

For any  $\alpha > 0$  we note  $\mathbf{M}_\alpha$  the set of probability distributions on  $(\mathbf{R}_+, \mathfrak{B})$  such that  $m_\mu = \alpha$  and  $\mathbf{M} := \bigcup_{\alpha > 0} \mathbf{M}_\alpha$  the set of probability distributions with nonzero, finite means.

In most cases in socioeconomics, we want inequality indicators not be affected by uniform rescaling. (E.g. the indicators need not change if the incomes are expressed in cents rather than in dollars. To take into consideration the fact that the marginal benefit of \$1 is not constant, we can apply a concave transformation  $U$  to the variable  $X$ ). We say that  $\mu$  and  $\nu$  are equal up to a scale factor of  $\alpha$  if and only if there exist two random variables  $X$  and  $Y$  on some probability space such that  $X = \alpha Y$  almost surely (a.s.),  $X \sim \mu$  and  $Y \sim \nu$ . We write  $\mu \equiv \nu$  iff  $\mu$  and  $\nu$  are equal up to any scale factor  $\alpha > 0$ .

## Part I

# Alternate definitions of Lorenz curve, Gini index and Hoover index

## 1 Alternate definitions of the Lorenz curve

We introduce the definition of the Lorenz curve formalized by [Gas71]. Then we discuss the relations that exist between this formalization and a more intuitive definition.

### 1.1 Integral of quantile definition and elementary properties

**Definition 1.1.** Let  $\mu \in \mathbf{M}$  and  $Q_\mu$  its quantile function. The *Lorenz function* of the measure  $\mu$  is:

$$L_\mu : [0, 1] \longrightarrow [0, 1]$$

$$p \longmapsto \frac{\int_0^p Q_\mu(t) dt}{\int_0^1 Q_\mu(t) dt}.$$

The *Lorenz concentration curve* of  $\mu$  is the representative curve of the function  $L_\mu$  in an orthonormal frame.

If  $X$  is a random variable with  $X \sim \mu$ , we allow to write  $L_X := L_\mu$ .

The facts that  $Q_\mu$  is not defined in 1 and that  $Q_\mu(1^-)$  need not be finite are not an issue in this definition. Indeed, the denominator integral is equal to  $m_\mu$  (see lemma B.5 in appendix). Thus, it is nonzero and finite. It follows that the numerator integral is also finite.

**Basic properties.** We give a few elementary properties of the Lorenz function (see for instance [Tho76]). Let  $\mu \in \mathbf{M}$ ,  $m$  its mean. We

have:

- $L_\mu$  is nondecreasing.
- $L_\mu(0) = 0$  and  $L_\mu(1) = 1$ .
- $L_\mu$  is continuous, and even absolutely continuous (see for instance [HMB61, chapter VII, theorem 1, p. 99]).
- For all  $t \in (0, 1]$ ,  $L_\mu$  has a left derivative  $\partial_- L_\mu(t) = \frac{Q_\mu(t)}{m}$ . Likewise, for all  $t \in [0, 1)$ , it has a right derivative  $\partial_+ L_\mu(t) = \frac{Q_\mu(t^+)}{m}$ . These results are inferred from the fact that  $Q_\mu$  is left-continuous (proposition B.2 in appendix) and has a right-limit everywhere.
- As a corollary,  $Q_\mu$  being nondecreasing,  $L_\mu$  is a convex function.
- Thus, for all  $p \in [0, 1]$ ,  $L_\mu(p) \leq p$ .
- The transformation  $\mu \mapsto L_\mu$  is scale-invariant, i.e.  $L_\mu = L_\nu$  as soon as  $\mu \equiv \nu$ .

**Finite case.** Let  $n \in \mathbf{N}^*$ . For any vector  $\mathbf{x} := (x_1, \dots, x_n) \in \mathbf{R}_+^n$ , let  $\mathbf{x}^\uparrow$  the vector with same components but reordered increasingly. Let  $\hat{\mu}$  the empirical measure associated with sample  $\mathbf{x}$ . Then for every  $k \in [0, n]$ , we have  $\int_0^{k/n} Q_\mu(p) dp = \frac{1}{n} \sum_{i=1}^k x_i^\uparrow$ . Hence:

$$L_{\hat{\mu}}\left(\frac{k}{n}\right) = \frac{\sum_{i=1}^k x_i^\uparrow}{\sum_{i=1}^n x_i}$$

and  $L_{\hat{\mu}}$  is affine of the intervals of form  $[\frac{i}{n}, \frac{i+1}{n}]$ .

### Majorization and the Lorenz order.

A vector  $\mathbf{x} := (x_1, \dots, x_n)$  is said to *majorize* another vector  $\mathbf{y} := (y_1, \dots, y_n)$  if  $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$  and for every  $k \in [1, n]$ ,  $\sum_{i=1}^k x_i^\uparrow \leq \sum_{i=1}^k y_i^\uparrow$ . If so, a distribution of incomes  $x_1, \dots, x_n$  is more unequal than  $y_1, \dots, y_n$  and has same mean. See [MOA11] for a complete course on majorization theory.

Let  $\mu$  and  $\nu$  the empirical measures associated with  $\mathbf{x}$  and  $\mathbf{y}$ . for every  $k \in [0, n]$ ,  $L_\mu(\frac{k}{n}) \leq L_\nu(\frac{k}{n})$ . Hence, for all  $p \in [0, 1]$ ,  $L_\mu(p) \leq L_\nu(p)$ .

We say that a measure  $\mu \in \mathbf{M}$  *Lorenz-dominates*  $\nu \in \mathbf{M}$  if  $L_\mu(p) \leq L_\nu(p)$  for all

$p \in [0, 1]$ , i.e.  $L_\mu \leq L_\nu$ . From what precedes, follows that Lorenz domination naturally extends majorization. In particular, one can say that  $\mu$  is more unequal than  $\nu$  if  $L_\mu \leq L_\nu$ . See [Arn87] and [MOA11, chapter 17.C] for a brief introduction to Lorenz order.

## 1.2 Some other intuitive definitions and how to deal with atoms

### 1.2.1 The pseudo-Lorenz function

The Lorenz function is often defined intuitively by economists as “*the function that maps  $p$  to the proportion of the total resource owned by the bottom 100 $p$ %-share of the total population*”. However, this definition is not appropriate.

**Definition 1.2.** Let  $\mu \in \mathbf{M}$ ,  $Q_\mu$  its quantile function. We call *pseudo-Lorenz function of  $\mu$*  the function  $\Lambda_\mu : [0, 1] \rightarrow [0, 1]$  such that for all  $p \in [0, 1)$ ,

$$\Lambda_\mu(p) := \frac{\int_0^{Q_\mu(p)} u d\mu(u)}{\int_0^\infty u d\mu(u)}.$$

and  $\Lambda_\mu(1) := 1$ .

$\Lambda_\mu$  formalises the intuitive definition presented above, using the quantiles formalism. Notice that the denominator is exactly  $m_\mu$ .

Do  $\Lambda$  and  $L$  coincide? It is clear that the answer is no as soon as  $\mu$  has nonzero atoms. In this case, most properties of  $L$  proven before do not hold. For instance, if  $\mu$  is a Dirac mass  $\mu := \delta_x$  (with  $x > 0$ ), then  $Q_\mu(0) = 0$  and  $Q_\mu(p) = x$  for all  $p \in (0, 1]$ , hence  $\Lambda_\mu(0) = 0$  and  $\Lambda_\mu(t) = 1$  for all  $t > 0$ . Thus,  $\Lambda_\mu$  is not continuous at 0.

Actually, there is a relation between  $\Lambda_\mu$  and  $L_\mu$ , as stated in the following lemma:

**Lemma 1.3.** *Let  $\mu \in \mathbf{M}$ . Then, for all  $p \in [0, 1)$ ,*

$$\Lambda_\mu(p) - L_\mu(p) = \frac{Q_\mu(p)}{m_\mu} \cdot [F_\mu(Q_\mu(p)) - p].$$

In order to ease the notation, we write  $Q$ ,  $F$ ,  $m$ ,  $L$  and  $\Lambda$  as an abbreviation for  $Q_\mu$ ,  $F_\mu$ , etc. if no ambiguity.

*Proof.* We perform a pushforward change-of-variable through  $Q$ . Let  $p \in [0, 1)$ . First notice that:

$$\int_0^{Q(p)} u \, d\mu(u) = \int_{0^+}^{Q(p)} u \, d\mu(u).$$

Then, applying the law of unconscious statistician (lemma B.4 in appendix) with the function

$$\begin{aligned} f &: \mathbf{R}_+ \longrightarrow \mathbf{R}_+ \\ x &\longmapsto x \cdot \mathbf{1}_{(0, Q(p)]}(x), \end{aligned}$$

we have:

$$\begin{aligned} \int_{0^+}^{Q(p)} x \, d\mu(x) &= \int_0^\infty f(x) \, d\mu(x) \\ &= \int_0^1 f(Q(u)) \, du \\ &= \int_0^1 Q(u) \cdot \mathbf{1}_{(0, Q(p)]}(Q(u)) \, du \\ \int_{0^+}^{Q(p)} x \, d\mu(x) &= \int_{Q^{-1}\langle(0, Q(p))\rangle} Q(u) \, du. \end{aligned}$$

Yet by Galois inequalities,

$$Q^{-1}\langle(0, Q(p))\rangle = (F(0), F(Q(p))].$$

Furthermore,  $Q(F(0)) = 0$  so  $Q(t) = 0$  for all  $t \in [0, F(0)]$ . Eventually, we get:

$$\int_0^{Q(p)} u \, d\mu(u) = \int_0^{F(Q(p))} Q(t) \, dt.$$

As  $p \leq F(Q(p))$ , we deduce that for all  $p \in [0, 1)$ ,

$$\Lambda(p) - L(p) = \frac{1}{m} \int_p^{F(Q(p))} Q(u) \, du \geq 0.$$

Yet  $Q$  is a constant function over  $[p, F(Q(p))]$ , since it is nondecreasing and takes same values at the endpoints of the interval. This concludes the proof.  $\blacksquare$

From this result and the immediate properties of the quantile function, we deduce the conditions under which  $\Lambda$  and  $L$  match:

**Proposition 1.4.** *Let  $\mu \in \mathbf{M}$ .*

*For all  $p \in [0, 1]$ ,  $L_\mu(p) \leq \Lambda_\mu(p)$ . Furthermore,  $L_\mu(p) = \Lambda_\mu(p)$  iff one of the following assumptions holds:*

(i)  $\exists x \in \mathbf{R}_+, p = F(x)$ .

(ii)  $Q(p) = 0$ .

(iii)  $p = 1$ .

Thus, the “intuitive” definition of Lorenz curve is valid as soon as  $\mu$  has no atom (except possibly in 0), but it fails in every other case.

### 1.2.2 Interpolling atomic Lorenz functions with nonatomic ones?

Since the intuitive definition is valid as soon as  $\mu$  is nonatomic, it might be tempting to deal with nondiffuse measures by approximating them with nonatomic ones, e.g. by allocating the mass of atoms on an interval, so the intuitive definition would hold.

Such a hack is, for instance, sometimes used to define the median of a measure with atoms. To illustrate this, let  $\mu = \frac{1}{2}(\delta_{\{0\}} + \delta_{\{1\}})$ . Some authors define the median in such a way that the median of  $\mu$  would be  $\frac{1}{2}$ . (Whereas  $Q_\mu(0.5)$  is equal to zero, not  $\frac{1}{2}$ ). This, however, does not lead to any satisfying result.

Note that we land on our feet with the integral-of-quantile-function definition. Indeed, one can imagine taking the individuals corresponding to an atom, splitting them into infinitely thin parts, and arbitrarily ordering them. For instance, if  $\mu = \frac{1}{3}\delta_{\{1\}} + \frac{2}{3}\delta_{\{2\}}$ , we would say that the bottom half of the total population is the set of individuals for which  $X = 1$  plus one quarter of those for which  $X = 2$ .

What looks like this approach of “joining the points” best is in fact given by the following approach.

### 1.2.3 Kendall’s parametric curve

A last way to formalize this definition, proposed by [KS45, section 2.23], is to see the Lorenz curve as a parametric curve.

**Definition 1.5 (Kendall curve).** Let  $\mu \in \mathbf{M}$ . The *Kendall curve* of  $\mu$  is the parametrized

curve  $\mathcal{K}_\mu$  of equation:

$$\mathcal{K}_\mu : \begin{cases} x &= \int_0^t d\mu(u) = F_\mu(t), \\ y &= \frac{\int_0^t u d\mu(u)}{\int_0^\infty u d\mu(u)}, \end{cases} \quad (t \in \mathbf{R}_+).$$

Do  $\mathcal{K}$  coincide with what we defined as the Lorenz curve, i.e. the graph of  $L$ ? In fact, if  $\mu$  has atoms then  $F$  has discontinuities, the curve jumps along the  $x$ -axis, so it is not the graph of a function. For instance, if  $\mu = \delta_x$  (with  $x > 0$ ), the Lorenz ‘‘curve’’ according to Kendall’s definition is the reunion of two points  $(0; 0)$  and  $(1; 1)$ . So let’s describe the relations between  $\mathcal{K}$  and  $L$ .

**$\mathcal{K}$  is included in the graph of  $L$ .** This point is easy to check. Let  $t \in \mathbf{R}_+$ . It suffices to prove that  $m \cdot L(F(t)) = \int_0^t u d\mu(u)$ . Using Galois inequalities and the pushforward formula (lemma B.4), we have:

$$\begin{aligned} m \cdot L(F(t)) &= \int_0^{F(t)} Q(u) du \\ &= \int_0^1 \mathbf{1}_{(u \leq F(t))} Q(u) du \\ &= \int_0^1 \mathbf{1}_{(Q(u) \leq t)} Q(u) du \\ &= \int_0^\infty \mathbf{1}_{(x \leq t)} x d\mu(x) \\ m \cdot L(F(t)) &= \int_0^t x d\mu(x). \end{aligned}$$

**What happens outside of  $F(\mathbf{R}_+)$ ?** So the only difference between  $\mathcal{K}$  and the graph of  $L$  lies in the values of  $x$  that are not reached by  $F(t), t \in \mathbf{R}_+$ . These points belong to two categories:

- $x = 1$  (if the support of  $\mu$  is not bounded).
- The  $x$  such that  $F$  is discontinuous at  $x$ . In other words, if  $x \in F(\mathbf{R}_+)$  and  $x < 1$ , then there exists  $t$  such that  $F(t^-) \leq x < F(t)$ . The first inequality can be either strict or an equality.

We successively deal with the three cases.

**Case  $x = 1$ .** As  $t \rightarrow \infty$ ,  $F(t) \rightarrow 1$  and  $\int_0^t u d\mu(u) \rightarrow \int_0^\infty u d\mu(u) = m$ . Hence,  $(1; L(1)) = (1; 1)$  is a limit point of the curve  $\mathcal{K}$ .

**Case  $F(t^-) = x < F(t)$ .** As in previous case, if  $u \rightarrow t$ , then  $F(u) \rightarrow F(t^-) = x$ . Furthermore, by Galois inequalities (inverse versions, see appendix B.1.2):

$$\begin{aligned} \frac{1}{m} \int_0^u v d\mu(v) &\xrightarrow[u < t]{u \rightarrow t} \frac{1}{m} \int_0^{t^-} v d\mu(v) \\ &= \frac{1}{m} \int_0^\infty \mathbf{1}_{(v < t)} v d\mu(v) \\ &= \frac{1}{m} \int_0^1 \mathbf{1}_{(Q(z) < t)} Q(z) dz \\ &= \frac{1}{m} \int_0^1 \mathbf{1}_{(z < F(t^-))} Q(z) dz \\ \frac{1}{m} \int_0^u v d\mu(v) &\xrightarrow[u < t]{u \rightarrow t} L(F(t^-)) = L(x). \end{aligned}$$

Hence,  $(x; L(x))$  is a limit point of  $\mathcal{K}$ .

**Case  $F(t^-) < x < F(t)$ .** By the Galois inequalities,  $x \leq F(t) \iff Q(x) \leq t$ . Furthermore,  $x > F(t^-) \iff Q(x) \geq t$ .

Thus,  $Q$  is constant over the interval  $(F(t^-), F(t)]$ . Hence,  $L$  is affine over the same interval; since it is continuous, it is affine over  $[F(t^-), F(t)]$ . It follows that  $(x, L(x))$  belongs to a line joining  $(F(t^-); L(F(t^-)))$  to  $(F(t); L(F(t)))$ .  $(F(t^-); L(F(t^-)))$  is either a point of  $\mathcal{K}$  or a limit point thereof (see case 2), while  $(F(t); L(F(t))) \in \mathcal{K}$ .

These results can be summarized as follow:

**Proposition 1.6.** *Let  $\mu \in \mathbf{M}$ .*

*The Lorenz curve of  $\mu$ , i.e. the graph of the function  $L_\mu$ , is the union of the following sets:*

- The Kendall curve  $\mathcal{K}_\mu$ ;
- The set of limit points of  $\mathcal{K}_\mu$ ;
- For every  $t \in \mathbf{R}_+$  such that  $F_\mu$  is not continuous at  $t$ , the segment joining the points of coordinates  $(F_\mu(t^-); L_\mu(F_\mu(t^-)))$  and  $(F_\mu(t); L_\mu(F_\mu(t)))$ .

We show the differences between  $L_\mu, \Lambda_\mu$  and  $\mathcal{K}_\mu$  for  $\mu := \frac{1}{2}(\mathcal{L}_{[0,1]} + \delta_{0.5})$  in fig. 1, p. 7.

## 2 The Lorenz curve and the mean fully characterize a distribution

The goal of this section is to formally prove the (not so trivial) fact that  $L_\mu = L_\nu$  only if  $\mu \equiv \nu$ .

### 2.1 Bijection between $\mathbf{M}$ and $\mathfrak{Q}$

In appendix B.1, we recall that quantile functions of measures taken in  $\mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$  are non-decreasing, left-continuous and take the value 0 in 0. In fact, these properties fully characterize the set of quantile functions.

Let  $\mathfrak{Q}$  the set of nondecreasing, left-continuous functions  $[0, 1] \rightarrow \mathbf{R}_+$  taking value 0 in 0.

If  $f$  is a function and  $\mu$  a measure, we note  $f_\#(\mu)$  the pushforward of  $\mu$  through  $f$  (see appendix A for details).

**Proposition 2.1.** *The following mappings are inverse bijections:*

$$\begin{array}{ccc} \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B}) & \xrightarrow{\sim} & \mathfrak{Q} \\ \mu & \mapsto & Q_\mu \\ q_\#(\mathcal{L}_{[0,1]}) & \xleftarrow{\sim} & q. \end{array}$$

*Proof.* It is a well known fact that for all  $\mu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$ ,  $(Q_\mu)_\#(\mathcal{L}_{[0,1]}) = \mu$  (see proposition B.3 in appendix). Hence, it suffices to prove that for every  $q \in \mathfrak{Q}$ ,  $q$  is the quantile function of the measure  $\mu = q_\#(\mathcal{L}_{[0,1]})$ .

First, notice that for all  $x \in \mathbf{R}_+$ ,

$$\begin{aligned} F_\mu(x) &= \mu([0, x]) \\ &= \mathcal{L}(q^{-1}([0, x])) \\ &= \mathcal{L}(\{u \in [0, 1] : 0 \leq q(u) \leq x\}) \\ F_\mu(x) &= \mathcal{L}(\{q \leq x\}), \end{aligned}$$

where we note:

$$\{q \leq x\} := \{u \in [0, 1] : q(u) \leq x\}.$$

As  $q$  is nondecreasing, the set  $\{q \leq x\}$  is an interval of form  $[0, s)$  or  $[0, s]$ , so  $F_\mu(x) = \sup\{q \leq x\}$ . As we assume that  $q$  is left-continuous,  $s$  belongs to this interval, so the c.d.f. of  $\mu$  is given by:

$$F_\mu(x) = \max\{q \leq x\}. \quad (2.2)$$

Now, fix  $p \in [0, 1]$ ; we need to prove that  $q(p) = Q_\mu(p)$ . We distinguish two cases:

- This is immediate for  $p = 0$ .

- Now assume  $p > 0$ . By definition of  $Q_\mu(p)$ , having  $Q_\mu(p) = q(p)$  is equivalent to having both the following assertions true:

- (i)  $F_\mu(q(p)) \geq p$ ;
- (ii)  $\forall r \in \mathbf{R}_+, r < q(p) \implies F_\mu(r) < p$ .

By (2.2), (i) is equivalent to having  $p \in \{q \leq q(p)\}$ , which is true.

Furthermore, by (2.2), (ii) is equivalent to

$$\forall r \in \mathbf{R}_+, r < q(p) \implies p \notin \{q \leq r\}$$

which is also true.  $\blacksquare$

Notice that this latest equivalence fails if we do not have a max but only a sup in (2.2), i.e. is  $q$  is not left-continuous.

### 2.2 Bijection between $\mathbf{M}$ and $\mathfrak{L} \times \mathbf{R}_+^*$

**Definition 2.3.**  $\mathfrak{L}$  is the subspace of functions  $\ell \in [0, 1]^{[0,1]}$  that satisfy all four following conditions:

- (i)  $\ell$  is continuous,
- (ii)  $\ell$  is convex,
- (iii)  $\ell(0) = 0$ ,
- (iv)  $\ell(1) = 1$ .

Notice that (i) can be replaced by: (i')  $\ell$  is continuous at 1. Indeed, if  $\ell$  is convex, then it is continuous over  $(0, 1)$ . Hypotheses (ii), (iii) and (iv) imply that,  $\ell(p) \leq p$  for all  $p \in (0, 1)$ ; hence  $\ell(0^+) = 0$  and  $\ell$  is continuous at 0.

**Proposition 2.4.** *The following mapping is a bijection  $\mathbf{M} \approx \mathfrak{L} \times \mathbf{R}_+^*$ :*

$$\begin{array}{ccc} \Phi & : & \mathbf{M} \xrightarrow{\sim} \mathfrak{L} \times \mathbf{R}_+^* \\ & & \mu \mapsto (L_\mu, m_\mu). \end{array}$$

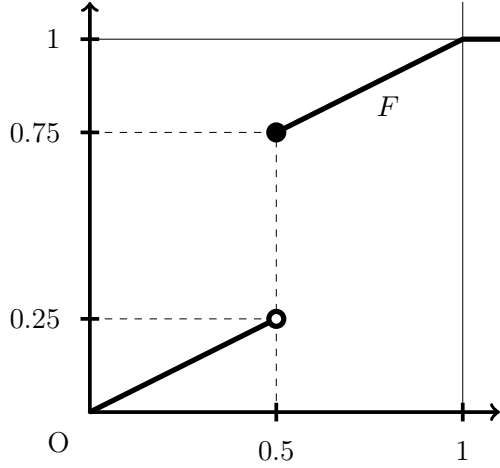
*Its inverse is:*

$$\begin{array}{ccc} \Phi^{-1} & : & \mathfrak{L} \times \mathbf{R}_+^* \xrightarrow{\sim} \mathbf{M} \\ & & (\ell, n) \mapsto (n \cdot \partial_- \ell)_\#(\mathcal{L}_{[0,1]}). \end{array}$$

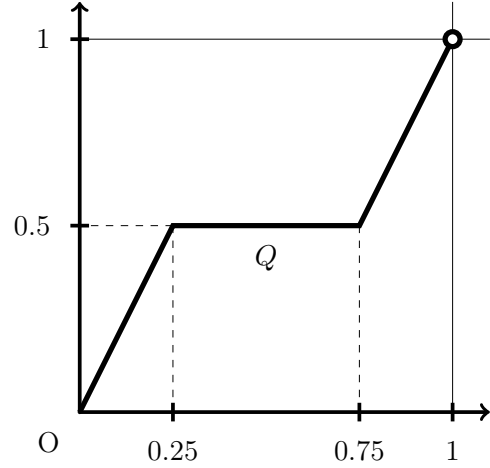
This means, in particular, that the space of Lorenz functions is *exactly*  $\mathfrak{L}$ .

To prove this proposition, we first need a technical lemma:

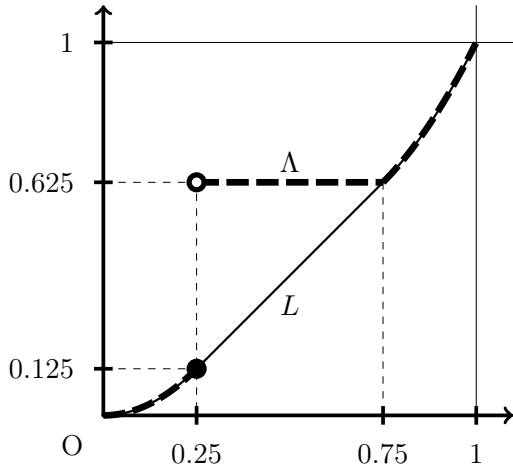




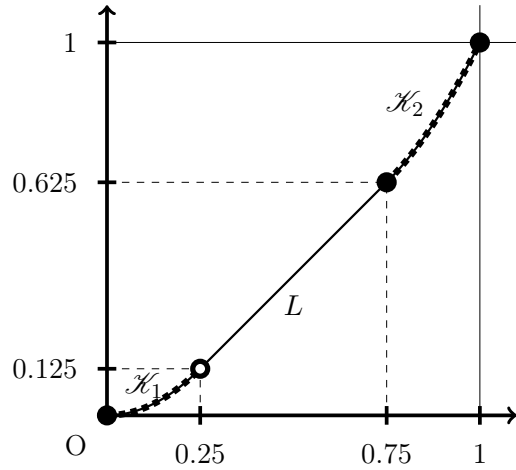
(a) Cumulative distribution function of  $\mu$ .  $F$  has a discontinuity in 0.5 and is right-continuous.



(b) Quantile function of  $\mu$ .



(c) The thin line represents the Lorenz function of  $\mu$  as we defined it in 1.1. The thick, dashed line is the graph of the pseudo-Lorenz  $\Lambda$ , which has a discontinuity in 0.25. These functions coincide at all  $p$  such that there exists an  $x \in \mathbf{R}_+$ , with  $p = F(x)$ .



(d) Here again, the thin line is the graph of the Lorenz function  $L$ . The thick dotted line is the Kendall curve  $\mathcal{K}$ . This one is divided in two connected components  $\mathcal{K}_1$  and  $\mathcal{K}_2$ .  $\mathcal{K}_2$  is closed, while  $\mathcal{K}_1$  does not include its end  $(0.25; 0.125)$ . These two components are included in the graph of  $L$ , the rest being a segment joining the ends of  $\mathcal{K}_1$  and  $\mathcal{K}_2$ .

Figure 1: Illustration of the differences between the Lorenz function  $L$ , the Kendall curve  $\mathcal{K}$  and the pseudo-Lorenz function  $\Lambda$  of a measure  $\mu = \frac{1}{2}(\mathcal{L}_{[0,1]} + \delta_{0.5})$ .

Consider any probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ . Let  $U$  a random variable uniformly distributed in  $[0, 1]$  and  $X$  Bernoulli with parameter 0.5 with  $X \perp\!\!\!\perp U$ . Let  $Y = 0.5$  if  $X = 0$  and  $Y = U$  if  $X = 1$ .  $\mu$  is the distribution of  $Y$ .

**Lemma 2.5.** *Let  $(a, b]$  be an interval of  $\mathbf{R}$  and  $f$  a convex function defined over  $(a, b]$ . The left derivative function  $x \mapsto \partial_- f(x)$  is left-continuous.*

*Proof.* For any  $a < x < y \leq b$ , let  $S(x, y) := \frac{f(y) - f(x)}{y - x}$ . As  $f$  is convex, the ‘‘chordal slope lemma’’ ensures that  $S$  is nondecreasing with respect to each variable. Hence, for all  $x \in (a, b]$ ,  $S(x^-, x)$  exists (it may be equal to  $+\infty$ ); it is by definition equal to  $\partial_- f(x)$ . Thus,  $x \mapsto \partial_- f(x)$  is nondecreasing. Thus, it admits a left-limit everywhere.

Fix  $z \in (a, b]$ . We immediately have  $\partial_- f(z^-) \leq \partial_- f(z)$ . Let us prove the opposite inequality.

For any  $x < y < z$ , we have by convexity:  $S(x, y) \leq \partial_- f(y)$ . Hence taking the left limit  $y \rightarrow z^-$ , we get  $S(x, z) \leq \partial_- f(z^-)$ . But then taking the left limit  $x \rightarrow z^-$ , we get  $S(z^-, z) \leq \partial_- f(z^-)$ ; in other words  $\partial_- f(z) \leq \partial_- f(z^-)$ .

Then,  $\partial_- f$  is left-continuous at  $z$ .  $\blacksquare$

*Proof of the proposition 2.4.* The proof lies in two parts: (i) proving  $\Phi$  is an injection, (ii) proving for all  $(L, m) \in \mathcal{L} \times \mathbf{R}_+^*$ ,  $\Phi(\mu) = (L, m)$  where  $\mu = (m \cdot \partial_- L)_\#(\mathcal{L}_{[0,1)})$ .

**$\Phi$  is an injection.** Assume  $\mu$  and  $\nu$  are measures such that  $L_\mu = L_\nu$  and  $m_\mu = m_\nu$ .

By taking the left derivative at any  $u \in (0, 1)$ :  $\frac{Q_\mu(u)}{m_\mu} = \frac{Q_\nu(u)}{m_\nu}$ . The equality also stands for  $u = 0$ . Thus,  $Q_\mu = Q_\nu$ . By proposition 2.1,  $\mu = \nu$ .

**$\Phi$  is a surjection.** Let  $\ell \in \mathcal{L}$  and  $m \in \mathbf{R}_+^*$ .

By convexity,  $\ell$  admits a finite left-derivative everywhere on  $(0, 1]$  and a finite right derivative everywhere on  $[0, 1)$ . Furthermore,  $\partial_- \ell$  and  $\partial_+ \ell$  are nondecreasing, and we have for all  $x \in (0, 1)$ ,

$$0 \leq \partial_+ \ell(0) \leq \partial_- \ell(x) \leq \partial_+ \ell(x).$$

We define  $q : [0, 1) \rightarrow \mathbf{R}_+$  by setting  $q(0) = 0$  and  $q(p) = \partial_- \ell(p)$  for  $p > 0$ .  $q$  is a nondecreasing function.

By lemma 2.5,  $q$  is left-continuous. So is the function  $n \cdot q$ . Hence, by proposition 2.1, the measure  $\mu = (n \cdot q)_\#(\mathcal{L}_{[0,1)})$  fullfills  $q = \frac{Q_\mu}{n}$ .

Now we need to prove that  $m_\mu = n$  and that  $\ell$  is the integral of function  $q$ . Contrary to what may seem at first glance, this is non-trivial. (See for instance the discussion in the preliminary section of [HMB61, chapter VI, p. 83].) This requires a few steps:

- First, as  $q$  is nondecreasing, it is  $\mathfrak{B}$ -measurable.
- $\ell$  is derivable almost everywhere, and we have  $\ell'(x) = q(x)$  at every  $x \in (0, 1)$  where  $\ell$  is derivable.
- Take any  $0 < x < 1$ . The function  $q$  is bounded by  $q(x)$  over the interval  $[0, x]$ . Thus,  $\ell$  is  $q(x)$ -Lipschitz over this interval, hence absolutely continuous over it.
- Hence, we infer from [HMB61, chapter VII, theorem 3, p. 100] that  $q$  is integrable over  $[0, x]$  and:

$$\int_0^x q(t) dt = \ell(x). \quad (2.6)$$

- Now we deal with  $x = 1$ . We set for every  $n \in \mathbf{N}$ :  $g_n = \mathbf{1}_{[0, 1-2^{-n}]}$ . Then the sequence of functions  $(q \cdot g_n)_{n \in \mathbf{N}}$  is nondecreasing and converges pointwise to  $q$  over  $[0, 1)$ . Hence, by monotone convergence theorem:

$$\begin{aligned} \int_0^{1^-} q(t) dt &= \int_0^{1^-} \lim_{n \rightarrow \infty} q(t) g_n(t) dt \\ &= \lim_{n \rightarrow \infty} \int_0^{1^-} q(t) g_n(t) dt \\ \int_{[0,1)} q(t) dt &= \lim_{n \rightarrow \infty} \ell(1 - 2^{-n}). \end{aligned}$$

By continuity of  $\ell$ , this limit is equal to  $\ell(1)$ , i.e. to 1. Then,  $q$  is integrable over  $[0, 1)$ . From proposition B.5, we get:

$$m_\mu = \int_0^{1^-} n \cdot q(t) dt = n.$$

- Eventually, for all  $x \in [0, 1)$ , (2.6) can be rewritten as

$$\frac{1}{m_\mu} \int_0^\infty Q_\mu(x) dx = \ell(x)$$

i.e.  $L_\mu(x) = \ell(x)$ .  $\blacksquare$

**Corollary 2.7.** *Let  $\mu, \nu \in \mathbf{M}$ .  $L_\mu = L_\nu$  if and only if  $\mu$  is a rescaling of  $\nu$ .*



### 2.3 The quotient set $\mathbf{M}/\mathbf{R}_+^*$

The relation  $\equiv$  is an equivalence relation. Let  $\mu \in \mathbf{M}$  and  $[\mu]$  the equivalence class of  $\mu$ . The mapping  $\nu \in [\mu] \mapsto m_\nu \in \mathbf{R}_+^*$  is one-to-one. Thus, we note  $\mathbf{M}/\mathbf{R}_+^*$  the quotient set  $\mathbf{M}/\equiv$ .

It is immediate to check that the mapping:

$$\begin{aligned} \mathbf{M} &\xrightarrow{\sim} (\mathbf{M}/\mathbf{R}_+^*) \times \mathbf{R}_+^* \\ \mu &\mapsto ([\mu], m_\mu) \end{aligned}$$

is a bijection.

Furthermore, for all  $\alpha > 0$ , the mapping  $\mu \in \mathbf{M}_\alpha \mapsto [\mu] \in \mathbf{M}/\mathbf{R}_+^*$  is a bijection.

Finally, an immediate corollary of proposition 2.4 is:

**Corollary 2.8.** *The mapping*

$$\begin{aligned} \Psi &: \mathbf{M}/\mathbf{R}_+^* \xrightarrow{\sim} \mathcal{L} \\ &[\mu] \mapsto L_\mu \end{aligned}$$

is a bijection.

In section 5, we will show that all these bijections are in fact homeomorphisms for some natural topologies.

## 3 Alternate definitions of the Gini and Hoover indicators

In the following section, we introduce two commonly used inequality indexes: the Gini index and the Hoover index.

These indexes are directly related to geometric properties of the Lorenz curve. Regarding Gini index, this relation is known since [Gin14] and commonly used by economists. As far as we know, only [Dor79] gives a proof which is valid in almost all generality, but some restrictions remain and the redaction of the proof is somehow surprising. Hence, we give a full proof, valid in all generality as soon as  $\mu \in \mathbf{M}$ . A more detailed motivation of this section is given in appendix C.

### 3.1 Gini and Hoover as mean differences and mean deviation

**Definition 3.1.** Let  $\mu \in \mathbf{M}$ ,  $X, X' \sim \mu$  independent and identically distributed (i.i.d.), and  $m = \mathbf{E}[X]$  the mean of  $\mu$ .

The *Gini coefficient* of  $\mu$  is the ratio:

$$G(\mu) := \frac{\mathbf{E}[|X - X'|]}{2m}.$$

The *Hoover coefficient* of  $\mu$  is the ratio:

$$H(\mu) := \frac{\mathbf{E}[|X - m|]}{2m}.$$

If  $X$  is a random variable on a probability space  $(\Omega, \mathcal{F})$  with distribution  $\mu$ , we allow to write  $G(X) = G(\mu)$  and  $H(X) = H(\mu)$ .

The Hoover coefficient is sometimes called the Pietra index.

**Elementary properties.** Let  $\mu \in \mathbf{M}$  and  $(X, X') \sim \mu \otimes \mu$ . Then:

- $G$  and  $H$  are scale-invariant.
- $G(\mu), H(\mu) \in [0, 1]$ , by triangle inequality.
- $H(\mu) \leq G(\mu)$ . Indeed, Jensen's inequality implies that

$$|\mathbf{E}[X - X'|X]| \leq \mathbf{E}[|X - X'| \mid X] \quad (3.2)$$

where  $\mathbf{E}[\bullet \mid X]$  denotes the expectation conditional to  $X$ .

Since  $\mathbf{E}[X|X] = X$  and  $\mathbf{E}[X'|X] = m_\mu$ , the left-hand side equal to  $|X - m_\mu|$ . Taking the expectation of the inequality concludes.

- The following assertions are equivalent:

- (i)  $G(\mu) = 0$ .
- (ii)  $H(\mu) = 0$ .
- (iii)  $\mu$  is a Dirac mass.

This can be deduced from the equality case of the triangle inequality in (3.2).

- The upper bound 1 is strict, i.e.  $G(\mu) < 1$ . Otherwise, we would have

$$\mathbf{E}[|X - X'|] = \mathbf{E}[|X|] + \mathbf{E}[|-X'|],$$

implying that  $X$  and  $-X'$  have same sign almost surely. This would imply that  $X = X' = 0$  a.s., which is absurd.

- But 1 is the best upper bound for both indicators, i.e.  $\sup_{\mu \in \mathbf{M}} G(\mu) = 1$ . For instance, for all  $\alpha \in [0, 1)$ , let:

$$\mu_\alpha = \alpha \cdot \delta_0 + (1 - \alpha) \cdot \delta_1,$$

then  $G(\mu_\alpha) = H(\mu_\alpha) = \alpha$ .

\*

Until the end of the subsection, we fix a distribution  $\mu \in \mathbf{M}$  and let  $m$ ,  $F$ ,  $Q$  and  $L$  respectively the mean, the c.d.f., the quantile function and the Lorenz function of  $\mu$ .

### 3.2 Gini coefficient as area between the Lorenz curve and the diagonal

**Theorem 3.3.** *Let  $\mu$  be a distribution in  $\mathbf{M}$ ,  $G(\mu)$  its Gini coefficient, and  $L_\mu$  its Lorenz function. We have:*

$$G(\mu) = 1 - 2 \int_0^1 L_\mu(p) dp.$$

In other words, the Gini coefficient is (up to a factor 2) the area between, on the one hand, the Lorenz curve of the considered definition, and on the other hand the diagonal. Notice that the diagonal is the Lorenz curve associated to a perfectly equal distribution.

We propose two proofs of this result. The first one is directly inspired from [Dor79], but more straightforward and with no hypotheses on  $\mu$  except  $0 < m_\mu < \infty$ . The second one is a new proof, whose general idea is courtesy of David Leturcq. We present it under the assumption that  $\mu$  has no atom. The general idea of the proof is still valid if  $\mu$  has atoms but needs a few refinements, that are fully detailed in appendix D.1.1.

#### 3.2.1 First proof (via $\int_0^1 (1 - F(t))^2 dt$ )

**Lemma 3.4 (Dorfman, 1979).** *Let  $\mu \in \mathbf{M}$ ,  $F$  its cumulative distribution function,  $G$  its Gini coefficient. Then:*

$$G = 1 - \frac{\int_0^\infty (1 - F(t))^2 dt}{\int_0^\infty (1 - F(t)) dt}.$$

*Proof.* Notice that for any real numbers  $a$  and  $b$ , we have:  $|a - b| = a + b - 2 \min(a, b)$ .

Let  $X$  and  $Y$  be independent random variables of same distribution  $\mu$  on  $(\Omega, \mathcal{F}, \mathbf{P})$ . We have:

$$\mathbf{E}[|X - Y|] = 2\mathbf{E}[X] - 2\mathbf{E}[\min(X, Y)].$$

Let  $Z = \min(X, Y)$ . We thus have:

$$G = \frac{\mathbf{E}[|X - Y|]}{2\mathbf{E}[X]} = 1 - \frac{\mathbf{E}[Z]}{\mathbf{E}[X]}.$$

Notice that:

$$\int_0^\infty (1 - F(t)) dt = \mathbf{E}[X]$$

(see lemma A.1 in appendix).

Now let's focus on the numerator. Let  $F_Z$  the c.d.f. of  $Z$ . We notice that for all  $t \in \mathbf{R}_+$ ,

$$\begin{aligned} 1 - F_Z(t) &= \mathbf{P}(Z > t) \\ &= \mathbf{P}((X > t) \cap (Y > t)) \\ 1 - F_Z(t) &= (1 - F(t))^2 \end{aligned}$$

since  $X \perp\!\!\!\perp Y$ .

Lemma A.1 now gives us that:

$$\mathbf{E}[Z] = \int_0^\infty (1 - F(t))^2 dt. \quad \blacksquare$$

*Proof of theorem 3.3.* Let

$$\begin{aligned} N &:= \int_0^\infty (1 - F(t))^2 dt \\ N &= \int_{t=0}^\infty \int_{u=F(t)}^1 2(1 - u) du dt. \end{aligned}$$

Since the Lebesgue measure is nonatomic, we can drop the lower bound of the inner integral, and then use Galois inequalities and Fubini's theorem:

$$\begin{aligned} N &= 2 \int_{t=0}^\infty \int_{u=F(x)^+}^1 (1 - u) du dt \\ &= 2 \int_{t=0}^\infty \int_{u=0}^1 (1 - u) \mathbf{1}_{u > F(t)} du dt \\ &= 2 \int_{t=0}^\infty \int_{u=0}^1 (1 - u) \mathbf{1}_{Q(u) > t} du dt \\ &= 2 \int_{u=0}^1 \int_{t=0}^{Q(u)^-} dt (1 - u) du \end{aligned}$$

$$N = 2 \int_0^1 Q(u)(1-u) \, du.$$

Given the definition of  $L(u) = \frac{1}{m} \int_0^u Q(t) \, dt$ , an integration by parts shows that:

$$N = 2m \int_0^1 L(u) \, du.$$

(For a general proof of the integration by part formula in case where the “primitives” are not differentiable *everywhere* but only absolutely continuous, see for instance [HMB61, chapter VII, theorem 4, p. 103].)

Injecting this latter expression of  $N$  in the numerator of the formula in lemma 3.4 concludes the proof. ■

### 3.2.2 Second proof (direct computation) for nonatomic measures

*Proof.* The proof lies in switching between integrals with respect to  $\mu$  and with respect to Lebesgue measure. As  $\mu$  is assumed to be diffuse, one can include or exclude the bounds of the integrals without having to be cautious.

Let  $I := m \int_0^1 L(p) \, dp$ . As  $\mu$  is diffuse, thanks to proposition 1.4, one has

$$I = \int_0^1 m \cdot \Lambda(p) \, dp = \int_{p=0}^1 \int_{u=0}^{Q(p)} u \, d\mu(u) \, dp.$$

Using Galois inequalities and Fubini’s theorem, the integral can be rewritten as:

$$\begin{aligned} I &= \int_{p=0}^1 \int_{u=0}^{\infty} u \mathbf{1}_{u < Q(p)} \, d\mu(u) \, dp \\ &= \int_{p=0}^1 \int_{u=0}^{\infty} u \mathbf{1}_{F(u) < p} \, d\mu(u) \, dp \\ &= \int_0^{\infty} \int_{p=F(u)}^1 u \, dp \, d\mu(u) \\ I &= \int_0^{\infty} u \cdot (1 - F(u)) \, d\mu(u) \end{aligned} \quad (3.5)$$

$$\begin{aligned} &= \int_0^{\infty} u \int_{s=u}^{\infty} d\mu(s) \, d\mu(u) \\ I &= \iint_{0 \leq u \leq s} u \, d\mu(u) \, d\mu(s). \end{aligned} \quad (3.6)$$

But from (3.5), we also deduce that:

$$I = \int_0^{\infty} u \, d\mu(u) - \int_0^{\infty} \int_{s=0}^u u \, d\mu(s) \, d\mu(u)$$

$$I = m - \iint_{0 \leq s \leq u} u \, d\mu(s) \, d\mu(u).$$

Then, performing the permutation of variables  $u \leftrightarrow s$ , we write:

$$I = \int_{0 \leq u \leq s} s \, d\mu(u) \, d\mu(s). \quad (3.7)$$

Summing up the equations (3.6) and (3.7) gives us:

$$I = \frac{m}{2} - \frac{1}{2} \iint_{0 \leq u < s} (s - u) \, d\mu(s) \, d\mu(u).$$

By symmetry, we have:

$$m \cdot G(\mu) = \frac{1}{2} \iint_{\mathbf{R}_+^2} |s - u| \, d\mu(u) \, d\mu(s)$$

$$m \cdot G(\mu) = \iint_{0 \leq u < s} (s - u) \, d\mu(s) \, d\mu(u).$$

Thus,

$$I = \frac{m}{2} - \frac{m}{2} \cdot G(\mu)$$

which concludes the proof. ■

## 3.3 Interpretations of the Hoover coefficient

### 3.3.1 Preliminary computations

Let:

$$P := \int_0^{m^-} (m - x) \, d\mu(x)$$

$$R := \int_{m^+}^0 (x - m) \, d\mu(x).$$

$P$  (resp.  $R$ ) is the share of income that the people poorer (resp. richer) than the average own above (resp. below) the average. We have:

$$H(\mu) = \frac{R(\mu) + P(\mu)}{2m}$$

Furthermore:

$$\begin{aligned} R - P &= \int_{m^+}^{\infty} x \, d\mu(x) - (1 - F(m))m \\ &\quad + \int_0^{m^-} x \, d\mu(x) - F(m^-)m \\ &= \int_0^{\infty} x \, d\mu(x) - \mu(\{m\})m \\ &\quad - m + mF(m) - mF(m^-) \end{aligned}$$

$$R - P = 0.$$

This implies in particular that:

$$H(\mu) = \frac{R(\mu)}{m} = \frac{P(\mu)}{m}. \quad (3.8)$$

### 3.3.2 Hoover coefficient as Robin Hood index

**Share redistributed.** Another classical way to see the Hoover coefficient is that it is the share of wealth to take to the people richer than the average and to redistribute to the poors in order to reach a perfect equality.

Indeed, the previous result simply means that the Hoover index is the relative share of wealth that the rich have above the average. This one is equal to the relative share of wealth that that the poors need to reach the average. Hence, taking a share  $H(\mu)$  of wealth (to the people above the average) and redistributing it (to the people below the average), we reach perfect equality.

This is why the Hoover coefficient is sometimes called “Robin Hood index”. (Even though Robin Hood would not steal from the riches to give to the poors, but from greedy government that was over-taxing their people.)

**Most efficient way to redistribute.** Is there a more efficient way to redistribute the money in order to reach the perfect equality? Our intuition doubts it.

The answer lies in optimal transportation theory. Searching for the most efficient way to redistribute money means solving the optimal transformation problem from distribution  $\mu$  to distribution  $\delta_m$  with nonsymmetric cost function  $c(x, y) = x - y$  if  $x > y$ ,  $c(x, y) = 0$  otherwise. (One only counts the taxes *taken*, not the aids and subsidies *paid*.) Adopting Kantorovitch’s optimal transport problem framework (see [Vil03, introduction]), we are faced to the minimization problem:

$$I = \min_{\pi \in \Pi(\mu, \delta_m)} \iint_{\mathbf{R}_+^2} c(x, y) \, d\pi$$

where  $\Pi(\mu, \delta_m)$  is the set of measures on  $(\mathbf{R}_+, \mathfrak{B})$  with marginals  $\mu$  and  $\delta_m$ .

However, as  $\delta_m$  is deterministic, there is only one measure in the set  $\Pi(\mu, \delta_m)$ , namely the tensor product  $\mu \otimes \delta_m$ . Hence, the minimum we are looking can be written:

$$\begin{aligned} I &= \int_{x=0}^{\infty} \left( \int_{y=0}^{\infty} \mathbf{1}_{x \geq y} (x - y) \, d\delta_m(y) \right) \, d\mu(x) \\ &= \int_{x=0}^{\infty} (x - m) \mathbf{1}_{x \geq m} \, d\mu(x) \\ &= \int_m^{\infty} (x - m) \, d\mu(x) \\ I &= R = m \cdot H(\mu). \end{aligned}$$

Thus,  $H(\mu)$  is the optimal share of incomes to be redistributed.

**Open question.** Can the Gini coefficient be interpreted in a similar, natural way?

At this point, the answers seems to be negative. This is a usual criticism against the Gini index: for instance, Piketty writes — among other criticisms — that the Gini coefficient gives “an abstract and sterile view of inequality” [Pik14, p. 408].

### 3.3.3 Hoover coefficient as maximum vertical distance between the Lorenz curve and the diagonal

Another classic interpretation of the Hoover coefficient involves the Lorenz curve.

**Proposition 3.9.** *Let  $\mu \in \mathbf{M}$ ,  $m$  its mean,  $L$  its Lorenz function and  $F$  its cumulative distribution function. Then:*

$$H(\mu) = F(m) - L(F(m)).$$

*Proof.* From (3.8), we can write:

$$\begin{aligned} H(\mu) &= \frac{1}{m} \int_0^{m^-} (m - x) \, d\mu(x) \\ &= \frac{1}{m} \int_0^m (m - x) \, d\mu(x) \\ H(\mu) &= F(x) - \frac{1}{m} \int_0^m x \, d\mu(x). \end{aligned}$$

Now, we apply the pushforward theorem (lemma B.4 in appendix) with function

$$\begin{aligned} f : \mathbf{R}_+ &\longrightarrow \mathbf{R}_+ \\ x &\longmapsto x \cdot \mathbf{1}_{x \leq m}. \end{aligned}$$

Hence,

$$\begin{aligned} \int_0^m x \, d\mu(x) &= \int_0^1 Q(p) \mathbf{1}_{Q(p) \leq m} \, dp \\ &= \int_0^1 Q(p) \mathbf{1}_{p \leq F(m)} \, dp \\ &= \int_0^{F(m)} Q(p) \, dp \\ \int_0^m x \, d\mu(x) &= m L(F(m)). \quad \blacksquare \end{aligned}$$

**Theorem 3.10.** *Let  $\mu$  be a distribution in  $\mathbf{M}$  and  $L$  its Lorenz function. We have:*

$$H(\mu) = \max_{p \in [0,1]} (p - L(p)).$$

*Proof.* Let  $\Phi : p \mapsto p - L(p)$ . Using proposition 3.9, it is enough to prove that the concave function  $\Phi$  reaches its maximum in  $F(m)$ .

As per  $L$ ,  $\Phi$  has a left-derivative  $\partial_- \Phi$  and a right-derivative  $\partial_+ \Phi$  in every point of  $(0, 1)$ . We have for  $t \in (0, 1)$ :  $\partial_- \Phi(t) = 1 - \frac{Q(t)}{m}$  and  $\partial_+ \Phi(t) = 1 - \frac{Q(t^+)}{m}$ .  $\Phi$  reaches its maximum in  $p_*$  if and only if  $\partial_- \Phi(p_*) \geq 0$  and  $\partial_+ \Phi(p_*) \leq 0$ . This is equivalent to having both  $Q(p_*) \leq m$  and  $Q(p_*^+) \geq m$ .

On the one hand, the Galois inequalities imply that  $Q(F(m)) \leq m$ . On the other hand, they ensure that if  $x > F(m)$ , then  $Q(x) > m$ . Taking the infimum,  $Q(F(m)^+) \geq m$ .  $\blacksquare$

## 4 Some direct applications of these results

### 4.1 $G$ and $H$ are nondecreasing with respect to Lorenz-domination

**Proposition 4.1.** *Let  $\mu, \nu \in \mathbf{M}$ . Assume that for all  $p \in [0, 1]$ ,  $L_\mu(p) \leq L_\nu(p)$ . Then:*

1.  $G(\nu) \leq G(\mu)$  and  $H(\nu) \leq H(\mu)$ ;
2. Furthermore,  $G(\mu) = G(\nu)$  if and only if  $\mu \equiv \nu$  ( $\mu = \nu$  up to some scale factor).

*Proof.* The first part is an immediate consequence of theorems 3.3 and 3.10.

Now assume  $G(\mu) = G(\nu)$ . Then, by theorem 3.3, we have  $\int_0^1 L_\nu(t) - L_\mu(t) \, dt = 0$ .  $L_\nu - L_\mu$  is nonnegative, continuous, so  $L_\mu = L_\nu$  on  $[0, 1]$ . Hence, by corollary 2.7,  $\mu \equiv \nu$ .  $\blacksquare$

From part 2 of the proposition, we can say that  $G$  is *strictly increasing* with respect to Lorenz-majorization. However,  $H$  is insensitive to redistribution either among the group of people with incomes higher (resp. lower) than the average, and thus not increasing. For instance, let

$$\mu := \frac{1}{4} (2\delta_0 + \delta_1 + \delta_3), \quad \nu := \frac{1}{4} (2\delta_0 + 2\delta_2)$$

then  $L_\mu \leq L_\nu$ ,  $H(\mu) = H(\nu) = \frac{1}{2}$ , but  $\mu \not\equiv \nu$ .

### 4.2 Extreme values of Gini index under constraint on Hoover index

Theorem 3.3 states that the Gini index of a measure  $\mu \in \mathbf{M}$  is two times the area of the

surface

$$S_\mu := \{(x, y) : x \in [0, 1], L_\mu(x) \leq y \leq x\}.$$

We have  $S_\mu = T \cap \Gamma_\mu$ , where  $T$  is the (full) triangle

$$T := \{(x, y) : 0 \leq y \leq x \leq 1\}$$

and  $\Gamma_\mu$  is the epigraph of  $L_\mu$ , i.e.

$$\Gamma_\mu := \{(x, y) : x \in [0, 1], L_\mu(x) \leq y \leq 1\}.$$

As  $L_\mu$  is a convex function,  $\Gamma_\mu$  is a convex surface; so is  $T$ . Hence, the surface  $S_\mu$  is convex.

Theorem 3.10 ensures that a measure  $\mu$  has a Hoover index of  $h$  if, and only if, its Lorenz curve touches the line of equation  $y = x - h$  but never goes below (see fig. 2a, p. 15). We can describe the measures reaching the extreme values.

**Proposition 4.2.** *Fix  $h \in (0, 1)$ .*

1. *We have:*

$$\{G(\mu) : \mu \in \mathbf{M}, H(\mu) = h\} = [h, 2h - h^2].$$

2. *The measures  $\mu$  for which*

$$G(\mu) = H(\mu) = h$$

*are exactly the bimodal measures of form:*

$$\mu = \alpha \cdot \delta_{m \cdot (1 - \frac{h}{\alpha})} + (1 - \alpha) \cdot \delta_{m \cdot (1 + \frac{h}{1 - \alpha})}$$

*where  $m > 0$  and  $\alpha \in [h, 1)$ .*

$m$  is the mean of the distribution, and  $\alpha$  a form parameter that is the relative share of the group of poors. In this configuration, the total group of poors owns a share  $\alpha - h$  of the total resource. At the same time, the group of rich (share  $1 - \alpha$  of the population) owns a share  $1 - \alpha + h$  of the resource.

*Proof.* Consider an orthonormal frame of origin  $O$ . Let  $I$  the point of coordinates  $(1; 1)$ . For any  $h$ , the line of equation  $y = x - h$  is the one containing the points  $A(h; 0)$  and  $B(1; 1 - h)$ . The surface  $S_\mu$  is delimited by the segment  $[OI]$  and the curve of  $L_\mu$ .

**Proof of the lower bound & the lower bound is reached (fig. 2b).** Minimizing  $G(\mu)$  under the constraint  $H(\mu) = h$  is equivalent to minimizing the surface of the epigraph of  $L_\mu$  under the constraint that  $L_\mu$  touches the line  $[AB]$ .

Assume  $L_\mu$  fillfills the constraint. There exists  $\alpha \in [h, 1)$  such that  $L$  passes through the point  $C_\alpha$  with coordinates  $(\alpha; \alpha - h)$ . Hence, the surface  $S_\mu$  must be convex and contain the points  $O$ ,  $I$  and  $C_\alpha$ . Yet, the surface with minimal area containing these three points is their convex envelope, i.e. the triangle  $OC_\alpha I$ . The surface of this triangle is  $\frac{h}{2}$  whatever the chosen value of  $\alpha$ .

Hence,  $G(\mu)$  is minimized by any  $\mu$  such that  $L_\mu$  is piecewise affine, which graph is made of the segments  $[OC_\alpha]$  and  $[C_\alpha I]$ . This corresponds to the measures of form:

$$\mu_{\alpha,m} = \alpha \cdot \delta_{m \cdot (1 - \frac{h}{\alpha})} + (1 - \alpha) \cdot \delta_{m \cdot (1 + \frac{h}{1 - \alpha})}$$

for some  $m > 0$ . The proposition 2.4 ensures that there is no other measures with such a graph.

Eventually, the set

$$\{\mu_{\alpha,m} : \alpha \in [h, 1), m > 0\}$$

is the set of functions minimizing  $G(\mu)$  under the constraint  $H(\mu) = h$ ; for these functions, we have  $G(\mu_{\alpha,m}) = h$ .

**Proof of the upper bound & proof that the upper bound is never reached (fig. 2c).**

If  $H(\mu) = h$ , then  $L_\mu$  must be contained within the trapezoid  $OABI$ . Thus, the area of  $S_\mu$  must be at most the area of  $OABI$ , i.e.  $h - \frac{h^2}{2}$ . Hence,  $G(\mu) \leq 2h - h^2$ .

We now prove that this value cannot be reached. As  $L_\mu$  is continuous at 1, there exist  $\eta > 0$  such that for all  $x \geq 1 - \eta$ ,  $L_\mu(x) > 1 - \frac{h}{2}$ . Let  $R$  be the rectangle:

$$R = [1 - \eta, 1] \times \left[1 - h, 1 - \frac{h}{2}\right].$$

We have that:

- The epigraph  $\Gamma_\mu$  cannot intersect  $R$ , so  $S_\mu \cap R = \emptyset$ .
- But  $R$  is contained in the (full) trapezoid  $OABI$ .

This implies that the surface of  $S_\mu$  is at most the surface of  $OABI$ , minus the surface of  $R$ , i.e.:

$$G(\mu) \leq 2h - h^2 - \eta h.$$

That is, the upper bound is never reached.

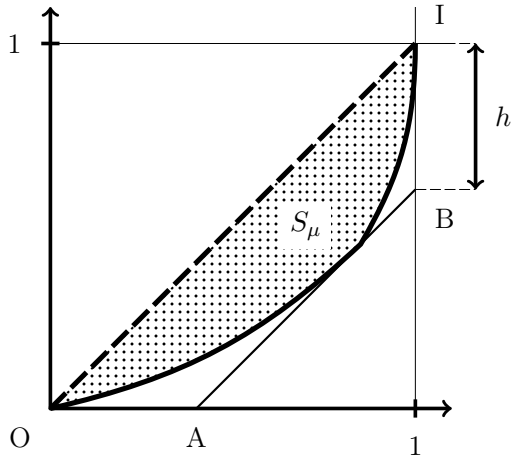
**Proof that every value between the upper bound and the lower bound is reached (fig. 2d).** In order to achieve any value in the interval  $(h, 2h - h^2)$ , we just propose a three-group distribution such that the Lorenz curve is composed of three segments  $[OA]$ ,  $[AC_\alpha]$  and  $[C_\alpha I]$  where  $C_\alpha$  has coordinates  $(\alpha; \alpha - h)$ , with  $h \leq \alpha < 1$  is to be determined (note that for  $h = \alpha$ ,  $C_h$  and  $A$  are combined). Such a  $\mu$  exists by proposition 2.4.

Some elementary geometry ensures that the area of the quadrilateral  $OAC_\alpha I$  is:

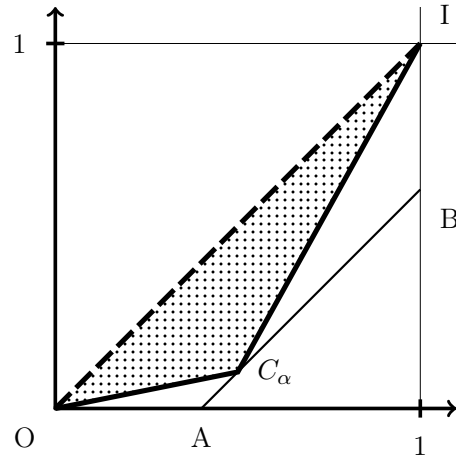
$$\mathcal{A}_\alpha = \frac{1}{2}(h + \alpha h - h^2).$$

The Gini index of the distribution is  $2\mathcal{A}_\alpha$ . Thus,  $G(\mu)$  reaches every value of  $h + \alpha h - h^2$  for  $h \leq \alpha < 1$ , i.e.  $G(\mu)$  reaches every value of the interval  $[h, 2h - h^2]$ . ■

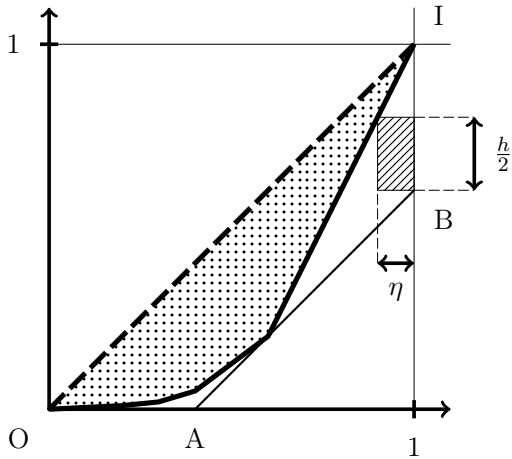




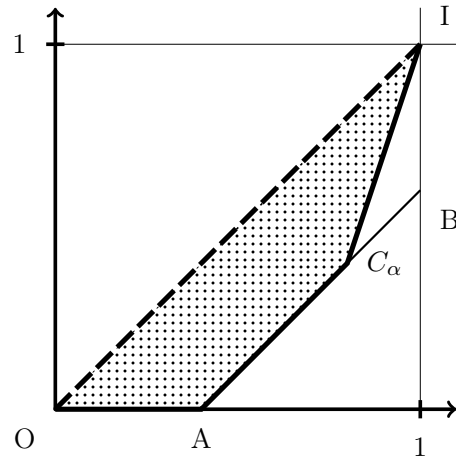
(a) To match the condition  $H(\mu) = h$ , the curve  $L_\mu$  must touch the line  $[AB]$  without crossing it. The dotted surface  $S_\mu$ , delimited by the line  $[OI]$  and the Lorenz curve  $L_\mu$ , has area  $\frac{1}{2}G(\mu)$ .



(b) To minimize the Gini index, the Lorenz curve must minimize the surface  $S_\mu$  while still touching the line  $[AB]$ . This condition is checked if and only if  $L_\mu$  touches the line in a single point  $C_\alpha$  and if surface  $S_\mu$  is the convex envelope of the points  $O$ ,  $C_\alpha$  and  $I$ , i.e. the dotted triangle.



(c) The dotted surface  $S_\mu$  needs be contained within the trapezoid  $OABI$ , which area is  $h - \frac{h^2}{2}$ . However, it need not include the hatched rectangle of dimensions  $\eta \times \frac{h}{2}$ .



(d) If  $\alpha$  moves between  $A$  (included) and  $B$  (excluded), then the dotted surface takes every possible area between  $\frac{h}{2}$  (included) and  $h - \frac{h^2}{2}$  (excluded).

Figure 2: Illustration of the geometric arguments used in the proof of corollary 4.2.

In all figures, we fix some  $h \in (0, 1)$ . The points  $A$  and  $B$  have coordinates  $(h, 0)$  and  $(1, 1 - h)$ . The line  $(AB)$  has equation  $y = x - h$ . The point  $C_\alpha$  is the point of this line having absciss  $\alpha \in [h, 1)$ .

## Part II

# Convergence properties of Lorenz curve, Gini index and Hoover index

## 5 Convergence of $L$ , $G$ and $H$ when the measures converge in $W_1$

It has been known for a long time that if  $\mu$  is a random variable and  $X_1, \dots, X_n$  is a random sample of  $\mu$  (i.i.d. variables of distribution  $\mu$ ), then the *empirical Gini index*  $G((X_i)_i)$  converges to the *theoretical Gini index*  $G(\mu)$ . Under some good hypotheses of integrability, we do even get the asymptotic normality of  $\sqrt{n}(G((X_i)_{1 \leq i \leq n}) - G(\mu))$  [Hoe48, section 9.b].

We do not intend to explore the asymptotic distributions of empirical indexes. However, we do not restrict on convergence of random samples: we will determine necessary and sufficient conditions for having uniform convergence of the Lorenz curves as soon as the underlying distributions converge.

### 5.1 $W_1$ distance, weak convergence and uniform integrability

Recall that a collection  $(U_i)_{i \in I}$  of real random variables defined on a space  $(\Omega, \mathcal{F}, \mathbf{P})$  is said *uniformly integrable (u.i.)* if

$$\sup_{i \in I} \mathbf{E}[|U_i| \mathbf{1}_{|U_i| > \alpha}] \xrightarrow{\alpha \rightarrow +\infty} 0.$$

It is known (see [Bil95, theorem 16.14, corollary]) that if  $U$  is a random variable,  $\mathbf{E}[|U_n - U|] \xrightarrow{n \rightarrow \infty} 0$  if and only if  $U_n \xrightarrow{n \rightarrow \infty} U$  in distribution and  $(U_n)_{n \in \mathbb{N}}$  is u.i.

Define the *Wasserstein-1 metric*  $W_1$  on  $\mathcal{M}_1(\mathbf{R}_+)$  as:

$$W_1(\mu, \nu) := \int_0^1 |Q_\mu(t) - Q_\nu(t)| dt.$$

Convergence with respect to metric  $W_1$  is equivalent to  $L^1$  convergence of quantile functions.

We say that  $(\mu_i)_{i \in \mathbb{N}}$  is *uniformly integrable* if the collection  $(Q_{\mu_i})_{i \in \mathbb{N}}$  is uniformly integrable

in the measure space  $([0, 1], \mathfrak{B}, \mathcal{L})$ . Keep in mind that this notion of uniform integrability of measures is nonstandard; most authors do only define uniform integrability for measurable functions or for random variables.

Finally, we say  $(\mu_n)_{n \in \mathbb{N}}$  *converges weakly* to  $\mu_\infty$  if  $F_{\mu_n}(x) \xrightarrow[n \rightarrow \infty]{} F_{\mu_\infty}(x)$  for all  $x$  at which  $F_{\mu_\infty}$  is continuous. We note  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathscr{W}} \mu_\infty$ . This is equivalent to having  $Q_{\mu_n} \xrightarrow[n \rightarrow \infty]{} Q_{\mu_\infty}$ ,  $\mathcal{L}$ -almost surely.

The pertinence of these concepts and the elementary properties are detailed in appendices B.2, B.3 and B.4.

### Theorem 5.1 (Scheffé–Lebesgue).

Let  $(\mu_n)_{n \in \mathbb{N}} \in \mathbf{M}^{\mathbb{N}}$  and  $\mu_\infty \in \mathbf{M}$ . The following are equivalent:

- (i)  $\mu_n \xrightarrow[n \rightarrow \infty]{W_1} \mu_\infty$ ;
- (ii)  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathscr{W}} \mu_\infty$  and  $m_{\mu_n} \xrightarrow[n \rightarrow \infty]{} m_{\mu_\infty}$ ;
- (iii)  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathscr{W}} \mu_\infty$  and  $(\mu_n)_{n \in \mathbb{N}}$  is uniformly integrable.

For a detailed proof, see propositions B.18 and B.25 in appendix.

### 5.2 Necessary and sufficient conditions for convergence of Lorenz curves

**Proposition 5.2.** Let  $(\mu_n)_{n \in \mathbb{N}}$  and  $\mu_\infty$  measures in  $\mathbf{M}$ . We denote  $L_\bullet$  their Lorenz functions and  $m_\bullet$  their means. The following assertions are equivalent:

- (i)  $\mu_n \xrightarrow[n \rightarrow \infty]{W_1} \mu_\infty$ ;

(ii)  $L_n \xrightarrow[n \rightarrow \infty]{} L_\infty$  pointwise over  $[0, 1]$  and  $m_{\mu_n} \xrightarrow[n \rightarrow \infty]{} m_{\mu_\infty}$ ;

(iii)  $L_n \xrightarrow[n \rightarrow \infty]{} L_\infty$  uniformly over  $[0, 1]$  and  $m_{\mu_n} \xrightarrow[n \rightarrow \infty]{} m_{\mu_\infty}$ .

We split the proof in two part.

*Proof of the easy implications.* (i)  $\implies$  (ii) is trivial using definition 1.1, triangle inequality and the definition of the  $W_1$  metric.

Implication (ii)  $\implies$  (iii) is a direct application of Dini's lemma, that we hereby recall. ■

**Lemma 5.3 (Dini).** *Let  $[a, b]$  be a segment. Let for  $n \in \mathbf{N}$ ,  $f_n$  and  $f_\infty : K \rightarrow \mathbf{R}$ . Assume:*

(i)  $f_n \xrightarrow[n \rightarrow \infty]{} f_\infty$  pointwise;

(ii)  $f_n$ 's are nondecreasing functions;

(iii)  $f_\infty$  is continuous on  $[a, b]$ .

Then  $f_n \xrightarrow[n \rightarrow \infty]{\|\cdot\|_\infty} f_\infty$  (uniform convergence on  $[a, b]$ ).

(See [PS78, problem 127] for a proof.)

Now the implication (iii)  $\implies$  (i) remains to be proven. To do so, we introduce the following lemma.

**Lemma 5.4.** *Let  $I$  an open interval of  $\mathbf{R}$ . Let  $F_n : I \rightarrow \mathbf{R}$  convex functions that uniformly converge to  $F_\infty$  over  $I$ . Let  $\delta_-$  the left-derivative of a function.*

$\partial_- F_n(x) \xrightarrow[n \rightarrow \infty]{} \partial_- F_\infty(x)$  pointwise for every  $x$  at which  $\partial_- F_\infty$  is continuous.

*Proof.* First, for all  $\varepsilon > 0$ , there exists  $N_\varepsilon \in \mathbf{N}$  such that for every  $n \geq N_\varepsilon$ ,  $\|F_n - F_\infty\|_\infty < \varepsilon^2$ .

Let, for every  $n \in \mathbf{N} \cup \{\infty\}$  and  $a, b, p \in I$  and  $p \in I$ ,

$$S_n(a, b) := \frac{F_n(b) - F_n(a)}{b - a}.$$

Notice that  $F_\infty$ , as a uniform limit of convex functions, is itself convex.

To ease the notations, let  $f_n := \partial_- F_n$  for every  $n \in \mathbf{N} \cup \{\infty\}$ .

**Minoring  $\liminf_n f_n(p)$ .** First, fix  $\varepsilon > 0$ . for every  $n \geq N_\varepsilon$ , we have:

$$\begin{aligned} & |S_n(p - \varepsilon, p) - S_\infty(p - \varepsilon, p)| \\ &= \frac{|F_n(p - \varepsilon) - F_n(p) - F_\infty(p - \varepsilon) + F_\infty(p)|}{\varepsilon} \\ &\leq \frac{|F_n(p - \varepsilon) - F_\infty(p - \varepsilon)|}{\varepsilon} + \frac{|F_n(p) - F_\infty(p)|}{\varepsilon} \\ &\leq \frac{\varepsilon^2}{\varepsilon} + \frac{\varepsilon^2}{\varepsilon} = 2\varepsilon. \end{aligned}$$

Hence, by convexity of  $F_n$ ,

$$S_\infty(p - \varepsilon, p) - 2\varepsilon \leq S_n(p - \varepsilon, p) \leq f_n(p).$$

Furthermore, we have by definition of  $f_\infty$ :

$$S_\infty(p - h, p) - 2h \xrightarrow[h > 0]{h \rightarrow 0} f_\infty(p) - 0.$$

Hence, for every  $\delta > 0$ , there exists  $\varepsilon_\delta > 0$  such that

$$S_\infty(p - \varepsilon_\delta) - 2\varepsilon_\delta > f_\infty(p) - \delta.$$

So for every  $\delta > 0$ , there exists  $N := N_{\varepsilon_\delta}$  such that  $\forall n \geq N_{\varepsilon_\delta}$ ,

$$f_n(p) > f_\infty(p) - \delta.$$

**Majoring  $\limsup_n f_n(p)$ .** Similarly, for all  $\varepsilon > 0$ , for every  $n \geq N_\varepsilon$ ,

$$|S_n(p, p + \varepsilon) - S_\infty(p, p + \varepsilon)| \leq 2\varepsilon.$$

By convexity of  $F_n$ ,

$$f_n(p) = \partial_- F_n(p) \leq S_n(p, p + \varepsilon).$$

Hence,

$$f_n(p) \leq S_\infty(p, p + \varepsilon) + 2\varepsilon.$$

By convexity of  $F_\infty$ , for all  $\eta > 0$ :

$$\partial_- F_\infty(p) \leq \partial_+ F_\infty(p) \leq \partial_- F_\infty(p + \eta).$$

Thus, since  $f_\infty$  is continuous at  $p$ ,

$$\partial_+ F_\infty(p) = \partial_- F_\infty(p^+) = \partial_- F_\infty(p) = f_\infty(p).$$

Hence,

$$S_\infty(p, p + h) + 2h \xrightarrow[h > 0]{h \rightarrow 0} f_\infty(p).$$

Hence, for all  $\delta > 0$ , there exists  $\varepsilon'_\delta$  such that  $S_\infty(p, p + \varepsilon'_\delta) + 2\varepsilon'_\delta < f_\infty(p) + \delta$ . Finally, for every  $n \geq N_{\varepsilon'_\delta}$ ,

$$f_n(p) < f_\infty + \delta. \quad \blacksquare$$

*Proof of the remaining part of theorem 5.2.*

Assume hypothesis (iii). The functions  $L_n$  are convex. As stated in section 1.1, for all  $p \in (0, 1]$  and for every  $n \in \mathbf{N} \cup \{\infty\}$ ,  $\partial_- L_{\mu_n}(p) = Q_{\mu_n}(p)/m_{\mu_n}$

As  $L_{\mu_n} \xrightarrow[n \rightarrow \infty]{} L_{\mu_\infty}$  uniformly, lemma 5.4 ensures that the sequence of functions  $(Q_{\mu_n}/m_{\mu_n})_{n \in \mathbf{N}}$  converges to  $Q_{\mu_\infty}/m_{\mu_\infty}$  pointwise over  $(0, 1)$ , except at the points of discontinuity of  $Q_{\mu_\infty}$ . The convergence also stands at 0.

As we assumed  $m_{\mu_n} \xrightarrow[n \rightarrow \infty]{} m_{\mu_\infty}$ , we get  $Q_{\mu_n} \xrightarrow[n \rightarrow \infty]{} Q_{\mu_\infty}$  pointwise over  $(0, 1)$ , except in the points of discontinuity of  $Q_\infty$ . Hence,  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathcal{W}} \mu_\infty$ . By Scheffé–Lebesgue, (i) stands. ■

### 5.3 Sufficient condition for Gini and Lorenz convergence

**Proposition 5.5.** *Let  $(\mu_n)_{n \in \mathbf{N}}$  and  $\mu_\infty$  measures of  $\mathbf{M}$ . If  $\mu_n \xrightarrow[n \rightarrow \infty]{W_1} \mu$ , then:*

1.  $L_{\mu_n} \xrightarrow[n \rightarrow \infty]{} L_{\mu_\infty}$  uniformly over  $[0, 1]$ .
2.  $G(\mu_n) \xrightarrow[n \rightarrow \infty]{} G(\mu_\infty)$ .
3.  $H(\mu_n) \xrightarrow[n \rightarrow \infty]{} H(\mu_\infty)$ .

*Proof.* Theorem 5.2 ensures that  $(L_{\mu_n})_{n \in \mathbf{N}}$  converges to  $L_{\mu_\infty}$  uniformly on  $[0, 1]$ . So does  $(\text{id} - L_{\mu_n})_{n \in \mathbf{N}}$  to  $\text{id} - L_{\mu_\infty}$ , where  $\text{id} : x \in [0, 1] \mapsto x$ .

Thus,

$$\int_0^1 p - L_{\mu_n}(p) \, dp \xrightarrow[n \rightarrow \infty]{} \int_0^1 p - L_{\mu_\infty}(p) \, dp ;$$

$$\max_{p \in [0, 1]} p - L_{\mu_n}(p) \xrightarrow[n \rightarrow \infty]{} \max_{p \in [0, 1]} p - L_{\mu_\infty}(p).$$

We conclude the proof by applying theorem 3.3 and theorem 3.10. ■

### 5.4 Topological formulation: homeomorphism between $(\mathbf{M}, W_1)$ and $\mathfrak{L} \times \mathbf{R}_+^*$

#### 5.4.1 Basic result

Recall that a topological space is called *sequential* if its sequentially closed sets are closed. A

mapping from a sequential space to any topological space is continuous iff it is sequentially continuous. Being first countable is a sufficient condition for being sequential; thus, metric spaces are sequential.

Now, we embed:

- $\mathbf{R}_+^*$  with its standard norm  $|\cdot|$ ;
- $\mathbf{M}$  with the metric  $W_1$ ;
- $\mathfrak{L}$  with the uniform norm  $\|\cdot\|_\infty$ .

$\mathbf{R}_+^*$ ,  $\mathbf{M}$  and  $\mathfrak{L}$ , embedded with the underlying topologies, are sequential spaces. Hence, propositions 2.4 and 5.2 can be reworded as follow:

**Theorem 5.6.** *The following mapping is a homeomorphism:*

$$\begin{aligned} \Phi : (\mathbf{M}, W_1) &\approx (\mathfrak{L}, \|\cdot\|_\infty) \times (\mathbf{R}_+^*, |\cdot|) \\ \mu &\mapsto (L_\mu, m_\mu). \end{aligned}$$

#### 5.4.2 $W_1$ topology on $\mathbf{M}/\mathbf{R}_+^*$

Now, let's focus on the quotient space  $\mathbf{M}/\mathbf{R}_+^*$  of distributions on  $(\mathbf{R}_+, \mathfrak{B})$  modulo equality up to a rescaling. Let  $\pi : \mathbf{M} \rightarrow \mathbf{M}/\mathbf{R}_+^*$  the canonical projection.  $\mathbf{M}/\mathbf{R}_+^*$  can be endowed with the quotient  $W_1$  topology:  $U$  is open in  $(\mathbf{M}/\mathbf{R}_+^*, W_1)$  iff  $\pi^{-1}\langle U \rangle$  is open in  $(\mathbf{M}, W_1)$ .

**A description of the  $W_1$  topology on  $\mathbf{M}/\mathbf{R}_+^*$ .** For  $\alpha > 0$  and  $\mu \in \mathbf{M}$ , call  $S_\alpha(\mu)$  the distribution of  $\alpha \cdot X$  where  $X \sim \mu$ . For  $\varepsilon > 0$  and  $\mu \in \mathbf{M}$ , call  $B(\mu, \varepsilon)$  the  $W_1$ -open ball of center  $\mu$  and radius  $\varepsilon$ .

Let  $U$  a  $W_1$ -open subset of  $\mathbf{M}$ . Call  $U' = \pi^{-1}\langle \pi\langle U \rangle \rangle$  its saturate for  $\equiv$ . If  $\mu \in U'$ , there exists  $\mu_0 \in U$  and  $\alpha > 0$  such that  $\mu = S_\alpha(\mu_0)$ . Since  $U$  is open, there exists  $\varepsilon > 0$  such that  $B(\mu_0, \varepsilon) \subseteq U$ . Now consider the open ball  $V' := B(\mu, \alpha\varepsilon)$ . It suffices to prove that  $V' \subseteq U'$ . For  $\nu \in V'$ , let  $\nu_0 = S_{1/\alpha}(\nu)$ . We have (see lemma B.6 in appendix for details):

$$\begin{aligned} W_1(\nu_0, \mu_0) &= \int_0^1 |Q_{\nu_0}(p) - Q_{\mu_0}(p)| \, dp \\ &= \int_0^1 \left| \frac{1}{\alpha} Q_\nu(p) - \frac{1}{\alpha} Q_\mu(p) \right| \, dp \\ &= \frac{1}{\alpha} W_1(\mu, \nu) \end{aligned}$$

$$W_1(\nu_0, \mu_0) < \varepsilon$$

thus  $\nu_0 \in B(\mu_0, \varepsilon) \subseteq U$ , so  $\nu \in U'$ . Hence,  $V' \subseteq U'$ . Hence,  $U'$  is open.

This proves that  $\pi : \mathbf{M} \rightarrow \mathbf{M}/\mathbf{R}_+^*$  is an open mapping. Hence, the  $W_1$  topology on  $\mathbf{M}/\mathbf{R}_+^*$  is generated by the elementary opens  $\pi\langle B(\mu_0, \varepsilon) \rangle$  for  $\mu_0 \in \mathbf{M}$  and  $\varepsilon > 0$ .

Notice that the quotient  $W_1$  metric need not induce the quotient  $W_1$  topology. In fact, it does not; one can check that the quotient  $W_1$  metric is a trivial pseudometric  $d([\mu], [\nu]) = 0$  for all  $\mu$  and  $\nu$ .

**Quotient mapping  $[\mu] \mapsto L_\mu$ .** The mapping  $\Psi : \mu \in \mathbf{M} \mapsto L_\mu$  is continuous and goes to the quotient. Hence, it induces a continuous mapping:  $\tilde{\Psi} : [\mu] \in \mathbf{M}/\mathbf{R}_+^* \mapsto L_\mu$ . One can directly check that its reciprocal function is given by:

$$\tilde{\Psi}^{-1}(\ell) = [\Phi^{-1}(\mu, 1)]$$

which is also continuous.

Finally, we have:

**Proposition 5.7.** *The following mappings are homeomorphisms:*

$$\begin{aligned} \tilde{\Psi} : \mathbf{M}/\mathbf{R}_+^* &\approx \mathfrak{L} \\ [\mu] &\mapsto L_\mu \\ \mathbf{M} &\approx \mathbf{M}/\mathbf{R}_+^* \times \mathbf{R}_+^* \\ \mu &\mapsto ([\mu], m_\mu) \\ \forall \alpha \in \mathbf{R}_+^*, \quad \mathbf{M}_\alpha &\approx \mathbf{M}/\mathbf{R}_+^* \\ \mu &\mapsto [\mu] \\ \forall \mu \in \mathbf{M}, \quad [\mu] &\approx \mathbf{R}_+^* \\ \nu &\mapsto m_\nu \end{aligned}$$

where  $\mathbf{M}$ ,  $\mathbf{M}_\alpha$ ,  $[\mu]$  and  $\mathbf{M}/\mathbf{R}_+^*$  are embedded with the  $W_1$  topology or the topology induced by it;  $\mathfrak{L}$  is embedded with the  $\|\cdot\|_\infty$  norm and  $\mathbf{R}_+^*$  is embedded with its standard topology.

### 5.4.3 Embedding $\mathfrak{L}$ with the pointwise convergence topology

What can we say if  $\mathfrak{L}$  is not embedded with the topology of uniform convergence, but pointwise

convergence, i.e. the topology induced by the product topology on  $\mathbf{R}^{[0,1]}$ ? As the pointwise convergence topology is not sequential, *a priori* proposition 5.2 tells nothing.

However, we can adapt Dini's theorem to prove that for given  $a, b \in \mathbf{R}$ , the  $\|\cdot\|_\infty$  topology and the product topology on  $\mathbf{R}^{[a,b]}$  induce the same topology on the subset of continuous nondecreasing functions.

**Proposition 5.8 (generalization of Dini).** *Let  $[a, b] \subseteq \mathbf{R}$ ,  $E := \mathbf{R}^{[a,b]}$  and  $\mathcal{C}^\uparrow \subseteq E$  the set of continuous, nondecreasing functions.*

*Let  $\tau_\infty$  the topology on  $E$  induced by the norm  $\|\cdot\|_\infty$  and  $\tau_\times$  the product topology on  $E$ . Let  $\tau'_\infty$  (resp.  $\tau'_\times$ ) the trace topology induced by  $\tau_\infty$  (resp.  $\tau_\times$ ) on  $\mathcal{C}^\uparrow$ .*

*We have  $\tau'_\infty = \tau'_\times$ .*

*Proof.* For  $f \in E$  and  $\varepsilon > 0$ , let  $B(f, \varepsilon)$  the  $\|\cdot\|_\infty$ -ball of center  $f$  and radius  $\varepsilon$ .

**$\tau_\infty$  is thinner than  $\tau_\times$ .** Let  $n \in \mathbf{N}^*$  and for  $i = 1, \dots, n$ :  $x_i \in [a, b]$ ,  $y_i \in \mathbf{R}$  and  $\varepsilon_i > 0$ . Let

$$U := \{g \in E : \forall i \in [1, n], |g(x_i) - y_i| < \varepsilon_i\}$$

an elementary open set of  $\tau_\times$ . Let  $g_0 \in U$ . for every  $i = 1, \dots, n$ , let  $a_i = g_0(x_i) - (y_i - \varepsilon_i)$  and  $b_i = (y_i + \varepsilon_i) - g_0(x_i)$ . Let:

$$m = \min\{a_i, b_i : i = 1, \dots, n\}.$$

Then  $g_0 \in B(g_0, m) \subseteq U$ , so  $U$  is  $\tau_\infty$ -open.

It follows that  $\tau_\infty$  is thinner than  $\tau_\times$ . Hence,  $\tau'_\infty$  is also thinner than  $\tau'_\times$ .

**$\tau'_\times$  is thinner than  $\tau'_\infty$ .** First, notice that the topology  $\tau'_\infty$  admits as a basis the trace balls of form  $B(f, \varepsilon) \cap \mathcal{C}^\uparrow$  with  $\varepsilon > 0$  and  $f \in \mathcal{C}^\uparrow$ .

Now consider a trace ball  $B(f, \varepsilon) \cap \mathcal{C}^\uparrow$  with  $f \in \mathcal{C}^\uparrow$ . It suffices to prove that there exists an elementary open  $U \in \tau_\times$  such that:

$$f \in U \quad \text{and} \quad U \cap \mathcal{C}^\uparrow \subseteq B(f, \varepsilon).$$

To do so, we use the same argument as for Dini's theorem. By Heine–Cantor theorem,  $f$  is uniformly continuous. Let  $\eta > 0$  such that for all  $x, y \in [a, b]$ ,  $|x - y| < \eta \implies |f(x) - f(y)| < \frac{\varepsilon}{5}$ .

Let  $a_0 := a < a_1 < \dots < a_n := b$  with  $a_i - a_{i-1} < \varepsilon$  for every  $i = 1, \dots, n$ . Let  $U$  the elementary  $\tau_\times$ -open:

$$\left\{ g \in E : \forall i \in [0, n], |g(a_i) - f(a_i)| < \frac{\varepsilon}{5} \right\}.$$

It is immediate that  $f \in U$ . Now assume  $g \in U \cap \mathcal{C}^\uparrow$ . Let  $x \in [a, b]$ . There exists  $i \in [0, n-1]$  such that  $a_i \leq x \leq a_{i+1}$ . We have:

$$\begin{aligned} & |g(x) - f(x)| \\ & \leq |g(x) - g(a_i)| + |g(a_i) - f(a_i)| \\ & \quad + |f(a_i) - f(x)| \\ & < |g(a_{i+1}) - g(a_i)| + \frac{2\varepsilon}{5} \\ & \leq |g(a_{i+1}) - f(a_{i+1})| + |f(a_{i+1}) - f(a_i)| \\ & \quad + |f(a_i) - g(a_i)| + \frac{2\varepsilon}{5} \end{aligned}$$

## 6 Applications of the $W_1$ convergence

The main practical interest of the proposition 5.5 is that it justifies that the Lorenz curves, the Gini and the Hoover are consistent with small perturbations and approximations. In the following section, we will show that the  $W_1$  convergence occurs in several cases, allowing us to apply proposition 5.5.

Let  $\mu \in \mathbf{M}$ . With increasing complexity, we deal with the following situations:

1.  $\mu$  is perturbed by a noise, which vanishes.
2.  $\mu$  is approximated by a random sample, and the size  $n$  of the sample increases.
3.  $\mu$  is approximated by a discrete distribution corresponding to  $\ell$  regularly-chosen quantiles, and  $\ell$  grows.
4.  $\mu$  is approximated by  $\ell$  regularly chosen quantiles of a sample of size  $n$ , as  $\ell$  and  $n$  grow. This corresponds to a more realistic situation: in practice, the public offices for statistics do only publish the quantiles of the population's revenue, based on a sample.
5.  $\mu$  is approximated with the kernel density estimate of a sample of size  $n$  with window

$< \varepsilon$ . ■

**Consequence.** Since  $\mathfrak{L} \subseteq \mathcal{C}^\uparrow$ , the theorem 5.6 remains valid if  $\mathfrak{L}$  is embedded with pointwise convergence topology.

### 5.4.4 $G$ and $H$ as continuous applications

To prove proposition 5.5, we merely used the fact that the mappings  $L_\mu \mapsto G(\mu)$  and  $L_\mu \mapsto H(\mu)$  are continuous. Thus, the mappings  $G : \mathbf{M} \rightarrow [0, 1)$  and  $H : \mathbf{M} \rightarrow [0, 1)$  are continuous, where  $\mathbf{M}$  is embedded with the  $W_1$  topology. Since  $G$  and  $H$  go to the quotient, the quotient mappings  $\tilde{G}$  and  $\tilde{H} : \mathbf{M}/\mathbf{R}_+^* \rightarrow [0, 1)$  are continuous too.

$\varepsilon$ , as  $n$  grows and  $\varepsilon$  shrinks.

**Notation.** In what follows, the notation  $\otimes$  refers to the tensor product of measures. Recall that if  $X \sim \mu$  and  $Y \sim \nu$  are independent, then  $(X, Y) \sim \mu \otimes \nu$ . Likewise,  $(X_n)_{n \in \mathbf{N}} \sim \mu^{\otimes \mathbf{N}}$  iff the  $X_n$ 's are i.i.d. random variables with distribution  $\mu$ , etc.

### 6.1 Vanishing noise

Consider a probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ .

Let  $X$  a nonnegative  $L^1$  random variable. Let  $(Y_n)_{n \in \mathbf{N}}$  a sequence of real, variables converging to 0  $\mathbf{P}$ -almost surely, such that  $(\mathbf{E}[|Y_n|])_{n \in \mathbf{N}}$  is bounded, and let for  $n \in \mathbf{N}$ :

$$Z_n := \max(X + Y_n, 0).$$

By dominated convergence theorem,  $\mathbf{E}[|Z_n - X|] \xrightarrow[n \rightarrow \infty]{} 0$ . Since  $L^1$  convergence of random variables implies  $W_1$  convergence of the underlying measures (see proposition B.18 in appendix), one can apply proposition 5.5. We have  $L_{Z_n} \xrightarrow[n \rightarrow \infty]{} L_X$  uniformly,  $G(Z_n) \xrightarrow[n \rightarrow \infty]{} G(X)$  and  $H(Z_n) \xrightarrow[n \rightarrow \infty]{} H(X)$ .

In particular, this is the case if  $Y_n$  represents a “noise” that decreases. (We cut in 0 to avoid dealing with negative values.)



For instance:

**Application 6.1.** Let  $(\Omega, \mathcal{F}, \mathbf{P})$  a probability space. Let  $X$  a nonnegative random variable with  $X \sim \mu \in \mathbf{M}$ .

Let  $Y \in L^1(\Omega, \mathcal{F}, \mathbf{P})$  and  $(\varepsilon_n)_{n \in \mathbf{N}}$  a series with limit 0. Let:

$$Z_n := \max(X + \varepsilon_n Y, 0)$$

and  $\nu_n$  the distribution of  $Z_n$ . Then:

$$1. \nu_n \xrightarrow[n \rightarrow \infty]{W_1} \mu.$$

$$2. L_{\nu_n} \xrightarrow[n \rightarrow \infty]{\|\cdot\|_\infty} L_\mu.$$

$$3. G(\nu_n) \xrightarrow[n \rightarrow \infty]{} G(\mu).$$

$$4. H(\nu_n) \xrightarrow[n \rightarrow \infty]{} H(\mu).$$

## 6.2 Sampling

Recall the following the following, fundamental theorem, proven in [Bil95, theorem 20.6].

**Theorem 6.2 (Glivenko–Cantelli).** Let  $\mu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$  and  $(X_n)_{n \in \mathbf{N}} \sim \mu^{\otimes \mathbf{N}}$ . Let  $\hat{\mu}_n$  be the empirical measure of  $(X_1, \dots, X_n)$ . Then, almost surely,  $\|F_{\hat{\mu}_n} - F_\mu\|_\infty \xrightarrow[n \rightarrow \infty]{} 0$ , where  $F_\bullet$  denotes the cumulative distribution functions associated to measures.

**Application 6.3.** Let  $\mu \in \mathbf{M}$ . Let  $X_1, \dots, X_n, \dots$  i.i.d with distribution  $\mu$ . Let  $\hat{\mu}_n$  the empirical measure of the subsample  $(X_1, \dots, X_n)$ . Then, almost surely,

$$1. \hat{\mu}_n \xrightarrow[n \rightarrow \infty]{W_1} \mu.$$

$$2. L_{\hat{\mu}_n} \xrightarrow[n \rightarrow \infty]{\|\cdot\|_\infty} L_\mu.$$

$$3. G(\hat{\mu}_n) \xrightarrow[n \rightarrow \infty]{} G(\mu).$$

$$4. H(\hat{\mu}_n) \xrightarrow[n \rightarrow \infty]{} H(\mu).$$

*Proof.* From Glivenko–Cantelli, we infer that  $\hat{\mu}_n \xrightarrow[n \rightarrow \infty]{} \mu$  weakly. Furthermore, by the strong law of large numbers,  $m_{\hat{\mu}_n} \xrightarrow[n \rightarrow \infty]{} m_\mu$ . Then, by Scheffé’s lemma,  $\hat{\mu}_n \xrightarrow[n \rightarrow \infty]{W_1} \mu$ . The three other assertions follow from proposition 5.5. ■

The results can be reworded in terms of consistency of estimators.

**Corollary 6.4.** Let  $\mu \in \mathbf{M}$ . Let  $X_1, \dots, X_n, \dots$  i.i.d with distribution  $\mu$ . Let  $\hat{\mu}_n$  the empirical measure of the subsample  $(X_1, \dots, X_n)$ . Then, almost surely,

(i) The following estimator of  $G_\mu$  is strongly consistent:

$$\widehat{G}_n := \frac{\sum_{i=1}^n \sum_{j=1}^n |X_i - X_j|}{2n \sum_{i=1}^n X_i}.$$

(ii) The following estimator of  $H_\mu$  is strongly consistent:

$$\widehat{H}_n := \frac{\sum_{i=1}^n \left| X_i - \frac{1}{n} \sum_{j=1}^n X_j \right|}{2 \sum_{i=1}^n X_i}.$$

(iii) The estimators  $L_n^{(x)}$  of  $L_\mu(x)$  are strongly consistent, uniformly in  $x \in [0, 1]$ , where

$$\widehat{L}_n^{(x)} := \frac{\sum_{i=1}^{nx} \left( X_{(1:n)}^\uparrow \right)_i}{\sum_{i=1}^n X_i},$$

with the following notations:

- $\left( X_{(1:n)}^\uparrow \right)_i$  is the  $i$ -th term of the subsequence  $(X_1, \dots, X_n)$ , reordered increasingly;
- $\sum_{i=1}^0 a_i = 0$ ;
- if  $x = k + f$  with  $k \in \mathbf{N}$  and  $0 < f < 1$ , then

$$\sum_{i=1}^x a_i := f \sum_{i=1}^{k+1} a_i + (1-f) \sum_{i=1}^k a_i.$$

**Uniform integrability of empirical measures.** The following result can be directly deduced from the  $W_1$  convergence of  $(\hat{\mu}_n)_{n \in \mathbf{N}}$ .

**Corollary 6.5.** Let  $\mu \in \mathbf{M}$  and  $(X_n)_{n \in \mathbf{N}^*}$  i.i.d. with distribution  $\mu$ . Let  $\hat{\mu}_n$  the empirical measure of  $(X_1, \dots, X_n)$ . Then, almost surely, the collection of measures  $(\hat{\mu}_n)_{n \in \mathbf{N}^*}$  is uniformly integrable.

An alternate proof is given in appendix [D.3.1](#).

### 6.3 Quantile approximation

For  $\ell \in \mathbf{N}^*$ , the  $\ell$ -quantile approximation of  $\mu$  is a discretization  $\tilde{\mu}_\ell$  that has  $\ell$  atoms in the 0th, 1st,  $\dots$ ,  $(\ell - 1)$ th  $\ell$ -quantiles of  $\mu$ , each with probability  $\frac{1}{\ell}$ .

**Lemma 6.6.** Let  $\mu \in \mathbf{M}$  and  $\ell \in \mathbf{N}^*$ . Consider the  $\ell$ -quantile approximation of  $\mu$ :

$$\tilde{\mu}_\ell := \frac{1}{\ell} \sum_{k=0}^{\ell-1} \delta_{Q_\mu(\frac{k}{\ell})}.$$

For all  $x \in \mathbf{R}_+$ ,

$$F_\mu(x) \leq F_{\tilde{\mu}_\ell}(x) \leq F_\mu(x) + \frac{1}{\ell}.$$

In other words,  $\mu$  dominates  $\hat{\mu}_\ell$  at first order and  $\|F_{\tilde{\mu}_\ell} - F_\mu\|_\infty \leq \frac{1}{\ell}$ .

*Proof.* Let  $x \in \mathbf{R}_+$ . Consider the unique integer  $k_0 \in [0, \ell]$  such that  $\frac{k_0}{\ell} \leq F_\mu(x) < \frac{k_0+1}{\ell}$ . We have:

$$\begin{aligned} F_{\tilde{\mu}_\ell}(x) &= \tilde{\mu}_\ell([0, x]) \\ &= \frac{1}{\ell} \sum_{k=0}^{\ell-1} \delta_{Q_\mu(\frac{k}{\ell})}([0, x]) \\ &= \frac{1}{\ell} \sum_{k=0}^{\ell-1} \mathbf{1}_{Q_\mu(\frac{k}{\ell}) \leq x} \\ F_{\tilde{\mu}_\ell}(x) &= \frac{1}{\ell} \sum_{k=0}^{\ell-1} \mathbf{1}_{\frac{k}{\ell} \leq F_\mu(x)} \end{aligned}$$

The inequality  $\frac{k}{\ell} \leq F_\mu(x)$  is true for  $k = 0, \dots, k_0$ , and false otherwise. Hence,

$$F_{\tilde{\mu}_\ell}(x) = \frac{k_0 + 1}{\ell}. \quad \blacksquare$$

**Application 6.7.** Let  $\mu \in \mathbf{M}$ . for every  $\ell \in \mathbf{N}^*$ , let

$$\tilde{\mu}_\ell := \frac{1}{\ell} \sum_{k=1}^{\ell} \delta_{Q_\mu(\frac{k}{\ell})}.$$

Then  $\tilde{\mu}_\ell \xrightarrow[\ell \rightarrow \infty]{W_1} \mu$ , and the conclusions of proposition [5.5](#) stand.

*Proof.* Lemma [6.6](#) ensures that  $F_{\tilde{\mu}_\ell} \xrightarrow[\ell \rightarrow \infty]{} F_\mu$  pointwise. Thus,  $\tilde{\mu}_\ell \xrightarrow[\ell \rightarrow \infty]{} \mu$  weakly.

Furthermore, by lemma [6.6](#), for every  $\ell \in \mathbf{N}^*$ ,  $F_\mu \leq F_{\tilde{\mu}_\ell}$ . It follows that  $Q_{\tilde{\mu}_\ell} \leq Q_\mu$  (see proposition [B.7](#) in appendix for details).

$Q_\mu$  is integrable over  $[0, 1)$  and has finite integral  $m_\mu$  (proposition [B.5](#)). Thus, the collection  $(Q_{\tilde{\mu}_\ell})_{\ell \in \mathbf{N}^*}$  is uniformly dominated by a  $L^1$  random variable; hence it is uniformly integrable.

By Scheffé–Lebesgue  $\tilde{\mu}_\ell \xrightarrow[n \rightarrow \infty]{W_1} \mu$ . ■

### 6.4 Quantile-of-sample approximation

In the application [6.7](#), the uniform integrability followed from the facts that all  $Q_{\tilde{\mu}_\ell}$  were dominated by the same  $L^1$  random variable. In fact, this result can be easily generalized: a collection of random random variables  $(X_j)_{j \in J}$  is u.i. as soon as there exists a collection of u.i. variables  $(Y_i)_{i \in I}$  and for every  $j \in J$ , there exists an  $i$  such that  $X_j \leq Y_i$  a.s. (just write it — or see proposition [B.24](#) in appendix for detailed proof).

**Notation.** If  $(u_{n,m})$  is a double-indexed sequence of numbers, we write:

$$u_{n,m} \xrightarrow[n \rightarrow \infty \perp\!\!\!\perp m \rightarrow \infty]{} u_\infty$$

if for every mappings  $\varphi, \psi : \mathbf{N} \rightarrow \mathbf{N}$  with limit  $\infty$  we have:

$$u_{\varphi(i), \psi(i)} \xrightarrow[i \rightarrow \infty]{} u_\infty.$$

*Mutatis mutandis* for other modes of convergence (weak or  $W_1$  convergence of measure, uniform or pointwise convergence of functions, etc.). *Mutatis mutandis* for convergence in the neighbourhood of a real number, etc.

**Application 6.8 (quantiles of sample).** Let  $\mu \in \mathbf{M}$  and  $(X_n)_{n \in \mathbf{N}} \sim \mu^{\otimes \mathbf{N}}$  random variables defined on a probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ . Let  $\hat{\mu}_n$  the empirical measure of  $(X_1, \dots, X_n)$ . For every  $\ell \in \mathbf{N}^*$ , let:

$$\hat{\mu}_{n,\ell} = \frac{1}{\ell} \sum_{k=0}^{\ell-1} \delta_{Q_{\hat{\mu}_n}(\frac{k}{\ell})}.$$

Then,  $\mathbf{P}$ -almost surely,

$$\hat{\mu}_{n,\ell} \xrightarrow[n \rightarrow \infty \perp \ell \rightarrow \infty]{W_1} \mu$$

and the conclusions of proposition 5.5 stands.

Notice that  $\hat{\mu}_{n,\ell}$  is the quantile approximation (with  $\ell$  quantiles) of the empirical distribution  $\hat{\mu}_n$ .

*Proof.* We note  $F_\bullet$  the cumulative distribution functions of the considered measures.

By lemma 6.6 applied to measure  $\hat{\mu}_n$ , for each  $n \in \mathbf{N}^*$ ,  $\|F_{\hat{\mu}_{n,\ell}} - F_{\hat{\mu}_n}\| \leq \frac{1}{\ell}$ . By Glivenko–Cantelli,  $\|F_{\hat{\mu}_n} - F_\mu\|_\infty$  converges to 0  $\mathbf{P}$ -a.s. as  $n \rightarrow \infty$ . Hence  $\mathbf{P}$ -a.s.,  $\|F_{\hat{\mu}_{n,\ell}} - F_\mu\|_\infty \rightarrow 0$ , i.e.

$$\hat{\mu}_{n,\ell} \xrightarrow[n \rightarrow \infty \perp \ell \rightarrow \infty]{\mathcal{W}} \mu.$$

Furthermore, by lemma 6.6, for each  $n \in \mathbf{N}^*$ ,  $\ell \in \mathbf{N}^*$ , the measure  $\hat{\mu}_{n,\ell}$  is stochastically dominated by  $\hat{\mu}_n$ . As we proved that  $(\hat{\mu}_n)_{n \in \mathbf{N}}$  is u.i. (corollary 6.5), so is  $(\hat{\mu}_{n,\ell})_{n \in \mathbf{N}, \ell \in \mathbf{N}^*}$ . Hence, by theorem 5.1, the  $\mathbf{P}$ -a.s. convergence of  $\hat{\mu}_{n,\ell}$  is  $W_1$ . ■

## 6.5 Kernel density estimation of a sample

The kernel density estimate (KDE), or Parzen–Rosenblatt estimate, is classical way of approximating a distribution using a sample, by convolving the empirical distribution thereof with a continuous random variable (the “density kernel”) with a scale parameter  $h$  (the bandwidth).

In what follows, we will prove that the Lorenz curve, Gini and Hoover indexes are consistent with the KDE under the following loose hypotheses:

- the kernel density is  $L^1$ ;
- the bandwidth  $h$  holds to 0 and the sample size  $n$  holds to  $+\infty$ , independently.

Note that to avoid tricks with negative values, we “cut” the kernel density estimate in zero. More precisely, we define:

**Definition 6.9 (cut-in-zero KDE).** Let  $\mu$  a measure over  $(\mathbf{R}, \mathfrak{B})$ ,  $K$  a density function over  $\mathbf{R}$ , and  $h > 0$ . Let  $(\Omega, \mathcal{F}, \mathbf{P})$  a probability space.

Let  $(X, Y)$  a couple of random variables over  $(\Omega, \mathcal{F}, \mathbf{P})$  with distribution  $\mu \otimes K\mathcal{L}$  (i.e.,  $X \sim \mu$ ,  $Y$  has density  $K$  with respect to the Lebesgue measure and  $X \perp\!\!\!\perp Y$ ).

We note  $\text{KDE}(\mu, K, h)$  the distribution of the random variable  $\max(X + hY, 0)$ .

In the most common case, when knowing a sample  $(X_1, \dots, X_n)$ , the original distribution  $\mu$  is approximated as  $\text{KDE}(\hat{\mu}_n, K, h)$  where  $\hat{\mu}_n$  is the empirical measure associated with the sample.

We first need two lemmas involving KDEs.

**Lemma 6.10.** Let  $\mu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$ ,  $h > 0$ ,  $K$  a density function over  $\mathbf{R}$  and  $G : x \mapsto \int_{-\infty}^x K(t) dt$ .

Then, for all  $t \in \mathbf{R}_+$ , the c.d.f. of  $\text{KDE}(\mu, K, h)$  can be written:

$$F_{\text{KDE}(\mu, K, h)}(t) = \int_{\mathbf{R}_+} G\left(\frac{t-x}{h}\right) d\mu(x).$$

*Proof.* We have for all  $t \in \mathbf{R}_+$ :

$$\begin{aligned} & F_{\text{KDE}(\mu, K, h)}(t) \\ &= \iint_{\substack{x \in \mathbf{R}_+ \\ y \in \mathbf{R}}} \mathbf{1}_{(\max(x+hy, 0) \leq t)} d\mu(x) K(y) dy \\ &= \iint_{\substack{x \in \mathbf{R}_+ \\ y \in \mathbf{R}}} \mathbf{1}_{(x+hy \leq t)} d\mu(x) K(y) dy \\ &= \int_{\mathbf{R}_+} \left( \int_{\mathbf{R}} \mathbf{1}_{(x+hy \leq t)} K(y) dy \right) d\mu(x). \end{aligned}$$

Let  $Y \sim K\mathcal{L}$  on a probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ . Then:

$$\begin{aligned} \int_0^\infty \mathbf{1}_{(x+hy \leq t)} K(y) dy &= \mathbf{P}(x + hY \leq t) \\ &= \mathbf{P}\left(Y \leq \frac{t-x}{h}\right) \\ \int_0^\infty \mathbf{1}_{(x+hy \leq t)} K(y) dy &= G\left(\frac{t-x}{h}\right). \quad \blacksquare \end{aligned}$$

**Lemma 6.11.** Let  $K : \mathbf{R} \rightarrow \mathbf{R}_+$  a density. For all  $\varepsilon > 0$ , there exists  $M_{K,\varepsilon} > 0$  such that for all  $t \in \mathbf{R}_+$ ,  $h \in \mathbf{R}$  and  $\mu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$ :

$$\begin{aligned} & |F_{\text{KDE}(\mu, K, h)}(t) - F_\mu(t)| \\ & < 2\varepsilon + 2\mu((t - M_{K,\varepsilon}h, t + M_{K,\varepsilon}h]) \end{aligned}$$

*Proof.* Let  $M \in \mathbf{R}_+$  such that  $\int_{-M}^M K(u) du > 1 - \varepsilon$ . We fix  $t, h$  and  $\mu$  for the rest of the proof of the lemma.

We note  $\nu := \text{KDE}(\mu, K, h)$ . Let  $G$  the c.d.f. associated with density  $K$ .

By lemma 6.10, write  $F_\nu(t) = A + B + C$  where:

$$\begin{aligned} A &:= \int_0^{(t-Mh)^+} G\left(\frac{t-x}{h}\right) d\mu(x); \\ B &:= \int_{(t-Mh)^-}^{(t+Mh)^+} G\left(\frac{t-x}{h}\right) d\mu(x); \\ C &:= \int_{(t+Mh)^-}^{\infty} G\left(\frac{t-x}{h}\right) d\mu(x). \end{aligned}$$

$G$  is nondecreasing, with values in  $[0, 1]$ , and we have  $G(-M) < \varepsilon$  and  $G(M) > 1 - \varepsilon$ . Hence, we can bound  $A, B$  and  $C$ :

- $(1 - \varepsilon)F_\mu(t - Mh) \leq A \leq F_\mu(t - Mh)$ , hence:

$$\begin{aligned} -\varepsilon &\leq -\varepsilon F_\mu(t - Mh) \\ &\leq A - F_\mu(t - Mh) \leq 0; \end{aligned}$$

- $0 \leq B \leq \mu((t - Mh, t + Mh])$ ;
- $0 \leq C \leq \varepsilon \mu((t + Mh, \infty]) \leq \varepsilon$ .

Thus:

$$\begin{aligned} & |F_\nu(t) - F_\mu(t)| \\ & \leq |F_\nu(t) - F_\mu(t - Mh)| \\ & \quad + |F_\mu(t - Mh) - F_\mu(t)| \\ & \leq |A + B + C - F_\mu(t - Mh)| \\ & \quad + \mu((t - Mh, t]) \\ & \leq |A - F_\mu(t - Mh)| + |B| + |C| \\ & \quad + \mu((t - Mh, t]) \\ & \leq \varepsilon + \mu((t - Mh, t + Mh]) + \varepsilon \\ & \quad + \mu((t - Mh, t]) \\ & \leq 2\varepsilon + 2\mu((t - Mh, t + Mh]). \end{aligned}$$

Hence, the lemma stands with  $M_{K,\varepsilon} := M$ .  $\blacksquare$

**Application 6.12.** Let  $\mu \in \mathbf{M}$  and  $K$  a probability density over  $\mathbf{R}$  which is  $L^1$ .

Let  $(\Omega, \mathcal{F}, \mathbf{P})$  a probability space,  $(X_n)_{n \in \mathbf{N}} \sim \mu^{\otimes \mathbf{N}}$ , and  $\hat{\mu}_n$  the empirical measure of  $(X_1, \dots, X_n)$ .

Then,  $\mathbf{P}$ -almost surely:

$$\text{KDE}(\hat{\mu}_n, K, h) \xrightarrow[n \rightarrow \infty \perp\!\!\!\perp h \rightarrow 0^+]{W_1} \mu$$

and the conclusions of proposition 5.5 stand.

*Proof.*  $F_\bullet$  denotes the cumulative distribution functions associated to measures.  $\mathbf{P}$ -almost surely,  $F_{\hat{\mu}_n} \xrightarrow[n \rightarrow \infty]{} F_\mu$  uniformly (Glivenko-Cantelli). We saw (corollary 6.5) that,  $\mathbf{P}$ -almost surely,  $(\hat{\mu}_n)_{n \in \mathbf{N}}$  is uniformly integrable.

Fix  $\omega \in \Omega$  such that both previous assertion stand.

**Weak convergence.** We prove that the distributions  $\text{KDE}(\hat{\mu}_n, K, h)$  weakly converge to  $\mu$  as  $h \rightarrow 0^+$  and  $n \rightarrow \infty$  independently.

Let  $t \in \mathbf{R}_+$  such that  $F_\mu$  is continuous at  $t$ . Let  $\varepsilon > 0$ . We chose  $N_\varepsilon > 0$  and  $H_{\varepsilon, t} > 0$  as follows:

- $N_\varepsilon$  is given by Glivenko-Cantelli, such that for every  $n \geq N_\varepsilon$ ,  $\|F_{\hat{\mu}_n} - F_\mu\|_\infty < \varepsilon$ .
- Let  $U_{\varepsilon, t} > 0$ , given by continuity of  $F_\mu$  at  $t$ , such that for all  $u$  with  $|u| < U_{\varepsilon, t}$ ,  $|F(t+u) - F(t)| < \varepsilon$ .

Thanks to lemma 6.11, there exists  $M_\varepsilon > 0$  such that for all  $\nu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$  and for all  $r > 0$ ,

$$\begin{aligned} & |F_{\text{KDE}(\nu, K, r)}(t) - F_\nu(t)| \\ & < 2\varepsilon + 2\nu((t - Mr, t + Mr]). \end{aligned}$$

We set:  $H_{\varepsilon, t} := U_{\varepsilon, t}/M_\varepsilon$ .

Now, let  $n \geq N_\varepsilon$ . We have:

$$\begin{aligned} & |F_{\text{KDE}(\hat{\mu}_n, K, h)}(t) - F_\mu(t)| \\ & \leq |F_{\text{KDE}(\hat{\mu}_n, K, h)}(t) - F_{\hat{\mu}_n}(t)| + |F_{\hat{\mu}_n}(t) - F_\mu(t)| \\ & \leq |F_{\text{KDE}(\hat{\mu}_n, K, h)}(t) - F_{\hat{\mu}_n}(t)| + \varepsilon. \end{aligned}$$

Assume  $0 < h < H_{\varepsilon, t}$ . By definition of  $H_{\varepsilon, t}$ :

$$\begin{aligned} & |F_{\text{KDE}(\hat{\mu}_n, K, h)}(t) - F_\mu(t)| \\ & \leq 3\varepsilon + 2\hat{\mu}_n((t - Mh, t + Mh]) \end{aligned}$$

$$\begin{aligned}
&= 3\varepsilon + 2F_{\hat{\mu}_n}(t + Mh) - 2F_{\hat{\mu}_n}(t - Mh) \\
&\leq 7\varepsilon + 2F_{\mu}(t + Mh) - 2F_{\mu}(t - Mh) \\
&\leq 11\varepsilon.
\end{aligned}$$

which proves the weak convergence.

**Uniform integrability.** Let  $X_n \sim \hat{\mu}_n$  ( $n \in \mathbf{N}$ ) and  $Y \sim K \mathcal{L}$  defined on a probability space  $(\Omega', \mathcal{F}', \mathbf{P}')$ .

The random variable  $Y$  is integrable. Hence, the collection  $\{hY : 0 < h \leq 1\}$  is u.i. Furthermore, the collection

$(\hat{\mu}_n)_{n \in \mathbf{N}^*}$  is u.i. Hence, by sum, the collection  $\{X_n + hY : n \in \mathbf{N}^*, 0 < h \leq 1\}$  is u.i. Eventually, the collection  $\{\max(X_n + hY, 0) : n \in \mathbf{N}^*, 0 < h \leq 1\}$  is u.i. too. This means that the collections of measures  $\{\text{KDE}(\hat{\mu}_n, K, h) : n \in \mathbf{N}^*, 0 < h \leq 1\}$  is u.i. (See appendix B.4.3 for details about operations on u.i. collections.)

**Conclusion.** By theorem 5.1, the  $W_1$  convergence holds. ■

## 7 Weaker assumptions

By Scheffé–Lebesgue,  $W_1$  convergence is equivalent to weak convergence plus (either uniform integrability or convergence of means). The goal of this section is to analyse what happens if only one of these hypotheses stands.

### 7.1 Weak convergence without $W_1$ convergence

#### 7.1.1 Weak convergence is not enough for convergence of $L$ , $G$ and $H$

We first give two counter-examples to illustrate the fact that, if  $\mu_n \xrightarrow[n \rightarrow \infty]{} \mu_\infty$  weakly but not in  $W_1$ , almost anything can happen to the Lorenz curves, Gini and Hoover indexes.

**Counter-example 7.1.** Let  $X_n$  random variables such that  $\mathbf{P}(X_n = 1) = 1 - \frac{1}{n^2}$  and  $\mathbf{P}(X_n = n^2) = \frac{1}{n^2}$ . Let  $\mu_n$  the distribution of  $X_n$ .

Then, Borel–Cantelli’s lemma states that the set of  $n$ ’s such that  $X_n \neq 1$  is finite, i.e.  $(X_n)_n$  holds to the deterministic random variable  $\mathbf{1}$   $\mathbf{P}$ -a.s. Thus,  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathcal{W}} \delta_1$  weakly. We have  $H(\delta_1) = G(\delta_1) = 0$  and  $L_{\delta_1} = \text{id}$ .

for every  $n \in \mathbf{N}$ , the expectation of  $X_n$  is  $m_{\mu_n} = \mathbf{E}[X_n] = 2 - \frac{1}{n^2}$ . Hence, the convergence of  $\mu_n$ ’s is not  $W_1$ .

Direct computations show that  $\mathbf{E}[|X_n - m_{\mu_n}|] \xrightarrow[n \rightarrow \infty]{} 2$ . Hence,  $H(\mu_n) \xrightarrow[n \rightarrow \infty]{} 1$  and  $G(\mu_n) \xrightarrow[n \rightarrow \infty]{} 1$ . Furthermore, the pointwise limit of  $L_{\mu_n}$ ’s is a discontinuous function  $\ell$  such that  $\ell(t) = 0$  for  $t < 1$  and  $\ell(1) = 1$ .

**Counter-example 7.2.** Consider the probability space  $(\Omega, \mathcal{F}, \mathbf{P}) := ([0, 1], \mathfrak{B}, \mathcal{L})$ .

For every  $n \in \mathbf{N}^*$  and  $\omega \in [0, 1)$ , let:

$$Y_n(\omega) = \begin{cases} 0 & \text{if } \omega \leq 0.5; \\ 1 & \text{if } 0.5 < \omega \leq 1 - \frac{1}{n^2}; \\ n^2 & \text{otherwise.} \end{cases}$$

The random variables  $Y_n$  converge  $\mathbf{P}$ -a.s. to a random variable  $Y_\infty$  such that  $Y_\infty(\omega) = 0$  if  $\omega \leq 0.5$ ,  $Y_\infty(\omega) = 1$  otherwise. Call  $\nu_n$  and  $\nu_\infty$  the underlying distributions.

We have:  $m_{\nu_\infty} = \mathbf{E}[Y_\infty] = 0.5$ . However,  $m_{\nu_n} = \mathbf{E}[Y_n] = \frac{3}{2} - \frac{1}{n^2}$ . Hence,  $\nu_n \xrightarrow[n \rightarrow \infty]{\mathcal{W}} \nu_\infty$  but not  $W_1$ .

$G(\nu_\infty) = H(\nu_\infty) = 0.5$ ,  $L_{\nu_\infty}(t) = 0$  if  $t \leq 0.5$  and  $L_{\nu_\infty}(t) = 2t - 1$  if  $t \geq 0.5$ .

However, direct computations show that  $G(\nu_n) \xrightarrow[n \rightarrow \infty]{} \frac{5}{6}$ ,  $H(\nu_n) \xrightarrow[n \rightarrow \infty]{} \frac{2}{3}$  and that  $L_{\nu_n}$  converges pointwise to a limit  $f$  such that  $f(t) = 0$  if  $t \geq 0.5$ ,  $f(t) = \frac{2}{3}t - \frac{1}{3}$  if  $0.5 \leq t < 1$  and  $f(1) = 1$ .

#### 7.1.2 Topological properties of $\Phi : (\mathbf{M}, \mathcal{W}) \longrightarrow \mathfrak{L} \times \mathbf{R}_+^*$

Now consider the bijection  $\Phi : \mathbf{M} \longrightarrow \mathfrak{L} \times \mathbf{R}_+^*$  of proposition 2.4, defined by  $\Phi(\mu) = (L_\mu, m_\mu)$ . Contrary to theorem 5.6, we embed  $\mathbf{M}$  with the topology of weak convergence  $\mathcal{W}$ .  $\mathcal{W}$  is at least as coarse than the topology induced by metric  $W_1$ . Hence,  $\Phi^{-1}$  is continuous. However, from the previous counterexamples, follows that  $\Phi$  is not continuous.

What can we say about the results of proposition 5.7 if  $\mathbf{M}$ , its subspaces and quotients are embedded with  $\mathscr{W}$  rather than with  $W_1$ ?

**Restrictions.** It is easy to see that some restrictions of  $\Phi$  are continuous, because  $W_1$  and  $\mathscr{W}$  induce the same topologies. For instance:

- Let  $\alpha > 0$ . Due to Scheffé's lemma the weak convergence of  $(\mu_n)_{n \in \mathbf{N}} \in \mathbf{M}_\alpha^{\mathbf{N}}$  to some  $\mu_\infty \in \mathbf{M}_\alpha$  is equivalent to  $W_1$  convergence.
- Let  $\mu \in \mathbf{M}_1$ . Let  $(\nu_n)_{n \in \mathbf{N}} \in [\mu]^{\mathbf{N}}$  weakly converging to some  $\nu_\infty \in [\mu]$ . We have  $Q_{\nu_n}(t) \xrightarrow[n \rightarrow \infty]{} Q_{\nu_\infty}(t)$  almost everywhere. For every  $n \in \mathbf{N} \cup \{\infty\}$ ,  $Q_{\nu_n} = m_{\nu_n} Q_{\mu_n}$ . As  $\{t \in [0, 1) : Q_\mu(t) \neq 0\}$  has strictly positive measure, it follows that  $m_{\nu_n} \xrightarrow[n \rightarrow \infty]{} m_{\nu_\infty}$ . Finally, due to Scheffé's lemma,  $\nu_n \xrightarrow[n \rightarrow \infty]{W_1} \nu_\infty$ .

Since  $W_1$  convergence always implies  $\mathscr{W}$  convergence and both topologies are metrizable, it follows that they induce the same topologies on  $\mathbf{M}_\alpha$  and on  $[\mu]$ .

**Quotient.** The quotient space  $\mathbf{M}/\mathbf{R}_+^*$  can be embedded with the quotient topology induced by  $\mathscr{W}$ , i.e. the thinnest topology making the mapping  $\mu \mapsto [\mu]$  continuous whence  $\mathbf{M}$  is embedded with  $\mathscr{W}$ .

This topology is coarser than the quotient topology induced by  $W_1$ . Hence, if  $\tilde{\Psi}$  is the quotient mapping:

$$\tilde{\Psi} : (\mathbf{M}/\mathbf{R}_+^*, \mathscr{W}) \longrightarrow \mathfrak{L} \\ [\mu] \longmapsto L_\mu$$

then  $\tilde{\Psi}^{-1}$  is continuous. However,  $\tilde{\Psi}$  is not continuous. This proves that the quotient topology induced by  $\mathscr{W}$  is strictly coarser than the quotient topology induced by  $W_1$ .

## 7.2 Convergence of means without weak convergence

In fact, in both previous counterexamples, the Lorenz curves converge in  $[0, 1]^{[0, 1]}$ , but the limit

is not a Lorenz curve itself, as it is not continuous at 1. On  $[0, 1)$ , they converge to a “shrunked” Lorenz curve. The following results explore what we may say about the limit behaviour of Lorenz curves in the neighbourhood of 1, with no assumption about weak convergence.

First, recall that if  $\ell_n$ 's are Lorenz curves and if they converge pointwise to any function  $\ell_\infty : [0, 1] \rightarrow \mathbf{R}$ , then  $\ell_\infty$  is nondecreasing, convex, continuous on  $[0, 1)$  with  $\ell_\infty(0) = 0$  and  $\ell_\infty(1) = 1$ . Furthermore,  $\ell_\infty$  is continuous at 1 if and only if  $\ell_\infty \in \mathfrak{L}$ , if and only if the convergence is uniform.

### 7.2.1 Pointwise convergence of Lorenz function implies weak convergences of the measures

The following proposition shows what may happen in the convergence in distribution holds.

**Proposition 7.3.** *Let  $(\mu_n)_{n \in \mathbf{N}} \in \mathbf{M}^{\mathbf{N}}$ . Let  $m_{\mu_n}$  the mean of  $\mu_n$ . Assume that  $(L_{\mu_n})$  converges pointwise to a function  $\ell : [0, 1] \rightarrow \mathbf{R}$  and  $(m_{\mu_n})$  converges to some limit  $\alpha \in \mathbf{R}_+$ . Then:*

1.  $(\mu_n)_{n \in \mathbf{N}}$  has a weak limit  $\mu_\infty$  in  $\mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$ .
2. The mean of  $\mu_\infty$  satisfies:

$$m_{\mu_\infty} = \ell(1^-) \cdot \alpha \leq \alpha < \infty.$$

3. If  $m_{\mu_\infty} > 0$ , i.e.  $\mu_\infty \in \mathbf{M}$ , then for all  $x \in [0, 1)$ :

$$\ell(x) = \ell(1^-) L_{\mu_\infty}(x).$$

*Proof.* Let  $m_n$ ,  $Q_n$  and  $L_n$  the mean, quantile function and Lorenz function of  $\mu_n$ .

**Case  $\alpha \cdot \ell(1^-) = 0$ .** First, assume that either  $\alpha = 0$  or  $\ell(1^-) = 0$ . In both cases,

$$\int_0^{1^-} Q_n(p) \, dp \xrightarrow[n \rightarrow \infty]{} 0$$

hence for all  $x \in [0, 1)$ ,

$$\int_0^x Q_n(p) \, dp \xrightarrow[n \rightarrow \infty]{} 0.$$



By Dini (lemma 5.3), the convergence is uniform over every interval of the form  $[0, b]$ ,  $b \in [0, 1)$ . By lemma 5.4,  $Q_n \xrightarrow[n \rightarrow \infty]{} 0$  pointwise over  $(0, b)$ . Hence,  $Q_n \xrightarrow[n \rightarrow \infty]{} 0$  pointwise over  $(0, 1)$ . The convergence also stands in 0. Thus the  $\mu_n$ 's weakly converge to the Dirac mass  $\delta_0$ , and the proposition stands.

**General case.** Now consider the function:

$$\begin{aligned} \tilde{\ell} : [0, 1] &\longrightarrow [0, 1] \\ x &\longmapsto \begin{cases} \frac{\ell(x)}{\ell(1^-)} & \text{if } x < 1 \\ 1 & \text{if } x = 1. \end{cases} \end{aligned}$$

$\tilde{\ell}$  is a convex, continuous function taking values 0 in 0 and 1 in 1. By proposition 2.4, there exists an unique measure  $\mu_\infty$  with mean  $\alpha \cdot \ell(1^-)$  such that  $\tilde{\ell}$  is the Lorenz function of  $\mu_\infty$ . It suffices to prove that  $(\mu_n)_{n \in \mathbf{N}}$  weakly converges to  $\mu_\infty$ . Let  $Q_\infty$  the quantile function of  $\mu_\infty$ .

For all  $x \in [0, 1)$ ,  $\frac{1}{\ell(1^-)} L_n(x) \xrightarrow[n \rightarrow \infty]{} \tilde{\ell}(x)$ . By Dini, the convergence is uniform over every compact of form  $[0, p]$ . Hence, by lemma 5.4, for every  $p \in (0, 1)$  such that  $Q_\infty = \partial_- \tilde{\ell}$  is continuous at  $p$ ,

$$\begin{aligned} \frac{1}{\ell(1^-)} \partial_- L_n(p) &\xrightarrow[n \rightarrow \infty]{} \partial_- \tilde{\ell}(p) \\ \frac{1}{\ell(1^-)} \frac{Q_n(p)}{m_n} &\xrightarrow[n \rightarrow \infty]{} \frac{Q_\infty(p)}{\ell(1^-) \cdot \alpha} \\ Q_n(p) &\xrightarrow[n \rightarrow \infty]{} Q_\infty(p). \end{aligned}$$

The convergence also stands for  $p = 0$ .

Finally,  $(\mu_n)_{n \in \mathbf{N}}$  weakly converges to  $\mu_\infty$ , which concludes the proof.  $\blacksquare$

## 7.2.2 Uniform integrability and the infimum of the Lorenz curves

Now we do not necessarily consider sequences of measures, but any collection. We have the following characterization:

**Proposition 7.4.** *Let  $(\mu_i)_{i \in I}$  a collection of measures in  $\mathbf{M}$ , with means  $m_i$  and Lorenz functions  $L_i$ .*

1. *If the pointwise  $\liminf_{i \in I} L_i$  is continuous at 1 and  $\sup_{i \in I} m_i < \infty$ , then  $(\mu_i)_{i \in I}$  is uniformly integrable.*

2. *If  $(\mu_i)_{i \in I}$  is uniformly integrable and  $\inf_{i \in I} m_i > 0$ , then the pointwise  $\inf_{i \in I} L_i$  is continuous at 1.*

Notice that to define the pointwise  $\liminf$  in point 2,  $I$  need not be countable. In all generality,  $\liminf L_i(x)$  is defined as the lowest accumulation point of the set  $\{L_i(x) : i \in I\}$ .

*Proof.* Recall that both  $\liminf_{i \in I} L_i$  and  $\inf_{i \in I} L_i$  take value 1 in 1.

**Continuity of pointwise liminf implies uniform integrability.** Let :

$$A := \sup_{i \in I} m_i < \infty.$$

Let  $L := \liminf_{i \in I} L_i$ . Let  $x < y$ . There exists  $(i_n)_{n \in \mathbf{N}}$  such that  $\lim_{n \in \mathbf{N}} L_{i_n}(y) = L(y)$ . Yet for every  $n \in \mathbf{N}$ ,  $L_{i_n}(x) \leq L_{i_n}(y)$ . Thus,

$$L(x) = \liminf_{i \in I} L_i(x) \leq \liminf_{n \in \mathbf{N}} L_{i_n}(x) \leq L(y).$$

Hence,  $L$  is nondecreasing.

Consider the probability space  $([0, 1], \mathfrak{B}, \mathcal{L})$ . Let  $x \in [0, 1)$ . For all  $i \in I$ , let  $Q_i$  the quantile function of  $\mu_i$ . We have:

$$\begin{aligned} (1-x) Q_i(x) &\leq \int_x^1 Q_i(x) dx \\ &\leq \int_0^1 Q_i(x) dx \leq A. \end{aligned}$$

Hence, we have for all  $u \in [0, 1]$ :

$$u \leq x \implies Q_i(u) \leq Q_i(x) \leq \frac{A}{1-x}. \quad (7.5)$$

Now let  $M > A$ . Taking the contraposition of (7.5) and letting  $x = 1 - \frac{A}{M}$ , we have:

$$Q(u) > M \implies u > 1 - \frac{A}{M}.$$

Thus, for all  $i \in I$ :

$$\begin{aligned} \mathbf{E}[Q_i \mathbf{1}_{Q_i > M}] &= \int_0^1 Q_i(u) \mathbf{1}_{Q_i(u) > M} du \\ &\leq \int_0^1 Q_i(u) \mathbf{1}_{(u > 1 - \frac{A}{M})} du \\ &= \int_{1 - \frac{1}{M}}^1 Q_i(u) du \end{aligned}$$

$$= m_i \left[ L_i(1) - L_i \left( 1 - \frac{A}{M} \right) \right]$$

$$\mathbf{E}[Q_i \mathbf{1}_{Q_i > M}] \leq A \left[ 1 - L_i \left( 1 - \frac{A}{M} \right) \right].$$

Hence, to prove that  $(\mu_i)_i$  is u.i., it suffices to prove that the rightmost side is arbitrarily small for  $M$  big enough, uniformly in  $i$ . Let  $\varepsilon > 0$ .

Since  $L$  is continuous at 1, there exists  $M > 0$  such that  $L \left( 1 - \frac{A}{M} \right) > 1 - \frac{\varepsilon}{A}$ . Furthermore, by definition of  $\liminf$ , there exists  $J \subseteq I$  such that  $I \setminus J$  is finite and for all  $j \in J$ ,

$$L_j \left( 1 - \frac{A}{M} \right) > L \left( 1 - \frac{A}{M} \right) - \frac{\varepsilon}{A}.$$

Then in particular, for all  $x > 1 - \frac{A}{M}$ , for all  $j \in J$ ,  $L_j(x) > 1 - 2\frac{\varepsilon}{A}$ .

Hence, for all  $j \in J$ ,  $\mathbf{E}[Q_j \mathbf{1}_{Q_j > M}] < 2\varepsilon$ , and  $I \setminus J$  is finite. This is enough to show that the  $\mu_i$ 's are u.i. (see lemma B.23 in appendix).

**Uniform integrability implies continuity of the pointwise infimum.** Let  $A := \sup_{i \in I} \frac{1}{m_i} < \infty$ .

Let  $\varepsilon > 0$ . Notice that for all  $M > 0$ ,  $x \in [0, 1)$  and  $i \in I$ , we have:

$$\int_x^1 Q_i(t) dt$$

$$= \int_x^1 Q_i(t) \mathbf{1}_{Q_i(t) > M} dt + \int_x^1 Q_i(t) \mathbf{1}_{Q_i(t) \leq M} dt$$

$$\leq \int_x^1 Q_i(t) \mathbf{1}_{Q_i(t) > M} dt + (1-x)M$$

As the  $Q_i$ 's are uniformly integrable, there exists  $M$  such that the first term is less than  $\varepsilon$ . Let  $\eta = \frac{\varepsilon}{M}$ . If  $x > 1 - \eta$ , then we have:

$$\int_x^1 Q_i(t) dt \leq 2\varepsilon.$$

## Acknowledgements

The author thanks Hippolyte d'Albis, Guillaume Conchon-Kerjean, David Leturcq and Anne Perrot, whose advices during the writing of this article were precious.

Furthermore, for all  $i \in I$ ,

$$1 - L_i(x) = \frac{1}{m_i} \int_x^1 Q_i(t) dt \leq 2A\varepsilon.$$

Finally,  $1 - \inf_{i \in I} L_i(x) \leq 2A\varepsilon$ . This proves that  $\inf_{i \in I} L_i$  is continuous at 1. ■

Notice that for all  $x \in [0, 1]$ ,

$$\inf_{i \in I} L_i(x) \leq \liminf_{i \in I} L_i(x) \leq 1.$$

Hence, if  $\inf L_i$  is continuous at 1, so is  $\liminf L_i$ . In other words, we have the following (less strong, but more elegant) characterization:

**Corollary 7.6.** *Let  $(\mu_i)_{i \in I}$  a collection of measures in  $\mathbf{M}$  with means  $m_i$  and Lorenz functions  $L_i$ . Assume that both  $m_i$ 's and  $1/m_i$ 's are bounded. The following are equivalent :*

- (i)  $(\mu_i)_{i \in I}$  is uniformly integrable;
- (ii) The pointwise infimum function  $\inf_{i \in I} L_i$  is continuous at 1;
- (iii) The pointwise liminf function  $\liminf_{i \in I} L_i$  is continuous at 1.

### 7.2.3 Conclusion

The following corollary can be directly deduced from each of the propositions 7.3 and 7.4 and the use of Scheffé–Lebesgue and proposition 2.4.

**Corollary 7.7.** *Assume  $(\mu_n)_{n \in \mathbf{N}} \in \mathbf{M}^{\mathbf{N}}$  is uniformly integrable,  $(L_{\mu_n})_{n \in \mathbf{N}}$  converges pointwise to some  $\ell \in [0, 1]^{[0, 1]}$  and  $(m_{\mu_n})_{n \in \mathbf{N}}$  converges to some  $\alpha \in \mathbf{R}_+^*$ .*

*Then  $\ell \in \mathfrak{L}$  and  $\mu_n \xrightarrow[n \rightarrow \infty]{W_1} \mu_\infty$ , where  $\mu_\infty = \Phi^{-1}(\ell, \alpha)$  is the only distribution with Lorenz curve  $\ell$  and mean  $\alpha$ .*

## References

- [Arn87] Barry C. Arnold. *Majorization and the Lorenz Order: A Brief Introduction*. Lecture Notes in Statistics 43. Springer-Verlag, 1987. ISBN: 978-0-3879-6592-5. DOI: [10.1007/978-1-4615-7379-1](https://doi.org/10.1007/978-1-4615-7379-1).
- [Atk70] Anthony B. Atkinson. “On the measurement of inequality”. In: *Journal of economic theory* 2.3 (1970), pp. 244–263. DOI: [10.1016/0022-0531\(70\)90039-6](https://doi.org/10.1016/0022-0531(70)90039-6).
- [Bil68] Patrick Billingsley. *Convergence of Probability Measures*. 1st edition. Wiley, 1968. ISBN: 978-0-4710-7242-3.
- [Bil95] Patrick Billingsley. *Probability and Measure*. 3rd edition. Wiley Series in Probability and Mathematical Statistics. Wiley, 1995. ISBN: 978-0-471-00710-4.
- [Bil99] Patrick Billingsley. *Convergence of Probability Measures*. 2nd edition. Wiley, 1999. ISBN: 978-0-4711-9745-4.
- [Cho08] Duangkamon Chotikapanich, ed. *Modeling Income Distributions and Lorenz Curves*. Economic Studies in Inequality, Social Exclusion, and Well-Being. Springer, 2008. ISBN: 978-0-387-72796-7. DOI: [10.1007/978-0-387-72796-7](https://doi.org/10.1007/978-0-387-72796-7).
- [CV11] Lidia Ceriani and Paolo Verme. “The origins of the Gini index: extracts from *Variabilità e Mutabilità* (1912) by Corrado Gini”. In: *Journal of Economic Inequalities* 10 (2011), pp. 421–443. DOI: [0.1007/s10888-011-9188-x](https://doi.org/10.1007/s10888-011-9188-x).
- [Dor79] Robert Dorfman. “A Formula for the Gini Coefficient”. In: *The Review of Economics and Statistics* 61.1 (1979), pp. 146–149. DOI: [10.2307/1924845](https://doi.org/10.2307/1924845).
- [DSS73] Partha Dasgupta, Amartya Sen, and David Starrett. “Notes on the measurement of inequality”. In: *Journal of Economic Theory* 6.2 (1973), pp. 180–187. DOI: [10.1016/0022-0531\(73\)90033-1](https://doi.org/10.1016/0022-0531(73)90033-1).
- [FRK78] John C.H. Fei, Gustav Ranis, and Shirley W.Y. Kuo. “Growth and the family distribution of income by factor components”. In: *The Quarterly Journal of Economics* 92.1 (1978), pp. 17–53. DOI: [10.2307/1885997](https://doi.org/10.2307/1885997).
- [Gas71] Joseph L. Gastwirth. “A general definition of the Lorenz curve”. In: *Econometrica* 36.9 (Nov. 1971), pp. 1037–1039. DOI: [10.2307/1909675](https://doi.org/10.2307/1909675).
- [Gas72] Joseph L. Gastwirth. “The estimation of the Lorenz curve and Gini index”. In: *The review of Economics and Statistics* 54 (Aug. 1972), pp. 306–316. DOI: [10.2307/1937992](https://doi.org/10.2307/1937992).
- [Gin05] Corrado Gini. *On the measurement of concentration and variability of characters*. Ed. by Giovanni Maria Giorgi. Trans. by Fulvio De Santis. Vol. LXIII. 2005, pp. 1–38.
- [Gin12] Corrado Gini. *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche*. Studi economico-giuridici pubblicati per cura della facoltà di Giurisprudenza della R. Università di Cagliari. 1912.
- [Gin14] Corrado Gini. *Sulla misura della concentrazione e della variabilità dei caratteri*. Vol. LXXIII.II. 1914, pp. 1203–1248.
- [HMB61] Stanisław Hartman, Jan Mikusiński, and Leo F. Boron. *The Theory of Lebesgue Measure and Integration*. English edition. Elsevier Science, 1961. ISBN: 978-1-483-28033-2.
- [Hoe48] Wassily Hoeffding. “A Class of Statistics with Asymptotically Normal Distribution”. In: *The Annals of Mathematical Statistics* 19.3 (1948), pp. 293–325. DOI: [10.1214/aoms/1177730196](https://doi.org/10.1214/aoms/1177730196).
- [KS45] Sir Maurice George Kendall and Alan Stuart. *The advanced theory of statistics 1*. 2nd edition. Charles Griffin, London, 1945. ISBN: 978-0-340-61430-3.
- [Le 22] Jean-François Le Gall. *Measure Theory, Probability, and Stochastic Processes*. 1st edition. Graduate Texts in Mathematics. Springer Cham, 2022. ISBN: 978-3-031-14204-8.
- [MOA11] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: Theory of Majorization and Its Applications*. 2nd edition. Springer Series in Statistics. Springer, 2011. ISBN: 978-0-3874-0087-7. DOI: [10.1007/978-0-387-68276-1](https://doi.org/10.1007/978-0-387-68276-1).
- [New70] David Newbery. “A theorem on the measurement of inequality”. In: *Journal of Economic Theory* 2.3 (1970), pp. 264–266. DOI: [10.1016/0022-0531\(70\)90040-2](https://doi.org/10.1016/0022-0531(70)90040-2).
- [Pik14] Thomas Piketty. *Capital in the Twenty-First Century*. Trans. by Arthur Goldhammer. Harvard University Press, 2014. ISBN: 978-8-9378-3458-5.

- [PS78] George Polya and Gabor Szegő. *Problems and theorems in analysis I*. Classics in Mathematics. Springer, 1978. ISBN: 978-3-540-63640-3.
- [Sen73] Amartya Sen. *On Economic Inequality*. Enlarged edition. Radcliffe lectures. Clarendon, Oxford, 1973. ISBN: 978-0-198-28193-1.
- [She72] Eytan Sheshinski. “Relation between a social welfare function and the Gini index of income inequality”. In: *Journal of Economic Theory* 4.1 (1972), pp. 98–100. DOI: [10.1016/0022-0531\(72\)90167-6](https://doi.org/10.1016/0022-0531(72)90167-6).
- [Tho76] W. A. Thompson Jr. “Fisherman’s luck”. In: *Biometrics* 32.2 (1976), pp. 265–271. DOI: [10.2307/2529497](https://doi.org/10.2307/2529497).
- [Vil03] Cédric Villani. *Topics in Optimal Transportation*. Vol. 58. Graduate Studies in Mathematics. American Mathematical Society, 2003. ISBN: 978-0-8218-3312-4.

# Appendices

## A Details of notations and basic definitions

The notations of this article are chosen to be consistent with [Le 22].

Let  $\mathbf{R}_+$  the set of nonnegative real numbers,  $\mathbf{R}_+^*$  the set of strictly positive numbers, and  $\overline{\mathbf{R}}_+$  the set  $\mathbf{R}_+ \cup \{\infty\}$ .

**Elementary notations.** If  $S$  is a set, we note  $\mathbf{1}_S$  the indicator function of  $S$ .  $\text{id}_S$  is the identity function of  $S$ . In the absence of ambiguity, we simply write  $\text{id}$ .

Let  $A$  an assertion. We note  $\mathbf{1}_{(A)} = 1$  if  $A$  is true, 0 otherwise. If no ambiguity, we simply write  $\mathbf{1}_A$ .

If  $f : A \rightarrow B$  is a mapping,  $\alpha \subseteq A$  and  $\beta \subseteq B$ , we note  $f\langle\alpha\rangle$  the direct image of  $\alpha$  and  $f^{-1}\langle\beta\rangle$  the inverse image of  $\beta$  through  $f$ .

If  $f$  is a real-valued function defined over a real interval  $I$  and  $x \in I$ , we note  $f(x^-)$  (resp.  $f(x^+)$ ) the left-limit (resp. right-limit) of  $f$  in  $x$  (if exists). We note  $\partial_- f(x)$  (resp.  $\partial_+ f(x)$ ) the left derivative (resp. right derivative) of  $f$  in  $x$  (if exists).

**Lebesgue integrals.** If  $\mathcal{X}$ , is a topological space, we note  $\mathfrak{B}(\mathcal{X})$  the  $\sigma$ -algebra of Borel sets of  $I$  (i.e. the  $\sigma$ -algebra generated by the open sets of  $\mathcal{X}$ ). If  $I$  is a subset of  $\mathbf{R}$ , we embed  $I$  with the trace topology induced by the standard topology of  $\mathbf{R}$ ; then  $\mathfrak{B}(I)$  is the trace  $\sigma$ -algebra of  $\mathfrak{B}(\mathbf{R})$  on  $I$ . We call  $\mathcal{L}_I : \mathfrak{B}(I) \rightarrow \mathbf{R}_+$  the Lebesgue measure over the measurable space  $(I, \mathfrak{B}(I))$ . If no ambiguity, we simply note  $\mathfrak{B}$  and  $\mathcal{L}$ .

For any Borel measure  $\nu$  over  $I$  and any measurable function  $f : I \rightarrow \mathbf{R}_+$ , if  $a, b \in I \cup \{-\infty, +\infty\}$  with  $a \leq b$ , we denote:

$$\int_a^{b^-} f(x) \, d\nu(x) := \int_{[a,b)} f \, d\nu = \int_I \mathbf{1}_{[a,b)} f \, d\nu$$

*mutatis mutandis* for intervals of form  $(a, b]$ ,  $[a, b]$  and  $(a, b)$ . If  $f$  is defined on the interval  $[a, b)$  but not in  $b$  and if no confusion is possible,

we allow to write:

$$\int_a^b f(x) \, d\nu(x) := \int_{[a,b)} f \, d\nu.$$

When the integral is computed with respect to the Lebesgue measure, we drop the symbol  $\mathcal{L}$ , i.e. we simply write:

$$\int_a^b f(x) \, dx := \int_{[a,b)} f \, d\mathcal{L}.$$

### Measures and probability distributions.

Let  $(\Omega, \mathcal{F})$  a measurable space. We note  $\mathcal{M}(\Omega, \mathcal{F})$  the set of measures over it, and  $\mathcal{M}_1(\Omega, \mathcal{F})$  the subset of probability measures, i.e. the measures  $\mu$  such that  $\mu(\Omega) = 1$ . If no ambiguity on the measurable space, we may simply write  $\mathcal{M}$  and  $\mathcal{M}_1$ .

Let  $(\Omega', \mathcal{F}')$  another measurable space,  $f : (\Omega, \mathcal{F}) \rightarrow (\Omega', \mathcal{F}')$  a measurable function,  $\nu \in \mathcal{M}(\Omega, \mathcal{F})$ . Then we note  $f_{\#}(\nu)$  the *push-forward of  $\nu$  through  $f$* , i.e. the measure over  $(\Omega', \mathcal{F}')$  such that for all  $A \in \mathcal{F}'$ ,  $f_{\#}(\nu)(A) = \nu(f^{-1}\langle A \rangle)$ .

Let  $\mu \in \mathcal{M}(\Omega, \mathcal{F})$ . A point  $\omega \in \Omega$  is called an *atom* of  $\mu$  if  $\mu(\{\omega\}) > 0$ . A measure  $\mu$  is called *nonatomic* or *diffuse* if it has no atom.

If  $\omega \in \Omega$ , we note  $\delta_\omega$  the *Dirac mass in  $\omega$* , i.e. the measure such that for any  $A \in \mathcal{F}$ ,  $\delta_\omega(A) := \mathbf{1}_A(\omega)$ .

Let  $(\Omega, \mathcal{F}, \mathbf{P})$  a probability space and  $(\mathcal{X}, d)$  a metric space. Fix any  $x_0 \in \mathcal{X}$ . For  $p > 0$ , a random variable  $\Omega \rightarrow \mathcal{X}$  is said to be  $L^p$  if  $\int_\Omega d(x_0, X)^p \, d\mathbf{P} < \infty$  (this does not depend on the choice of  $x_0$ ).

Let  $X : (\Omega, \mathcal{F}) \rightarrow (A, \mathcal{G})$  a random variable. If  $\mu \in \mathcal{M}_1(A, \mathcal{G})$ , then we note  $X \sim \mu$  when the distribution of  $X$  is  $\mu$ , i.e. when  $X_{\#}(\mathbf{P}) = \mu$ . Recall that in such case, the *law of the unconscious statistician* [Le 22, proposition 8.5] states that for every measurable (resp.  $\mu$ -integrable) function  $f : A \rightarrow \overline{\mathbf{R}}_+$  (resp.  $f : A \rightarrow \mathbf{R}$ ):

$$\mathbf{E}[f(X)] = \int_\Omega f(X(\omega)) \, d\mathbf{P}(\omega) = \int_A f \, d\mu.$$

**Cumulative distribution functions.** Let  $\mu$  a probability distribution over  $(\mathbf{R}_+, \mathfrak{B})$ . We note  $F_\mu : \mathbf{R}_+ \rightarrow [0, 1)$  the *cumulative distribution function (c.d.f.)* of  $\mu$ , i.e. the function such that for all  $x \in \mathbf{R}_+$ ,

$$F_\mu(x) := \mu([0, x]).$$

Recall this function is nondecreasing, right-continuous. Its left-discontinuities correspond to the atoms of  $\mu$ . It admits a limit equal to 1 in  $+\infty$ .

If  $\mu$  and  $\nu$  are two measures over  $(\mathbf{R}_+, \mathfrak{B})$  such that  $F_\mu = F_\nu$ , then they match over all the sets of form  $[0, x]$ . Yet, the class of sets of form  $[0, x]$  is closed under finite intersections and generates  $\mathfrak{B}(\mathbf{R}_+)$  as a  $\sigma$ -algebra. Furthermore, the class of the sets  $A$  such that  $\mu(A) = \nu(A)$  is a monotone class. Hence, the monotone class theorem [Le 22, theorem 1.18, p. 13] ensures that  $\mu$  and  $\nu$  match on *every* set. In other words,  $F_\mu$  fully characterizes  $\mu$ .

**First-order stochastic domination.** Let  $\mu, \nu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$ . We say that  $\mu$  *stochastically dominates  $\nu$  at first-order* (FSD) if for all  $x \in \mathbf{R}_+$ ,  $F_\nu(x) \geq F_\mu(x)$ .

Notice that most authors add an hypothesis that  $F_\nu$  and  $F_\mu$  differ in at least one point. For the whole article, we choose to keep this non-strict definition.

**Mean and expectation.** If  $X$  and  $Y$  are real-valued, nonnegative random variables over the same probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ , we note  $\mathbf{E}[X]$  the expectation of  $X$  and  $\mathbf{E}[X | Y]$  the expectation of  $X$  conditional to  $Y$ . We allow them to be equal to  $+\infty$ .

If  $\mu$  is a probability measure on  $(\mathbf{R}_+, \mathfrak{B})$ , we note:

$$m_\mu := \int_0^\infty x \, d\mu(x) \in [0, \infty]$$

the *average* or *mean* of  $\mu$ . Notice that if  $X \sim \mu$ , then  $\mathbf{E}[X] = m_\mu$ .

**Lemma A.1.** Let  $\mu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$ ,  $m_\mu$  its mean,  $F_\mu$  its c.d.f. The following equality holds in  $\overline{\mathbf{R}_+}$ :

$$m_\mu = \int_0^\infty (1 - F_\mu(x)) \, dx.$$

*Proof.* We use Fubini's theorem so we can swap two integrals:

$$\begin{aligned} \int_0^\infty (1 - F_\mu(x)) \, dx &= \int_0^\infty \mu(x, \infty) \, dx \\ &= \int_{x=0}^\infty \int_{u=x^+}^\infty d\mu(u) \, dx \\ &= \int_{u=0}^\infty \int_{x=0}^{u^-} dx \, d\mu(u) \\ &= \int_0^\infty u \, d\mu(u) \end{aligned}$$

$$\int_0^\infty (1 - F(x_\mu)) \, dx = m_\mu. \quad \blacksquare$$

**Measures on product spaces.** Let  $(X, \mu)$  and  $(X', \Sigma', \mu')$  two measure spaces. We note  $\Sigma \otimes \Sigma'$  the product  $\sigma$ -algebra on  $X \times X'$ , and  $\mu \otimes \mu' \in \mathcal{M}(X \times X', \Sigma \otimes \Sigma')$  the tensor product of measures  $\mu$  and  $\mu'$ , *mutatis mutandis* for products indexed by collections.

Notice that  $\mu \otimes \mu'$  is the distribution of a couple of independent variables of distributions  $\mu$  and  $\mu'$ ;  $\mu^{\otimes \mathbf{N}}$  the distribution of a countable collection of i.i.d. variables, each one having distribution  $\mu$ ; etc.

**Rescalings.** For all  $\alpha > 0$ , let  $S_\alpha : x \in \mathbf{R}_+ \mapsto \alpha \cdot x$ . If  $\mu \in \mathbf{M}_1(\mathbf{R}_+, \mathfrak{B})$  and  $X$  is a random variable with distribution  $\mu$  on some probability space,  $(S_\alpha)_\#(\mu)$  is the distribution of  $\alpha \cdot X$ . We say that  $(S_\alpha)_\#(\mu)$  is the *rescaling of  $\mu$  of factor  $\alpha$* .

The relation:

$$\mu \equiv \nu \iff \exists \alpha > 0, \mu = (S_\alpha)_\#(\nu)$$

(“ $\mu$  and  $\nu$  are equal up to a rescaling”) is an equivalence relation on  $\mathbf{M}$ . Since every equivalence class  $[\mu]$  is one-to-one with  $\mathbf{R}_+$  (take  $\nu \mapsto m_\nu$ ), we write  $\mathbf{M}/\mathbf{R}_+^*$  for the quotient  $\mathbf{M}/\equiv$ .



## B Some elementary results and prerequisites

We assume all results about elementary theory of measure (e.g. chapters 1–3 and 5 in [Le 22]) and the foundations of probability theory (chapter 8 in [Le 22]). The following appendix recalls and gives a proof of the most important results in probability theory needed for the present article.

### B.1 Quantile functions

#### B.1.1 Definition and elementary properties

Let  $\mu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$  and  $F_\mu$  its c.d.f. For all  $p \in [0, 1)$ , the set

$$S_p = \{q \in \mathbf{R}_+ : F_\mu(q) \geq p\}$$

is nonempty, because  $F$  has a limit equal to 1 in  $+\infty$ . Hence, it admits a greatest lower bound  $\inf S_x$ . By right-continuity of  $F$ ,  $\inf S_x \in S_x$ . Hence, we can define the *quantile function* of  $\mu$  as follows:

**Definition B.1.** Let  $\mu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$ ,  $F_\mu$  its cumulative distribution function. The *quantile function* of  $\mu$  is the function  $Q_\mu : [0, 1) \rightarrow \mathbf{R}_+$  such that for all  $p \in [0, 1)$ ,

$$Q_\mu(p) := \min \{q \in \mathbf{R}_+ : F_\mu(q) \geq p\}.$$

We immediately have:

- $Q_\mu(0) = 0$ .
- $Q_\mu$  is nondecreasing.

Furthermore, we can show that:

- If  $F_\mu$  is a surjection, then  $Q_\mu$  is strictly increasing.
- For all  $p \in [0, 1)$ ,  $F_\mu(Q_\mu(p)) \geq p$ , with equality if and only if  $p \in F_\mu(\mathbf{R}_+)$ .

#### B.1.2 Galois inequalities and immediate applications

The definition of the quantile function immediately leads to the following equivalence, known as the ‘‘Galois inequalities’’:

$$\begin{aligned} &\text{for all } p \in [0, 1), q \in \mathbf{R}_+, \\ Q_\mu(p) \leq q &\iff p \leq F_\mu(q), \\ q < Q_\mu(p) &\iff F_\mu(q) < p. \end{aligned}$$

We directly deduce from the Galois inequalities that for all  $\mu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$ :

- For all  $x \in \mathbf{R}_+$ ,  $Q_\mu(F_\mu(x)) \leq x$ , and the inequality is strict if and only if there exists  $\varepsilon > 0$  such that  $F(x - \varepsilon) = F(x)$ .
- $Q_\mu \circ F_\mu \circ Q_\mu = Q_\mu$  and  $F_\mu \circ Q_\mu \circ F_\mu = F_\mu$ .
- $Q_\mu$  is constant on all interval of form  $[p, F_\mu(Q_\mu(p))]$  (which may be a singleton).
- If  $Q_\mu$  is strictly increasing, then  $F_\mu(\mathbf{R}_+)$  contains  $[0, 1)$ .

An important proposition is the following.

**Proposition B.2 (continuity of  $Q_\mu$ ).** Let  $\mu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$ .

1.  $Q_\mu$  is left-continuous at every point of  $(0, 1)$ .
2. Furthermore,  $Q_\mu$  is right-continuous at  $p \in [0, 1)$  if and only if  $F_\mu^{-1}(\{p\})$  contains at most one element.

*Proof.* 1. Let  $p \in (0, 1)$ . As  $Q_\mu$  is nondecreasing, it admits a left-limit  $Q(p^-)$  in  $p$ . We immediately have  $Q(p^-) \leq Q(p)$ .

Furthermore, for all  $\pi < p$ ,  $Q(p^-) \geq Q(\pi)$ . By Galois inequalities,  $F(Q(\pi^-)) \geq \pi$ . Taking the lowest upper bound for  $\pi < p$ , it follows that  $F(Q(p^-)) \geq p$ , so  $Q(p^-) \geq Q(p)$ .

Hence,  $Q(p^-) = Q(p)$ .

2.
  - Assume that  $F_\mu^{-1}(\{p\})$  is empty. Then  $F_\mu(Q_\mu(p)) > p$  and  $Q_\mu(F_\mu(Q_\mu(p))) = Q_\mu(p)$ . Hence,  $Q_\mu$  is constant between  $p$  and  $F_\mu(Q_\mu(p))$ , thus in particular right-continuous at  $p$ .
  - Suppose that  $F_\mu^{-1}(\{p\})$  is a singleton  $\{x\}$ . Then  $F_\mu(x) = p$ . By right-continuity of  $F_\mu$ , for every  $n \in \mathbf{N}$ , there exists  $\varepsilon_n$  such that  $p < F_\mu(x + \varepsilon_n) < p + 2^{-n}$ . Then,  $Q_\mu(p^+) = \lim_{n \rightarrow \infty} Q_\mu(F_\mu(x + \varepsilon_n))$ . But  $Q_\mu(F_\mu(x + \varepsilon_n)) \leq x + \varepsilon_n$ .

Since  $\varepsilon_n \xrightarrow[n \rightarrow \infty]{} 0$ , it implies that  $Q(p^+) \leq Q(p)$ . Hence,  $Q_\mu$  is right-continuous at  $p$ .

- Assume  $F_\mu^{-1}(\{p\})$  contains at least two elements  $x < y$ . Then  $Q_\mu(p) \leq x$  and for all  $\varepsilon > 0$ ,  $Q_\mu(p + \varepsilon) > y$ , so  $Q_\mu(p^+) \geq y$ . Hence,  $Q_\mu$  is not continuous at  $p$ . ■

**Inverse Galois.** The Galois inequalities are invalid if the sign is changed. However, for  $q \in \mathbf{R}_+$ ,  $p \in [0, 1)$ , one has:

$$\begin{aligned} Q_\mu(p) < q &\iff \exists \varepsilon > 0, Q_\mu(p) \leq q - \varepsilon \\ &\iff \exists \varepsilon > 0, p \leq F_\mu(q - \varepsilon) \\ Q_\mu(p) < q &\iff p \leq F_\mu(q^-). \end{aligned}$$

Notice that since  $F_\mu$  must be constant in the neighbourhood of  $q$ , the latter inequality may be an equality.

Hence, taking the negation, one has:

$$q \leq Q_\mu(p) \iff F_\mu(q^-) < p.$$

By the same reasoning,

$$F_\mu(q) \leq p \iff q < Q_\mu(p^+).$$

### B.1.3 Characterizing measures from quantile function

Let  $\mu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$ . Since  $Q_n$  is nondecreasing, it is a measurable function:

$$Q_\mu : ([0, 1), \mathfrak{B}) \longrightarrow (\mathbf{R}_+, \mathfrak{B}).$$

**Lemma B.3 ( $Q_\mu$  characterizes  $\mu$ ).** Let  $\mu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$  and  $Q_\mu$  its quantile function. Then,  $\mu = (Q_\mu)_\#(\mathcal{L}_{[0,1)})$ .

In other words,  $Q_\mu$  is a random variable on the space  $([0, 1), \mathfrak{B}, \mathcal{L})$ , and we have:  $Q_\mu \sim \mu$ .

This means, in particular, that a probability measure over  $(\mathbf{R}_+, \mathfrak{B})$  is fully characterized by its quantile function.

*Proof.* Let us work in the probability space  $([0, 1), \mathfrak{B}, \mathcal{L})$ . By definition, the random variable  $Q_\mu$  has distribution  $\nu := (Q_\mu)_\#(\mathcal{L}_{[0,1)})$ . It suffices to prove that it also has distribution  $\mu$ .

Thanks to Galois inequalities,  $F_\nu$  is such that for all  $x \in \mathbf{R}_+$ :

$$\begin{aligned} F_\nu(x) &:= \nu([0, x]) = (Q_\mu)_\# \mathcal{L}([0, x]) \\ &= \mathcal{L}(\{\omega \in [0, 1) : Q_\mu(\omega) \leq x\}) \\ &= \mathcal{L}(\{\omega \in [0, 1) : \omega \leq F_\mu(x)\}) \\ &= \mathcal{L}([0, F_\mu(x)]) \\ F_\nu(x) &= F_\mu(x). \end{aligned}$$

Since they have same c.d.f,  $\mu$  and  $\nu$  are equal. ■

Thanks lemma to this lemma and LOTUS, the integrals with respect to  $\mu$  can be expressed as integrals with respect  $\mathcal{L}$  implying the function  $Q_\mu$ :

**Proposition B.4 (LOTUS with  $Q_\mu$ ).** Let  $\mu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$ ,  $Q_\mu$  its quantile function. Let  $f : \mathbf{R}_+ \longrightarrow \mathbf{R}_+$  is a measurable function. Then the following equality holds in  $\overline{\mathbf{R}_+}$ :

$$\int_0^1 f(Q_\mu(p)) \, dp = \int_0^\infty f(x) \, d\mu(x).$$

### B.1.4 Several properties that can be characterized through the quantile function

The three following results consist in characterizing some properties of measures through their quantile functions.

**Proposition B.5.** Let  $\mu \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$ ,  $m_\mu$  its mean, and  $Q_\mu$  its quantile function. We have, in  $\overline{\mathbf{R}_+}$ :

$$m_\mu = \int_0^1 Q_\mu(u) \, du.$$

*Proof.* This is a direct application of proposition B.4 where  $f$  is the identity function. ■

**Proposition B.6 (rescaling).** Let  $\mu$  and  $\nu$  elements of  $\mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$ ,  $\alpha \in \mathbf{R}_+^*$ .  $\nu$  is a rescaling of  $\mu$  of factor  $\alpha$  if, and only if,  $Q_\nu = \alpha \cdot Q_\mu$ .

*Proof.* Assume  $Q_\nu = \alpha \cdot Q_\mu$ . Thanks to lemma B.3, we know that  $Q_\nu$  and  $Q_\mu$  can be seen as random variables over the probability space  $([0, 1), \mathfrak{B}, \mathcal{L})$  such that  $Q_\nu \sim \nu$  and  $Q_\mu \sim \mu$ . Thus,  $\nu$  is a rescaling of  $\mu$  with a factor  $\alpha$ .

Conversely, assume that there exists a random variable  $X$  on any probability space  $(\Omega, \mathcal{F}, \mathbf{P})$  such that  $X \sim \mu$  and  $\alpha X \sim \nu$ . Let  $F_\mu$  and  $F_\nu$  be the c.d.f. of those distributions. We have for any  $t \in [0, 1)$ :

$$\begin{aligned} Q_\nu(t) &= \min\{q : F_\nu(q) \geq p\} \\ &= \min\{q : \mathbf{P}(\alpha X \leq q) \geq p\} \\ &= \alpha \min\{q : \mathbf{P}(\alpha X \leq \alpha q) \geq p\} \\ &= \alpha \min\{q : \mathbf{P}(X \leq q) \geq p\} \\ &= \alpha \min\{q : F_\mu(q) \geq p\} \\ Q_\nu(t) &= \alpha Q_\mu(t). \quad \blacksquare \end{aligned}$$

**Proposition B.7 (FSD).** *Let  $\mu$  and  $\nu$  elements of  $\mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})$ ,  $F_\mu$  and  $F_\nu$  their c.d.f., and  $Q_\mu$  and  $Q_\nu$  their quantile functions. The three following statements are equivalent:*

- (i) *For all  $x \in \mathbf{R}_+$ ,  $F_\mu(x) \leq F_\nu(x)$ , i.e.  $\mu$  stochastically dominates  $\nu$  at first order.*
- (ii) *For all  $p \in [0, 1)$ ,  $Q_\mu(p) \geq Q_\nu(p)$ .*
- (iii) *There exists a probability space  $(\Omega, \mathcal{F}, \mathbf{P})$  and two random variables  $X$  and  $Y$  over it such that  $X \sim \mu$ ,  $Y \sim \nu$  and  $X \geq Y$   $\mathbf{P}$ -almost surely.*

*Proof.*

- Assume (i). Let  $p \in [0, 1)$ . Setting  $x := Q_\mu(p)$ , we have,

$$F_\mu(Q_\mu(p)) \leq F_\nu(Q_\mu(p)).$$

By Galois inequalities, we get

$$Q_\nu(F_\mu(Q_\mu(p))) \leq Q_\mu(p).$$

Furthermore,  $p \leq F_\mu(Q_\mu(p))$ .  $Q_\nu$  being nondecreasing, we get:

$$Q_\nu(p) \leq Q_\nu(F_\mu(Q_\mu(p))).$$

Hence, (ii) stands.

- Assume (ii). Consider the probability space  $([0, 1), \mathfrak{B}, \mathcal{L})$ . Let  $X = Q_\mu$  and  $Y = Q_\nu$ . According to lemma B.3, they match the conditions of (iii).
- Assume (iii). Let  $t \in \mathbf{R}_+$ . Then with probability 1,  $\mathbf{1}_{(X \leq t)} \geq \mathbf{1}_{(Y \leq t)}$ . Taking the expectation, we get  $\mathbf{P}(X \leq t) \geq \mathbf{P}(Y \leq t)$ , which proves (i). ■

## B.1.5 Extensions to $[0, 1]$ and $[0, \infty]$

We rigorously defined  $F$  and  $Q$  as mappings between the sets  $[0, 1)$  and  $[0, \infty)$ .

In fact, it can be sensible to “extend by left-continuity” the function  $Q$  to 1, by letting  $Q(1) := Q(1^-)$ . This notation is consistent with the definition, since we have:

$$Q(1) = \min\{q \in \overline{\mathbf{R}_+} : F(q) \geq 1\}$$

if we set  $F(\infty) := \mu([0, \infty)) = 1$ .

With these notations,  $F$  and  $Q$  are extended to mappings between the sets  $[0, 1]$  and  $[0, \infty]$ . However, until the end of the article, to avoid confusions, we chose to keep the rigorous definitions of  $F$  and  $Q$ , and we will restrict ourselves to the right-open intervals  $[0, 1)$  and  $\mathbf{R}_+$ .

## B.2 Weak convergence of measures

### B.2.1 Generalities

For this subsection, fix  $(\mathcal{X}, d)$  a separable, complete metric space. We call  $C_b(\mathcal{X})$  the set of continuous, bounded functions  $\mathcal{X} \rightarrow \mathbf{R}$ . If no ambiguity, we simply note  $\mathcal{M}_1$  for the set  $\mathcal{M}_1(\mathcal{X}, \mathfrak{B}(\mathcal{X}))$  of probability measures on  $(\mathcal{X}, \mathfrak{B}(\mathcal{X}))$ .

**Definition B.8.** 1. The *weak topology* on  $\mathcal{M}_1$  is the topology  $\mathscr{W}$  generated by the elementary balls of form

$$\left\{ \nu \in \mathcal{M}_1 : \left| \int_{\mathcal{X}} f \, d\nu - \int_{\mathcal{X}} f \, d\mu \right| < \varepsilon \right\}$$

for given  $f \in C_b(\mathcal{X})$ ,  $\mu \in \mathcal{M}_1$  and  $\varepsilon > 0$ . In other words, the open sets are arbitrary unions of finite intersections of elementary balls.

2. We call *weak convergence* the convergence with respect to the topology  $\mathscr{W}$ . We write  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathscr{W}} \mu_\infty$  if  $(\mu_n)_{n \in \mathbf{N}}$  weakly converges to  $\mu_\infty$ . It is immediate to check that this is equivalent to having

$$\int_{\mathcal{X}} f \, d\mu_n \xrightarrow[n \rightarrow \infty]{} \int_{\mathcal{X}} f \, d\mu_\infty$$

for all  $f \in C_b(\mathcal{X})$ .

3. Let  $X_1, \dots, X_n, \dots, X_\infty$  random variables taking values in  $\mathcal{X}$ , each being defined on a probability space  $(\Omega_n, \mathcal{F}_n, \mathbf{P}_n)$ . Let  $\mu_1, \dots, \mu_n, \dots, \mu_\infty$  their distributions. We say that  $(X_n)_{n \in \mathbf{N}}$  converges in distribution to  $X_\infty$  if  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathcal{W}} \mu_\infty$ . By LOTUS, this is equivalent to having, for all  $f \in C_b(\mathcal{X})$ ,

$$\int_{\Omega_n} f(X_n) d\mathbf{P}_n \xrightarrow[n \rightarrow \infty]{} \int_{\Omega_\infty} f(X_\infty) d\mathbf{P}_\infty.$$

Most authors only define the weak convergence for sequences. However, *a priori*, there is no reason for  $(\mathcal{M}_1, \mathcal{W})$  to be sequential, i.e. there is no guarantee that convergent sequences fully characterize the topology.

We immediately have the following fact:

**Proposition B.9.** *Let  $(\Omega, \mathcal{F}, \mathbf{P})$  a probability space and  $X_1, \dots, X_n, \dots, X_\infty$  random variables such that  $X_n \xrightarrow[n \rightarrow \infty]{} X_\infty$ ,  $\mathbf{P}$ -almost surely. Then  $X_n \xrightarrow[n \rightarrow \infty]{} X_\infty$  in distribution.*

*Proof.* Let  $f \in C_b(\mathcal{X})$ . By continuity of  $f$ ,  $f(X_n) \xrightarrow[n \rightarrow \infty]{} f(X_\infty)$   $\mathbf{P}$ -a.s. Furthermore,  $(f(X_n))_{n \in \mathbf{N}}$  is uniformly bounded by  $\max f$ . Hence, by dominated convergence theorem,

$$\int_{\Omega} f(X_n) d\mathbf{P} \xrightarrow[n \rightarrow \infty]{} \int_{\Omega} f(X_\infty) d\mathbf{P}. \quad \blacksquare$$

**Portmanteau's theorem.** The following strong version of *Portmanteau's theorem* gives other characterizations of the weak topology on  $\mathcal{M}_1$ .

**Theorem B.10 (Portmanteau).** *The following collections of subsets of  $\mathcal{M}_1(\mathfrak{B})$  each generate  $\mathcal{W}$  as a subbase:*

- (i)  $\{\nu \in \mathcal{M}_1 : |\int_{\mathcal{X}} f d\nu - \int_{\mathcal{X}} f d\mu| < \varepsilon\}$  for  $f \in C_b(\mathcal{X})$  uniformly continuous,  $\mu \in \mathcal{M}_1$  and  $\varepsilon > 0$ .
- (ii)  $\{\nu \in \mathcal{M}_1 : \nu(F) < \mu(F) + \varepsilon\}$  for  $\mu \in \mathcal{M}_1$ ,  $F$  closed subset of  $\mathcal{X}$  and  $\varepsilon > 0$ .
- (iii)  $\{\nu \in \mathcal{M}_1 : \nu(G) > \mu(G) - \varepsilon\}$  for  $\mu \in \mathcal{M}_1$ ,  $G$  open subset of  $\mathcal{X}$  and  $\varepsilon > 0$ .

- (iv)  $\{\nu \in \mathcal{M}_1 : |\nu(A) - \mu(A)| < \varepsilon\}$  for  $\mu \in \mathcal{M}_1$ ,  $A \in \mathfrak{B}(\mathcal{X})$  such that  $\mu(\partial B) = 0$  (where  $\partial B$  is the topological boundary of  $B$ ) and  $\varepsilon > 0$ .

Each characterization of the topology immediately gives a characterization of the weakly convergent sequences of  $\mathcal{M}_1$ . For instance, from point (iv), follows that  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathcal{W}} \mu_\infty$  if and only if for all  $A \in \mathfrak{B}(\mathcal{X})$  such that  $\mu_\infty(\partial A) = 0$ ,  $\mu_n(A) \xrightarrow[n \rightarrow \infty]{} \mu_\infty(A)$ .

For a proof, see [Bil68, appendix III, theorem 3]. The book does not explicitly state point (i); however, it is a free consequence of Billingsley's proof.

From (i), follows this implication:

**Proposition B.11.** *Let  $(\Omega, \mathcal{F}, \mathbf{P})$  a probability space and  $X_1, \dots, X_n, \dots, X_\infty$  random variables with values in  $\mathcal{X}$ . Let  $\mu_1, \dots, \mu_\infty$  their distributions.*

*If  $\mathbf{E}[d(X_n, X_\infty)] \xrightarrow[n \rightarrow \infty]{} 0$ , then  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathcal{W}} \mu_\infty$ .*

*Proof.* Let  $f : \mathcal{X} \rightarrow \mathbf{R}$  bounded, uniformly continuous. Let  $M := \sup f$ . Chose  $\varepsilon > 0$ .

By uniform continuity of  $f$ , there exists  $\delta > 0$  such that for all  $x, y \in \mathcal{X}$ ,

$$d(x, y) < \delta \implies |f(x) - f(y)| \leq \varepsilon.$$

For each  $n \in \mathbf{N}$ , let  $A_n := \{d(X_n, X_\infty) < \delta\}$ . We have:

$$\begin{aligned} & |\mathbf{E}[f(X_n)] - \mathbf{E}[f(X_\infty)]| \\ & \leq \mathbf{E}[|f(X_n) - f(X_\infty)|] \\ & = \int_{A_n} |f(X_n) - f(X_\infty)| d\mathbf{P} \\ & \quad + \int_{\Omega \setminus A_n} |f(X_n) - f(X_\infty)| \\ & \leq \varepsilon \mathbf{P}(A_n) + 2M \mathbf{P}(\Omega \setminus A_n). \end{aligned}$$

Yet  $\mathbf{P}(A_n) \leq 1$ , and by Markov's inequality,

$$\mathbf{P}(\Omega \setminus A_n) \leq \frac{\mathbf{E}[d(X_n, X_\infty)]}{\delta}.$$

Hence, if  $N$  is chosen such that for every  $n \geq N$ ,  $\mathbf{E}[d(X_n, X_\infty)] < \frac{\delta \varepsilon}{2M}$ , we have:

$$|\mathbf{E}[f(X_n)] - \mathbf{E}[f(X_\infty)]| \leq 2\varepsilon.$$

Thus,  $\mathbf{E}[f(X_n)] \xrightarrow{n \rightarrow \infty} \mathbf{E}[f(X_\infty)]$ . By Portmanteau's theorem, this is enough to state the convergence in distribution.  $\blacksquare$

**Metric inducing  $\mathscr{W}$ .** A nice property of  $\mathscr{W}$  is that it is metrizable. For all  $A \in \mathfrak{B}(\mathcal{X})$  and  $\varepsilon > 0$ , let:

$$A^\varepsilon := \{x \in \mathcal{X} : \exists a \in A, d(x, a) < \varepsilon\}$$

which is immediately an open set.

**Definition B.12 (Prokhorov metric).** For all  $\mu, \nu \in \mathcal{M}_1$ , we call  $d_P(\mu, \nu)$  the infimum of positive  $\varepsilon$  such that the inequalities  $\mu(A) < \nu(A^\varepsilon) + \varepsilon$  and  $\nu(A) < \mu(A^\varepsilon) + \varepsilon$  stand for all  $A \in \mathfrak{B}(\mathcal{X})$ .

We note  $d_P : \mathcal{M}_1(\mathcal{X})^2 \rightarrow \mathbf{R}_+$  the Prokhorov metric.

It is known that  $d_P$  is a metric, and that  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathscr{W}} \mu_\infty$  if and only if  $d_P(\mu_n, \mu_\infty) \rightarrow 0$  (see for instance [Bil99, p. 72, remark (i)]). However, this is not enough to prove that  $d_P$  induces the topology  $\mathscr{W}$ . This more difficult result is proven in [Bil68, appendix III, theorem 5]\*.

Now, we know that  $\mathscr{W}$  is metrizable, hence sequential.

## B.2.2 The real line

Now we restrict on real numbers. For this article, we only need to consider  $\mathcal{X} = \mathbf{R}_+$ , but all the following results can be extended to  $\mathcal{X} = \mathbf{R}$ .

We give a few characterization of weak convergence that will freely be used in the article.

**Theorem B.13.** Let  $\mu_1, \dots, \mu_n, \dots, \mu_\infty$  distributions on  $(\mathbf{R}_+, \mathfrak{B})$ ,  $F_{\mu_\bullet}$  their cdf and  $Q_{\mu_\bullet}$  their quantile functions. The following are equivalent:

- (i)  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathscr{W}} \mu_\infty$ ;
- (ii)  $F_{\mu_n}(x) \xrightarrow[n \rightarrow \infty]{} F_{\mu_\infty}(x)$  for every  $x \in \mathbf{R}_+$  where  $F_{\mu_\infty}$  is continuous;
- (iii)  $F_{\mu_n}(x) \xrightarrow[n \rightarrow \infty]{} F_{\mu_\infty}(x)$  for  $\mathcal{L}$ -almost all  $x \in \mathbf{R}_+$ ;

(iv)  $Q_{\mu_n}(p) \xrightarrow[n \rightarrow \infty]{} Q_{\mu_\infty}(p)$  for every  $p \in [0, 1)$  where  $Q_{\mu_\infty}$  is continuous;

(v)  $Q_{\mu_n}(p) \xrightarrow[n \rightarrow \infty]{} Q_{\mu_\infty}(x)$  for  $\mathcal{L}$ -almost all  $p \in [0, 1)$ ;

(vi) There exist a probability space  $(\Omega, \mathcal{F}, \mathbf{P})$  and nonnegative random variables  $X_1, \dots, X_n, \dots, X_\infty$  such that  $X_n \xrightarrow[n \rightarrow \infty]{} X_\infty$   $\mathbf{P}$ -almost surely,  $X_n \sim \mu_n$  for every  $n \in \mathbf{N}^*$  and  $X_\infty \sim \mu_\infty$ .

*Proof.* • (i)  $\implies$  (ii) is a direct consequence of Portmanteau's theorem (point (iv) of theorem B.10), taking  $\mu = \mu_\infty$  and  $A = [0, x]$ , thus  $\mu_\infty(\partial A) = \mu_\infty(\{x\}) = 0$  by continuity of  $F_{\mu_\infty}$ . (Notice that 0 is not in  $\partial A$ , since  $[0, x]$  is open in  $\mathbf{R}_+$ .)

• Assume (ii). Since  $F_{\mu_\infty}$  is nondecreasing, the set of its discontinuity points is at most countable. Hence, it has Lebesgue measure 0, i.e. (iii) stands.

• Assume (iii). Consider the probability space  $(\mathbf{R}_+, \mathfrak{B}, \mathbf{P})$  where  $\mathbf{P}(dt) = e^{-t} \mathcal{L}(dt)$  ( $\mathbf{P}$  is the measure with density  $t \mapsto e^{-t}$  with respect to  $\mathcal{L}$ ). The  $F_{\mu_\bullet}$ 's can be seen as random variables on  $(\mathbf{R}_+, \mathfrak{B}, \mathbf{P})$ .

Let  $C \subseteq \mathbf{R}_+$  the set of  $\omega \in \mathbf{R}_+$  such that  $F_{\mu_n}(\omega) \xrightarrow[n \rightarrow \infty]{} F_{\mu_\infty}(\omega)$ . We know that  $\mathcal{L}(C^c) = 0$ , hence  $\mathbf{P}(C^c) = 0$ , thus  $\mathbf{P}(C) = 1$ .

Consider  $x \in \mathbf{R}_+$  such that  $Q_{\mu_\infty}$  is continuous at  $x$ . Let  $D$  the set of  $\omega \in \mathbf{R}_+$  such that  $F_{\mu_\infty}(\omega) = x$ . By proposition B.2,  $D$  is either a singleton or empty. Hence,  $\mathcal{L}(L_x) = 0$ , thus  $\mathbf{P}(L_x) = 0$ .

Now, for all  $\omega \in C \setminus D$ , we have  $F_{\mu_n}(\omega) \xrightarrow[n \rightarrow \infty]{} F_{\mu_\infty}(\omega) \neq x$ . Since the function  $\mathbf{1}_{[0, x]}$  is continuous everywhere but in  $x$ ,

$$\mathbf{1}_{[0, x]}(F_{\mu_n}(\omega)) \xrightarrow[n \rightarrow \infty]{} \mathbf{1}_{[0, x]}(F_{\mu_\infty}(\omega)).$$

\*The proof has been removed in the Second Edition of the book and can only be found in the 1968 edition. [Bil68] generalizes the result by releasing the separability hypothesis.

Since  $\mathbf{P}(C \setminus D) = 1$  and  $\mathbf{1}_{[0,x]}(F_{\mu_n})$  is bounded by 1, by dominated convergence theorem, we have:

$$\begin{aligned} \mathbf{E}[\mathbf{1}_{[0,x]}(F_{\mu_n})] &\xrightarrow{n \rightarrow \infty} \mathbf{E}[\mathbf{1}_{[0,x]}(F_{\mu_\infty})] \\ \mathbf{P}(F_{\mu_n} < x) &\xrightarrow{n \rightarrow \infty} \mathbf{P}(F_{\mu_\infty} < x). \end{aligned}$$

However, for  $n \in \mathbf{N} \cup \{\infty\}$ ,  $\omega \in \mathbf{R}_+$ ,  $F_{\mu_n}(\omega) < x \iff \omega < Q_{\mu_n}(x)$ . Hence,

$$\mathbf{P}([0, Q_{\mu_n}(x)]) \xrightarrow{n \rightarrow \infty} \mathbf{P}([0, Q_{\mu_\infty}(x)])$$

The function  $t \mapsto \mathbf{P}([0, t]) = 1 - e^{-t}$  has inverse  $p \mapsto -\ln(1 - p)$ , which is continuous over  $\mathbf{R}_+$ . Hence,

$$Q_{\mu_n}(x) \xrightarrow{n \rightarrow \infty} Q_{\mu_\infty}(x),$$

i.e. (iv) is true.

- (iv)  $\implies$  (v) is proven exactly the same way as (ii)  $\implies$  (iii).
- (v)  $\implies$  (vi) is immediate since for every  $n \in \mathbf{N}^* \cup \{\infty\}$ ,  $Q_{\mu_n}$  is a random variable on  $([0, 1], \mathfrak{B}, \mathcal{L})$  with distribution  $\mu_n$  (see lemma B.3).
- (vi)  $\implies$  (i) is a direct consequence of proposition B.9.  $\blacksquare$

## B.3 $W_1$ metric and $W_1$ convergence

### B.3.1 The $W_1$ metric

For this subsection, we work on the set  $\mathbf{M}'$  of probability measures on  $(\mathbf{R}_+, \mathfrak{B})$  with finite expectations, i.e.  $\mathbf{M}' = \mathbf{M} \cup \{\delta_0\}$ . In the article, the following results will be restricted to  $\mathbf{M}$ , since we exclude the distribution  $\delta_0$ .

#### Definition B.14 (Wasserstein-1 metric).

Let  $\mu, \nu \in \mathbf{M}'$ . We let  $W_1(\mu, \nu) := \|Q_\mu - Q_\nu\|_1$ , i.e.

$$W_1(\mu, \nu) = \int_0^1 |Q_\mu(t) - Q_\nu(t)| dt.$$

It is immediate that  $W_1$  is a pseudometric. Now assume  $W_1(\mu, \nu) = 0$ . This implies that  $Q_\mu - Q_\nu = 0$  almost everywhere. However,  $Q_\mu - Q_\nu$  is left-continuous. Hence  $Q_\mu - Q_\nu = 0$

everywhere. By proposition 2.1,  $\mu = \nu$ . Hence,  $W_1$  is a metric.

The name of *Wassertein-p metric* more usually refers to the optimal cost of transportation for  $L^p$  cost in the Kantorovitch optimal transportation problem framework: for  $(\mathcal{X}, d)$  a Polish space and  $\mu, \nu$  measures on  $(\mathcal{X}, \mathfrak{B})$ ,  $W_p(\mu, \nu)$  is equal to:

$$\min_{\pi \in \Pi(\mu, \nu)} \left( \int_{(x,y) \in \mathcal{X}^2} d(x,y)^p d\pi(x,y) \right)^{\frac{1}{p}}$$

where  $\Pi(\mu, \nu)$  is the set of distributions on  $\mathcal{X}^2$  with marginals  $\mu$  and  $\nu$ . Happily those definitions are consistent. For a proof that for the minimum has value  $\left( \int_0^1 |Q_\mu(p) - Q_\nu(p)|^p dp \right)^{\frac{1}{p}}$ , see for instance [Vil03, theorem 2.18 and remarks 2.19].

Another way to express the  $W_1$  metric is the following:

**Proposition B.15.** *Let  $\mu, \nu \in \mathbf{M}'$ ,  $F_\mu$  and  $F_\nu$  their cdf. Then  $W_1(\mu, \nu) = \|F_\mu - F_\nu\|_1$ .*

*Proof.* We have:

$$\begin{aligned} &\int_0^\infty |F_\mu(x) - F_\nu(x)| dx \\ &= \int_{x=0}^\infty \left( \int_{y=F_\mu(x)}^{F_\nu(x)} 1 dy + \int_{y=F_\nu(x)}^{F_\mu(x)} 1 dy \right) dx \\ &= \int_{x=0}^\infty \left( \int_{y=F_\mu(x)^+}^{F_\nu(x)} 1 dy + \int_{y=F_\nu(x)^+}^{F_\mu(x)} 1 dy \right) dx \\ &= \int_{x=0}^\infty \int_{y=0}^{1^-} \left( \begin{array}{c} \mathbf{1}_{F_\mu(x) < y \leq F_\nu(x)} \\ + \mathbf{1}_{F_\nu(x) < y \leq F_\mu(x)} \end{array} \right) dy dx \\ &= \int_{x=0}^\infty \int_{y=0}^{1^-} \left( \begin{array}{c} \mathbf{1}_{Q_\nu(y) \leq x < Q_\mu(x)} \\ + \mathbf{1}_{Q_\mu(y) \leq x < Q_\nu(x)} \end{array} \right) dy dx \\ &= \int_{y=0}^{1^-} \int_{x=0}^\infty \left( \begin{array}{c} \mathbf{1}_{Q_\nu(y) \leq x < Q_\mu(x)} \\ + \mathbf{1}_{Q_\mu(y) \leq x < Q_\nu(x)} \end{array} \right) dx dy \\ &= \int_0^{1^-} |Q_\mu(y) - Q_\nu(y)| dy. \end{aligned}$$

using, in order, the nonatomicity of Lebesgue measure, the Galois inequalities, Fubini's theorem, and making the same computations backwards.  $\blacksquare$



### B.3.2 The $W_1$ convergence and its characterization

By definition, the  $W_1$  convergence of a sequence  $(\mu_n)_{n \in \mathbf{N}}$  to a limit  $\mu_\infty$  is equivalent to the  $L^1$  convergence of the random variables  $(Q_{\mu_n})_{n \in \mathbf{N}}$  to  $Q_{\mu_\infty}$  in the probability space  $([0, 1], \mathfrak{B}, \mathcal{L})$ . We give some other characterizations.

**Lemma B.16 (Scheffé).** *Let  $(\mathcal{X}, \mathcal{F}, \mu)$  a probability space and  $f_1, \dots, f_n, \dots, f_\infty$  measurable functions  $\mathcal{X} \rightarrow \mathbf{R}_+$ .*

*If  $f_n \xrightarrow[n \rightarrow \infty]{} f_\infty$   $\mu$ -almost everywhere and  $\int_{\mathcal{X}} f_n d\mu \xrightarrow[n \rightarrow \infty]{} \int_{\mathcal{X}} f_\infty d\mu < \infty$ , then:*

$$\int_{\mathcal{X}} |f_n - f_\infty| d\mu \xrightarrow[n \rightarrow \infty]{} 0.$$

*Proof.* For  $g : \mathcal{X} \rightarrow \mathbf{R}$ , call  $g_+ = \max(g, 0)$  and  $g_- = \max(-g, 0)$ . We have  $g = g_+ - g_-$  and  $|g| = g_+ + g_-$ . Hence,  $|g| = 2g_+ - g$ . It follows that:

$$\|g\|_1 = \int_{\mathcal{X}} |g| d\mu = 2 \int_{\mathcal{X}} g_+ d\mu - \int_{\mathcal{X}} g d\mu. \quad (\text{B.17})$$

Since  $f_n \xrightarrow[n \rightarrow \infty]{} f_\infty$   $\mu$ -a.e. and  $x \mapsto x_+$  is continuous, it follows that  $(f_\infty - f_n)_+ \xrightarrow[n \rightarrow \infty]{} 0$ ,  $\mu$ -a.e. Furthermore,

$$0 \leq (f_\infty - f_n)_+ \leq f_\infty.$$

Since  $f_\infty$  is integrable, by dominated convergence theorem,

$$\int_{\mathcal{X}} (f_\infty - f_n)_+ d\mu \xrightarrow[n \rightarrow \infty]{} 0.$$

Furthermore, by hypothesis:

$$\int_{\mathcal{X}} (f_\infty - f_n) d\mu \xrightarrow[n \rightarrow \infty]{} 0.$$

Injecting this in (B.17) with  $g = f_\infty - f_n$ , we get  $\|f_\infty - f_n\|_1 \xrightarrow[n \rightarrow \infty]{} 0$ .  $\blacksquare$

Scheffé's lemma gives us an important characterization of  $W_1$ -convergence:

**Proposition B.18.** *Let  $(\mu_n)_{n \in \mathbf{N}} \in \mathbf{M}^{\mathbf{N}}$ ,  $\mu_\infty \in \mathbf{M}$  and  $m_{\mu_1}, \dots, m_{\mu_\infty}$  their means. The*

*following assertions are equivalent:*

(i)  $\mu_n \xrightarrow[n \rightarrow \infty]{W_1} \mu_\infty$ .

(ii)  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathcal{W}} \mu_\infty$  and  $m_{\mu_n} \xrightarrow[n \rightarrow \infty]{} m_{\mu_\infty}$ .

(iii) *There exists a probability space  $(\Omega, \mathcal{F}, \mathbf{P})$  and nonnegative random variables  $X_1, \dots, X_n, \dots, X_\infty$  such that  $X_n \sim \mu_n$  for every  $n \in \mathbf{N} \cup \{\infty\}$  and  $\mathbf{E}[|X_n - X_\infty|] \xrightarrow[n \rightarrow \infty]{} 0$ .*

*Proof.* • (i)  $\implies$  (iii) : take  $\Omega = [0, 1]$ ,  $\mathcal{F} = \mathfrak{B}$  and  $\mathbf{P} = \mathcal{L}$  and for all  $x \in \mathbf{N} \cup \{\infty\}$ , let  $X_n = Q_{\mu_n}$ .

• Assume (iii). By triangle inequality,  $m_{\mu_n} = \mathbf{E}[X_n] \xrightarrow[n \rightarrow \infty]{} \mathbf{E}[X_\infty] = m_{\mu_\infty}$ . Furthermore, since  $L^1$  convergence implies weak convergence (proposition B.11),  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathcal{W}} \mu_\infty$ . Hence, (ii) stands.

• (ii)  $\implies$  (i) lies in applying Scheffé's lemma to the  $Q_{\mu_n}$ 's.  $\blacksquare$

## B.4 Uniform integrability

### B.4.1 Definition

Recall the definition of uniform integrability of random variables. We restrict ourselves to nonnegative-valued random variables. In order to simplify the following redaction, we also consider “uniformly integrable measures”, i.e. measures which are the distributions of a uniformly integrable collection of random variables — this concept is nonstandard.

**Definition B.19 (u.i. for random vars).** Let  $(\Omega_i, \mathcal{F}_i, \mathbf{P}_i)_{i \in I}$  a collection of probability spaces and for all  $i \in I$ ,  $U_i : \Omega_i \rightarrow \mathbf{R}$  a random variable.

The collection  $(U_i)_{i \in I}$  is said *uniformly integrable (u.i.)* if

$$\sup_{i \in I} \int_{\Omega_i} |U_i| \mathbf{1}_{|U_i| > \alpha} d\mathbf{P}_i \xrightarrow[\alpha \rightarrow +\infty]{} 0.$$

**Definition B.20 (u.i. for measures).** Let  $I$  a set and  $(\mu_i)_{i \in I} \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})^I$ . We say that  $(\mu_i)_{i \in I}$  is *uniformly integrable* if

$$\sup_{i \in I} \int_{\mathbf{R}_+} x \mathbf{1}_{x > \alpha} d\mu_i(x) \xrightarrow[\alpha \rightarrow +\infty]{} 0.$$

These definitions are coherent, thanks to the following proposition:

**Proposition B.21.** *Let  $I$  a set and  $(\mu_i)_{i \in I} \in \mathcal{M}_1(\mathbf{R}_+, \mathfrak{B})^I$ . The following are equivalent:*

- (i)  $(\mu_i)_{i \in I}$  is uniformly integrable.
- (ii) There exists probability spaces  $(\Omega_i, \mathcal{F}_i, \mathbf{P}_i)_{i \in I}$  and random variables  $U_i : \Omega_i \rightarrow \mathbf{R}_+$  such that for all  $i \in I$ ,  $U_i \sim \mu_i$  and  $(U_i)_{i \in I}$  is uniformly integrable.
- (iii) For all probability spaces  $(\Omega_i, \mathcal{F}_i, \mathbf{P}_i)_{i \in I}$  and random variables  $U_i : \Omega_i \rightarrow \mathbf{R}_+$  such that  $U_i \sim \mu_i$ ,  $(U_i)_{i \in I}$  is uniformly integrable.

*Proof.* Immediate consequence of LOTUS and the fact that  $Q_\mu \sim \mu$ .  $\blacksquare$

In particular, a collection  $(\mu_i)_{i \in I}$  of probability measures is u.i. if, and only if, the collection of random variables  $(Q_{\mu_i})_{i \in I}$  defined on the probability space  $([0, 1], \mathfrak{B}, \mathcal{L})$  is u.i.

We will freely use the following technical lemma.

**Lemma B.22.** *Let  $(U_i)_{i \in I}$  real, integrable random variables on probability spaces  $(\Omega_i, \mathcal{F}_i, \mathbf{P}_i)$ .  $(U_i)_{i \in I}$  is uniformly integrable if and only if for all  $\varepsilon > 0$ , there exists  $\alpha_\varepsilon > 0$  and  $J_\varepsilon > 0$  such that  $I \setminus J_\varepsilon$  is finite and*

$$\sup_{i \in J_\varepsilon} \mathbf{E}_i[|U_i| \mathbf{1}_{|U_i| > \alpha_\varepsilon}] < \varepsilon. \quad (\text{B.23})$$

*Proof.* Assume  $(U_i)_{i \in I}$  is u.i. Let  $\varepsilon > 0$ . By definition, there exists  $\alpha > 0$  such that:

$$\sup_{i \in I} \mathbf{E}_i[|U_i| \mathbf{1}_{|U_i| > \alpha}] < \alpha$$

and (B.23) holds for  $\beta_\varepsilon = \alpha$  and  $J_\varepsilon := I$ .

Conversely, assume that for all  $\varepsilon > 0$ , there exists  $J_\varepsilon$  and  $\beta_\varepsilon$  such that  $I \setminus J_\varepsilon$  is finite and (B.23) holds.

Let  $j \in I \setminus J_\varepsilon$ . Since  $|U_j|$  is integrable, by dominated convergence theorem,

$$\mathbf{E}[|U_j| \mathbf{1}_{|U_j| > \alpha}] \xrightarrow{\alpha \rightarrow +\infty} 0$$

hence there exists  $\alpha_j$  such that

$$\mathbf{E}[|U_j| \mathbf{1}_{|U_j| > \alpha_j}] < \varepsilon.$$

Now if we take  $A = \max(\beta_\varepsilon, \max_{j \in I \setminus J_\varepsilon} \alpha_j)$ , for all  $\alpha > A$  and  $i \in I$ :

- either  $i \in J_\varepsilon$  and

$$\begin{aligned} & \mathbf{E}[|U_i| \mathbf{1}_{|U_i| > \alpha}] \\ & < \sup_{i \in J_\varepsilon} \mathbf{E}_i[|U_i| \mathbf{1}_{|U_i| > \alpha_\varepsilon}] < \varepsilon; \end{aligned}$$

- or  $i \in I \setminus J_\varepsilon$ , and

$$\mathbf{E}[|U_i| \mathbf{1}_{|U_i| > \alpha}] < \mathbf{E}[|U_j| \mathbf{1}_{|U_j| > \alpha_j}] < \varepsilon.$$

Hence,  $(U_i)_{i \in I}$  is u.i.  $\blacksquare$

One can directly check that if a collection  $(U_i)_{i \in I}$  of real random variables is u.i., then for all  $i \in I$ ,  $\mathbf{E}_i[|U_i|] < \infty$ . Hence, only need to deal with integrable random variables and measures in  $\mathbf{M}$ . In fact, uniform integrability even implies that  $\mathbf{E}_i[|U_i|]$  are *uniformly* bounded. Indeed, take  $\alpha > 0$  such that  $\sup_{i \in I} \mathbf{E}[|U_i| \mathbf{1}_{|U_i| > \alpha}] < 1$ ; we have for all  $i \in I$ ,  $\mathbf{E}[|U_i|] < \alpha \mathbf{P}_i(\{|U_i| \leq \alpha\}) + 1 \leq \alpha + 1$ . The converse is false.

#### B.4.2 Some sufficient conditions for having uniform integrability

**Uniform  $L^p$  bound.** On the other hand, for  $p > 1$ , if there exists  $M > 0$  such that  $\mathbf{E}_i[|X_i|^p] < M$  for all  $i \in I$ , then  $(X_i)_{i \in I}$  is u.i., the converse being false. Indeed,

$$\mathbf{1}_{(|X_i| > \alpha)} |X_i| = \mathbf{1}_{(|X_i|^{p-1} > \alpha^{p-1})} |X_i| \leq \frac{|X_i|^p}{\alpha^{p-1}}$$

hence

$$\mathbf{E}_i[\mathbf{1}_{(|X_i| > \alpha)} |X_i|] \leq \frac{M}{\alpha^{p-1}} \xrightarrow{n \rightarrow \infty} 0.$$

**$L^1$  singleton.** If  $X$  is an integrable random variable on any probability space, then the collection  $\{X\}$  is uniformly integrable. Indeed,  $|X| \mathbf{1}_{|X| > n} \xrightarrow{n \rightarrow +\infty} 0$  pointwise and those functions are dominated by  $|X|$ , which is integrable. Hence, the dominated convergence theorem states that

$$\mathbf{E}[\mathbf{1}_{X > n} X] \xrightarrow{\alpha \rightarrow \infty} \mathbf{E}[0] = 0.$$

### B.4.3 Operations on uniformly integrable collections

**Multiplication by bounded scalars.** Assume  $(X_i)_{i \in I}$  is u.i. and  $(a_j)_{j \in I}$  is bounded. Immediately,  $(a_j X_i)_{i \in I, j \in J}$  is u.i.

**Sum.** Let  $X_i, i \in I$  and  $Y_j, j \in J$  real random variables defined on the same probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ . Assume  $(X_i)_{i \in I}$  is u.i. and  $(Y_j)_{j \in J}$  is u.i. Then  $(X_i + Y_j)_{i \in I, j \in J}$  is u.i.

Indeed, for all  $(a, b) \in \mathbf{R}$  and  $c \in \mathbf{R}_+$ , the following statement holds:

$$|a + b| \mathbf{1}_{(|a+b|>2c)} \leq 2|a| \mathbf{1}_{(|a|>c)} + 2|b| \mathbf{1}_{(|b|>c)}$$

(this is trivial if  $|a + b| \leq 2c$ ; otherwise one can directly check it in cases  $|a| \geq |b|$  and  $|a| \leq |b|$ ).

Thus:

$$\begin{aligned} & \mathbf{E}[|X_i + Y_j| \mathbf{1}_{|X_i + Y_j| > 2\alpha}] \\ & \leq 2\mathbf{E}[|X_i| \mathbf{1}_{|X_i| > \alpha}] + 2\mathbf{E}[|Y_j| \mathbf{1}_{|Y_j| > \alpha}] \\ & \leq 4\epsilon \end{aligned}$$

for  $\alpha$  large enough.

**Proposition B.24 (u.i. and FSD).** Let  $(\mu_i)_{i \in I}$  and  $(\nu_j)_{j \in J}$  collections of measures in  $\mathbf{M}$ . Assume that:

- (i) The collection  $(\mu_i)_{i \in I}$  is uniformly integrable;
- (ii) For all  $j \in J$ , there exists a  $\iota(j) \in I$  such that  $\nu_j$  is stochastically dominated at first order by  $\mu_{\iota(j)}$ .

Then  $(\nu_j)_{j \in J}$  is uniformly integrable.

*Proof.* Consider the probability space  $([0, 1], \mathfrak{B}, \mathcal{L})$ . For every  $i \in I, j \in J$ , let the random variables  $A_i = Q_{\mu_i}$  and  $B_j = Q_{\nu_j}$ .

The proposition B.7 and the hypothesis of stochastic dominance ensure that for all  $j \in J$ ,  $B_j \leq A_{\iota(j)}$ .

Fix  $\alpha \in \mathbf{R}_+$  and  $j \in J$ .  $\mathcal{L}$ -almost surely,

$$\begin{aligned} \mathbf{1}_{(B_j > \alpha)} & \leq \mathbf{1}_{(A_{\iota(j)} > \alpha)} \\ B_j \mathbf{1}_{(B_j > \alpha)} & \leq A_{\iota(j)} \mathbf{1}_{(A_{\iota(j)} > \alpha)}. \end{aligned}$$

Hence,

$$\mathbf{E}[B_j \mathbf{1}_{(B_j > \alpha)}] \leq \mathbf{E}[A_{\iota(j)} \mathbf{1}_{(A_{\iota(j)} > \alpha)}].$$

Taking the lowest upper bound over  $j \in J$ , we get:

$$\sup_{j \in J} \mathbf{E}[B_j \mathbf{1}_{(B_j > \alpha)}] \leq \sup_{i \in I} \mathbf{E}[A_i \mathbf{1}_{(A_i > \alpha)}]$$

The hypothesis of uniform integrability of collection  $(\mu_i)_{i \in I}$  ensures that the right-hand side converges to 0 as  $\alpha \rightarrow \infty$ . ■

### B.4.4 Uniform integrability and $W_1$ convergence

The main result that justifies to use uniform integrability in this article is the following, which is a measure-based adaptation of [Bil95, theorem 16.14].

**Proposition B.25.** Let  $(\mu_n)_{n \in \mathbf{N}} \in \mathbf{M}^{\mathbf{N}}$  and  $\mu_\infty$  in  $\mathbf{M}$ . Assume that  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathcal{W}} \mu_\infty$ . The following assertions are equivalent:

$$(i) \quad m_{\mu_n} \xrightarrow[n \rightarrow \infty]{} m_{\mu_\infty}.$$

$$(ii) \quad (\mu_n)_{n \in \mathbf{N}} \text{ is uniformly integrable.}$$

*Proof.* Consider the probability space  $([0, 1], \mathfrak{B}, \mathcal{L})$ . For every  $n \in \mathbf{N}$ , let  $Q_n := Q_{\mu_n}$  and  $Q := Q_{\mu_\infty}$ .

For  $\alpha > 0$ , let  $f_\alpha : x \mapsto x \mathbf{1}_{x \leq \alpha}$  and  $g_\alpha : g \mapsto x \mathbf{1}_{x > \alpha}$ .

For every  $n \in \mathbf{N}$ , by lemma B.5,

$$\begin{aligned} m_{\mu_n} & = \mathbf{E}[f_\alpha(Q_n)] + \mathbf{E}[g_\alpha(Q_n)]; \\ m_{\mu_\infty} & = \mathbf{E}[f_\alpha(Q)] + \mathbf{E}[g_\alpha(Q)]. \end{aligned} \quad (\text{B.26})$$

Furthermore, by proposition B.3,  $Q_n \xrightarrow[n \rightarrow \infty]{} Q$   $\mathcal{L}$ -a.s. Hence, if  $\mu(\{\alpha\}) = 0$ , then:

$$f_\alpha(Q_n) \xrightarrow[n \rightarrow \infty]{} f_\alpha(Q); \quad g_\alpha(Q_n) \xrightarrow[n \rightarrow \infty]{} g_\alpha(Q).$$

Since  $\mathbf{E}[f_\alpha(Q_n)] \leq \alpha$  for every  $n \in \mathbf{N}$ , by dominated convergence theorem,

$$\mathbf{E}[f_\alpha(Q_n)] \xrightarrow[n \rightarrow \infty]{} \mathbf{E}[f_\alpha(Q)]. \quad (\text{B.27})$$

Now we can prove both implications.

(i)  $\implies$  (ii). For every  $n \in \mathbf{N}$ , by (B.26):

$$\mathbf{E}[g_\alpha(Q_n)] \leq \mathbf{E}[g_\alpha(Q)] + |m_{\mu_\infty} - m_{\mu_n}| + |\mathbf{E}[f_\alpha(Q_n)] - \mathbf{E}[f_\alpha(Q)]|.$$

Fix  $\varepsilon > 0$ . Since  $\{Q\}$  is u.i., there exists  $\alpha_1 > 0$  such that for all  $\alpha > \alpha_1$ ,  $\mathbf{E}[g_\alpha(Q)] < 0$ . Since  $\mu$  has at most countably many atoms, fix  $\alpha_2 > \alpha_1$  such that  $\mu(\{\alpha_2\}) = 0$ . By (B.27) applied to  $\alpha := \alpha_2$  and (i), there exists  $N \in \mathbf{N}$  such that for every  $n \geq N$ ,

$$\mathbf{E}[g_\alpha(Q_n)] < 3\varepsilon.$$

Hence, by lemma B.22, (ii) holds.

(ii)  $\implies$  (i). For every  $n \in \mathbf{N}$ , by (B.26):

$$|m_{\mu_\infty} - m_{\mu_n}| \leq |\mathbf{E}[f_\alpha(Q_n)] - \mathbf{E}[f_\alpha(Q)]| + \mathbf{E}[g_\alpha(Q_n)] + \mathbf{E}[g_\alpha(Q)].$$

By the same reasoning, for all  $\varepsilon > 0$ , there exists some  $\alpha_1$  such that for all  $\alpha > \alpha_1$ ,  $\mathbf{E}[g_\alpha(Q)] < \varepsilon$ . Assume (ii); there exists  $\alpha_2$  such that for every  $n \in \mathbf{N}$  and  $\alpha > \alpha_2$ ,  $\mathbf{E}[g_\alpha(Q_n)] < \varepsilon$ .

Hence, take  $\alpha_3 > \max(\alpha_1, \alpha_2)$  such that  $\mu(\{\alpha_3\}) = 0$ . There exists  $N \in \mathbf{N}$  such that for every  $n \geq N$ ,  $|\mathbf{E}[f_{\alpha_3}(Q_n)] - \mathbf{E}[f_{\alpha_3}(Q)]| < \varepsilon$ . Finally,

$$|m_{\mu_n} - m_{\mu_\infty}| < 3\varepsilon.$$

Thus, (i) holds.  $\blacksquare$

## C Motivation of the alternate definitions of the Gini index (or: Why Is The Section 3.2 Useful?)

The Gini inequality index is one of the best-known mathematical objects used to measure economic inequalities. This index was first approached by Corrado Gini in his 1912 book *Variabilità e Mutabilità* [Gin12]<sup>†</sup>. In this paper, Gini considers  $n$  nonnegative, not all zero quantities  $(a_i)_{1 \leq i \leq n}$  (for instance, the incomes of  $n$  individuals) and gives several formula to express the *mean difference between the  $n$  quantities*, i.e. the value

$$\frac{1}{n^2} \sum_{1 \leq i, j \leq n} |a_i - a_j|.$$

Normalizing this mean difference of a non-negative series by its arithmetic mean and by a factor of two gives a number between 0 (included) and 1 (excluded). This value is precisely the *Gini index* as defined, without ambiguity, by economists and sociologists:

$$G((a_i)_{1 \leq i \leq n}) = \frac{1}{n} \frac{\sum_{1 \leq i, j \leq n} |a_i - a_j|}{2 \sum_{1 \leq i \leq n} a_i}.$$

The value 0 is reached if and only if the series  $(a_i)_i$  is constant, and 1 is the limit of more and more *concentrated* distributions.

If  $A$  is a random variable which follows the empirical distribution associated to<sup>‡</sup> the sample  $(a_i)_{1 \leq i \leq n}$ , then:

$$G((a_i)_i) = \frac{\mathbf{E}[|A - A'|]}{2\mathbf{E}[A]},$$

where  $\mathbf{E}$  denotes the expectation, and  $A'$  is an independent copy of  $A$ .

Modelization leads us to consider the case of variables with densities. For instance, the income of households could be, very roughly, modeled as a random variable following a gamma or a lognormal distribution<sup>§</sup>. Thus, one can extend the definition of  $G$ , and set:

$$G(A) = \frac{\mathbf{E}[|A - A'|]}{2\mathbf{E}[A]} \quad (\text{C.1})$$

for any nonnegative random variable for which this quantity is defined — the latter is equivalent to having  $0 < \mathbf{E}[A] < \infty$ . Notice that  $G$  does only depend of the distribution of  $A$ ; thus we can write  $G(\mu)$  for a probability measure  $\mu$ .

\*

<sup>†</sup>See [CV11] for a partial translation and comment in English.

<sup>‡</sup>I.e. if the probability of the event  $(A = x)$  is the number of occurrences of  $x$  in series  $(a_i)_{1 \leq i \leq n}$ , divided by  $n$ .

<sup>§</sup>See [Cho08] for a collection of articles about probability distributions for modelling household incomes.

It is well known that the Gini index has a strong relation with the *Lorenz curve*. This one is, roughly, defined as follows: “ $L(u)$  is the share of the total income earned by the  $100u$  % of the total population” — this definition is challenged in section 1.2.

Gini himself noticed in [Gin14]<sup>¶</sup> that the inequality index, for a sample  $(a_i)_i$ , is equal (up to a multiplicative factor of two) to the area between the Lorenz curve associated to this sample and the one corresponding to a situation of strict equality, i.e. the diagonal line  $y = x$ . In other words, one has:

$$G = 1 - \frac{1}{2} \int_0^1 L(u) du. \quad (\text{C.2})$$

However, Gini did only define this index for discrete distributions. Hence, he proved that (C.1) = (C.2) only when  $A$  is discrete. Some proofs were early proposed when  $A$  has a density (understand: with respect to the Lebesgue measure), for instance [KS45, section 2.33].

From this and on, it seems that most authors would use freely the result that (C.1) = (C.2). For instance, [Atk70] uses (C.2) with continuous variables as a definition; [New70] uses (C.2) as a definition for every kinds of variables; [She72] and [DSS73] use definition (C.1) with discrete variables; [Sen73] and [FRK78] use both definitions in the discrete case. [Gas72] states that the equality holds as soon as  $F$  is increasing on its support and cites [KS45] for a proof, but this latter proof does only stand if  $F$  has a density. Furthermore, while most of these authors only deal with discrete variables, the definition of the Lorenz curve they use does not correctly deal with the case of atoms, as shown in lemma 1.3.

Hence, two issues appear when using the equality (C.1) = (C.2):

1. Limiting to either distributions with density with respect to Lebesgue measure or empirical distributions of samples seems artificially restrictive. One may want to

consider more complex distributions, such as mixture models of a sample and a continuous random variable. For instance, it is sensible to model the distribution of the gross income as a mixture of a Dirac in 0 (people with no income) and a random variable with a density such as a gamma or a lognormal.

2. Having two completely different proofs for the same equality depending if we use discrete or continuous distributions is somehow unsatisfying.

The most general proof of (C.1) = (C.2) we found is [Dor79]. In this paper, Dorfman proves, in substance, that both quantities  $\mathbf{E}[|X - X'|]$  and  $\int_0^1 L(u) du$  are related to  $\int_0^\infty (1 - F(t))^2 dt$ , where  $F$  is the c.d.f. of the considered variable; he then proposes to consider

$$G = 1 - \frac{1}{\mathbf{E}[X]} \int_0^\infty (1 - F(t))^2 dt \quad (\text{C.3})$$

as a formula for the Gini coefficient, either for discrete and for continuous variables. Dorfman first defines  $G$  with equation (C.2) and derives formula (C.3) — which is a rather complex proof. Then, at the end of the article, a bit *out of the blue*, he reminds that  $G$  can also be defined through (C.1) and derives (C.3) with a really straightforward proof. Such a redacting is somehow surprising, and tends to suggest that the author assumes the equality (C.1) = (C.2) was already proven in *every case*. Furthermore, Dorfman needs to keep a few restrictions<sup>‡</sup> on  $F$ .

\*

Hence, we propose two proofs of (C.1) = (C.2) that deal with the widest possible spectrum of nonnegative distributions, i.e. every distributions on  $\mathbf{R}_+$  with finite, nonzero mean.

To achieve the first proof, we begin with deriving (C.1) = (C.3) using the straightforward argument by Dorfman. Then, we propose

<sup>¶</sup>See [Gin05] for a translation in English. The relation between the Gini index (noted  $\Delta$ ) and the area between the Lorenz curve and the diagonal (noted  $R$ ) is derived in [Gin05, section 9, pp. 27–29].

<sup>‡</sup>The set of its discontinuities needs have no point of accumulation, and  $F$  needs be differentiable between its discontinuities. It is true that most *nice* probability distributions used in economics match these constraints; however there is no *a priori* reason to be restrictive.

a different, more straightforward derivation of (C.3) = (C.2).

We also propose a second proof of the equality, more technical, based on an idea of David

Leturcq.

In order to carry this out, we need to use a convenient definition of Lorentz curve, proposed by [Gas71].

## D Alternate proofs of some results

### D.1 Two alternate proofs of the relation between Gini index and Lorenz curve

#### D.1.1 Direct computation dealing with atoms

*Proof.* We adapt the proof of theorem 3.3 presented in section 3.2.2 in the case where  $\mu$  may have atoms. It becomes necessary to make some “splits” to isolate the effects of the atoms, that will, at the end, sum up to 0.

**First split.** Let  $I := m \int_0^1 L(p) dp$ . As  $\mathcal{L}$  is diffuse, we have:

$$I = m \int_0^{1^-} L(p) dp.$$

Let  $\Lambda$  the pseudo-Lorenz function of  $\mu$ . Using lemma 1.3, we write  $I = J - K$ , with:

$$J := \int_0^{1^-} m \cdot \Lambda(p) dp ;$$

$$K := \int_0^{1^-} Q(p) \cdot [F(Q(p)) - p] dp.$$

We first deal with the integral  $K$ . As  $F$  is nondecreasing, the set of its discontinuities is at most countable. Let  $\{x_i : i \in \mathcal{A}\}$  an enumeration of this set. We note for all  $i \in \mathcal{A}$ ,  $\ell_i := F(x_i^-)$  and  $r_i := F(x_i)$ . Then:

$$[0, 1) = F(\mathbf{R}_+) \sqcup \coprod_{i \in \mathcal{A}} (\ell_i, r_i)$$

where  $\sqcup$  and  $\coprod$  denote disjoint unions.

Recall that if  $p \in F(\mathbf{R}_+)$ , then  $p = F(Q(p))$ . Hence, using the  $\sigma$ -additivity of the integral and the elementary properties of  $F$  and  $Q$ , one can write:

$$K = \sum_{i \in \mathcal{A}} \int_{\ell_i}^{r_i} Q(p) \cdot [F(Q(p)) - p] dp$$

$$= \sum_{i \in \mathcal{A}} \int_{\ell_i}^{r_i} x_i \cdot (r_i - p) dp$$

$$= \sum_{i \in \mathcal{A}} x_i \cdot \frac{(r_i - \ell_i)^2}{2}$$

$$K = \sum_{i \in \mathcal{A}} \frac{x_i \cdot \mu(\{x_i\})^2}{2}.$$

**Second split.** Now, let us consider the integral

$$J = \int_{p=0}^{1^-} \int_{u=0}^{Q(p)} u d\mu(u) dp.$$

We first isolate the upper bound of the inner integral, and we write  $J = J_1 + J_2$ , where:

$$J_1 := \int_0^{1^-} \int_{u=0}^{Q(p)^-} u d\mu(u) dp ;$$

$$J_2 := \int_0^{1^-} Q(p) \cdot \mu(\{Q(p)\}) dp.$$

The value of  $\mu(\{Q(p)\})$  is nonzero if and only if  $p$  belongs to one of the semi-open intervals  $(\ell_i, r_i]$ . If so,  $Q(p) = x_i$ . Thus, we have:

$$J_2 = \sum_{i \in \mathcal{A}} \int_{\ell_i}^{r_i} x_i \cdot \mu(\{x_i\}) dp$$

$$= \sum_{i \in \mathcal{A}} x_i \mu(\{x_i\}) (r_i - \ell_i)$$

$$= \sum_{i \in \mathcal{A}} x_i \mu(\{x_i\})^2$$

$$J_2 = 2K.$$

Let us deal with  $J_1$ . To do so, we make the same computations than in the “simplified” proof (equations (3.5) to (3.7)). We need to be careful about the endpoints of the integrals.

$$J_1 = \int_{p=0}^{1^-} \int_{u=0}^{\infty} u \mathbf{1}_{u < Q(p)} d\mu(u) dp$$



$$\begin{aligned}
&= \int_{p=0}^{1^-} \int_{u=0}^{\infty} u \mathbf{1}_{F(u) < p} d\mu(u) dp \\
&= \int_0^{\infty} \int_{p=F(u)^+}^1 u dp d\mu(u) \\
J_1 &= \int_0^{\infty} u \cdot (1 - F(u)) d\mu(u) \quad (\mathbf{D.1}) \\
&= \int_0^{\infty} u \int_{s=u^+}^{\infty} d\mu(s) d\mu(u) \\
J_1 &= \iint_{0 \leq u < s} u d\mu(u) d\mu(s). \quad (\mathbf{D.2})
\end{aligned}$$

From (D.1), we also deduce that:

$$\begin{aligned}
J_1 &= \int_0^{\infty} u d\mu(u) - \int_0^{\infty} \int_{s=0}^u u d\mu(s) d\mu(u) \\
J_1 &= m - \iint_{0 \leq s \leq u} u d\mu(s) d\mu(u).
\end{aligned}$$

**Third split.** We write  $J_1 = m - J_{11} - J_{12}$ , with:

$$\begin{aligned}
J_{11} &:= \iint_{0 \leq s < u} u d\mu(s) d\mu(u); \\
J_{12} &:= \iint_{0 \leq s = u} u d\mu(s) d\mu(u).
\end{aligned}$$

Using Fubini's theorem, we directly find

$$\begin{aligned}
J_{12} &= \int_0^{\infty} u \int_0^{\infty} \mathbf{1}_{\{u\}}(s) d\mu(s) d\mu(u) \\
&= \int_0^{\infty} u \mu(\{u\}) d\mu(u) \\
&= \sum_{i \in \mathcal{A}} x_i \mu(\{x_i\})^2 \\
J_{12} &= 2K.
\end{aligned}$$

Then, performing the permutation of variables  $u \leftrightarrow s$ , we write:

$$J_{11} = \int_{0 \leq u < s} s d\mu(u) d\mu(s).$$

We get:

$$J_1 = m - \iint_{0 \leq u < s} s d\mu(s) d\mu(u) - 2K. \quad (\mathbf{D.3})$$

**Summation.** Then, summing up the equations (D.2) and (D.3) gives us:

$$J_1 = \frac{m}{2} - \frac{1}{2} \iint_{0 \leq u < s} (s - u) d\mu(s) d\mu(u) - K.$$

The last thing to notice is that, by symmetry, we have:

$$m \cdot G(\mu) = \frac{1}{2} \iint_{\mathbf{R}_+^2} |s - u| d\mu(u) d\mu(s)$$

$$m \cdot G(\mu) = \iint_{0 \leq u < s} (s - u) d\mu(s) d\mu(u).$$

Thus,

$$J_1 = \frac{m}{2} - \frac{m}{2} \cdot G(\mu) - K.$$

Hence,

$$I = J_1 + J_2 - K = \frac{m}{2} - \frac{m}{2} \cdot G(\mu)$$

which concludes the proof in the general case.  $\blacksquare$

### D.1.2 Approximating atoms with diffuse measures 3.3

Corollary 5.5 gives another path for proving theorem 3.3.

Indeed, the proof of theorem 3.3 could be simplified in the case where  $\mu$  is nonatomic (section 3.2.2). The general case can then be proven using approximations with nonatomic measures.

We give the sketch of an alternative proof based on this idea. For what follows, we let  $G$  the Gini index as we defined with the expectation ( $\mathbf{E}[|X - X'|]/2m$ ), and  $\Gamma$  the alternative definition using Lorenz curve ( $1 - 2 \int_0^1 L(p) dp$ ):

1. Prove that  $G(\mu) = \Gamma(\mu)$  as soon as  $\mu$  has no atom. This can be done either with the simplified proof of section 3.2.2 or using [KS45]'s proof.
2. Now take  $\mu \in \mathbf{M}$ , with no assumption. We approximate  $\mu$  with nonatomic measures as follows. Consider a probability space  $(\Omega, \mathcal{F}, \mathbf{P})$  with random variables  $X, X' \sim \mu$  and  $\varepsilon_n, \varepsilon'_n \sim \mathcal{U}([0, 2^{-n}])$  (uniform over  $[0, 2^{-n}]$ ), all of them being mutually independent. Let  $\mu_n$  the distribution of  $X + \varepsilon_n$  and  $\nu_n$  the distribution of  $\varepsilon_n$ .

3. For all  $a \in \mathbf{R}_+$ , by Fubini, we have:

$$\begin{aligned}
& \mu_n(\{a\}) \\
&= \int_{\Omega} \mathbf{1}_{(X(\omega)+\varepsilon_n(\omega)=a)} d\mathbf{P}(\omega) \\
&= \iint_{\mathbf{R}_+ \times \mathbf{R}_+} \mathbf{1}_{(x+\varepsilon=a)} d(\mu \otimes \nu_n)(x, \varepsilon) \\
&= \int_{x \in \mathbf{R}_+} \int_{\varepsilon \in \mathbf{R}_+} \mathbf{1}_{(x+\varepsilon=a)} d\nu_n(\varepsilon) d\mu(x) \\
&= \int_{x \in \mathbf{R}_+} \int_{(\varepsilon=a-x)} d\nu_n(\varepsilon) d\mu(x) \\
&= \int_{\mathbf{R}_+} 0 d\mu(x) \\
&= 0.
\end{aligned}$$

Hence,  $\mu_n$  is atomless. Hence,  $G(\mu_n) = \Gamma(\nu_n)$ .

4.  $(X+Y_n)_{n \in \mathbf{N}}$  converges  $\mathbf{P}$ -almost surely to  $X$ ; so does it in distribution, i.e.  $(\mu_n)_{n \in \mathbf{N}}$  weakly converges to  $\mu$ . Furthermore,  $m_{\mu_n} = m_{\mu} + 2^{-n-1} \xrightarrow{n \rightarrow \infty} m_{\mu}$ . Hence,  $\mu_n \xrightarrow[n \rightarrow \infty]{W_1} \mu$ .

The proof of corollary 5.5 shows that  $\Gamma(\mu_n) \xrightarrow[n \rightarrow \infty]{} \Gamma(\mu)$ .

5. Consider the random variables:

$$Z_n := |X + Y_n - X' - Y'_n|.$$

One has:  $Z_n \xrightarrow[n \rightarrow \infty]{} |X - X'|$ ,  $\mathbf{P}$ -a.s. Furthermore, for every  $n \in \mathbf{N}$ ,  $0 \leq Z_n \leq X + 2$ . As  $\mu \in \mathbf{M}$ ,  $X$  has finite expectation; so does  $X + 2$ . Hence, the dominated convergence theorem states that  $(Z_n)_{n \in \mathbf{N}} \xrightarrow[n \rightarrow \infty]{} \mathbf{E}[|X - X'|]$ . We have then proven that  $G(\mu_n) \xrightarrow[n \rightarrow \infty]{} G(\mu)$ .

6. Hence,  $G(\mu) = \Gamma(\mu)$ .  $\blacksquare$

The same reasoning exactly can be performed for the equivalence of both definitions of Hoover index (proposition 3.9). Indeed, as the tricky part of the proof of this theorem is to deal with a possible atom of  $\mu$  in  $\{m_{\mu}\}$ , one can first prove the proposition assuming  $\mu$  is diffuse, and then approximating it by the same  $\mu_n$  we used for Gini.

## D.2 Convergence of the Hoover index of a sample

We provide a different, less straightforward but more elementary proof of point 3 of application 6.3, that does not need the framework of Lorenz functions nor theorem 5.5. This proof is courtesy of Guillaume Conchon-Kerjean.

*Proof.* Let  $n \in \mathbf{N}^*$ . We have:

$$H_{\hat{\mu}_n} = \frac{\sum_{i=1}^n |X_i - m_{\hat{\mu}_n}|}{2nm_{\hat{\mu}_n}}$$

The inverse triangular inequality states that for every  $i \in \{1, \dots, n\}$ ,

$$||X_i - m_{\hat{\mu}_n}| - |X_i - m_{\mu}|| \leq |m_{\hat{\mu}_n} - m_{\mu}|.$$

Thus, summing these inequalities and applying the triangular inequality,

$$\begin{aligned}
& \left| \frac{1}{n} \sum_{i=1}^n |X_i - m_{\hat{\mu}_n}| - \frac{1}{n} \sum_{i=1}^n |X_i - m_{\mu}| \right| \\
& \leq \frac{1}{n} \sum_{i=1}^n ||X_i - m_{\hat{\mu}_n}| - |X_i - m_{\mu}|| \\
& \leq |m_{\hat{\mu}_n} - m_{\mu}|.
\end{aligned}$$

On the one hand, the law of large numbers applied to random variables  $(X_n)_{n \in \mathbf{N}^*}$  states that almost surely,

$$|m_{\hat{\mu}_n} - m_{\mu}| \xrightarrow[n \rightarrow \infty]{} 0.$$

On the other hand, the law of large numbers applied to random variables  $(|X_n - m_{\mu}|)_{n \in \mathbf{N}^*}$  states that a.s.,

$$\frac{1}{n} \sum_{i=1}^n |X_i - m_{\mu}| \xrightarrow[n \rightarrow \infty]{} \mathbf{E}[|X_1 - m_{\mu}|].$$

Hence, a.s.,

$$\frac{1}{n} \sum_{i=1}^n |X_i - m_{\hat{\mu}_n}| \xrightarrow[n \rightarrow \infty]{} \mathbf{E}[|X_1 - m_{\mu}|].$$

Thus, a.s.  $H(\hat{\mu}_n) \xrightarrow[n \rightarrow \infty]{} H(\mu)$ .  $\blacksquare$

### D.3 Miscellaneous tool theorems

#### D.3.1 Uniform integrability of samples

We give a different proof of the corollary 6.5, not using either Glivenko–Cantelli’s theorem nor Scheffé’s lemma.

*Proof.* Let  $(\Omega, \mathcal{F}, \mathbf{P})$  the departure space of the random variables  $X_i$ ’s. For every  $\omega \in \Omega$  and  $n \in \mathbf{N}^*$ , we call  $\hat{\mu}_{n,\omega}$  the empirical measure over  $(\mathbf{R}_+, \mathfrak{B})$  associated with the sample  $X_1(\omega), \dots, X_n(\omega)$ , i.e.:

$$\hat{\mu}_{n,\omega} = \frac{1}{n} \sum_{i=1}^n \delta_{X_i(\omega)}.$$

Let  $A$  the set of  $\omega \in \Omega$  such that the collection  $(\hat{\mu}_{n,\omega})_{n \in \mathbf{N}^*}$  is u.i. It suffices to find a measurable subset of  $A$  of probability 1.

For  $\omega \in \Omega$  fixed, let:

$$\begin{aligned} f_{n,\omega} : \mathbf{R}_+ &\longrightarrow \mathbf{R}_+ \\ x &\longmapsto \frac{1}{n} \sum_{k=1}^n X_k(\omega) \mathbf{1}_{(X_k(\omega) > x)} \end{aligned}$$

In other words,  $f_{n,\omega}(x) = m$  is the expectation of  $Y \mathbf{1}_{\{Y > x\}}$ , where  $Y$  is any random variable of distribution  $\hat{\mu}_{n,\omega}$ . (Where  $Y$  is defined on a probability space which has nothing to do with  $(\Omega, \mathcal{F}, \mathbf{P})$ .) Thus, by definition,

$$A = \left\{ \omega \in \Omega : \sup_{n \in \mathbf{N}} f_{n,\omega}(x) \xrightarrow{x \rightarrow +\infty} 0 \right\}.$$

By lemma B.22,

$$A = \left\{ \omega \in \Omega : \begin{array}{l} \forall \varepsilon > 0, \exists N \in \mathbf{N}, \\ \forall n \geq N, f_{n,\omega}(x) < \varepsilon \end{array} \right\}.$$

Let:

$$\begin{aligned} g : \mathbf{R}_+ &\longrightarrow \mathbf{R}_+ \\ x &\longmapsto \mathbf{E} [X_1 \mathbf{1}_{(X_1 > x)}], \end{aligned}$$

and for all  $x \in \mathbf{R}_+$ ,

$$B_x := \left\{ \omega \in \Omega : f_{n,\omega}(x) \xrightarrow{n \rightarrow \infty} g(x) \right\}.$$

For all  $x \in \mathbf{R}_+$ , the law of large numbers applied to the i.i.d. random variables  $(X_n \mathbf{1}_{(X_n > x)})_{n \in \mathbf{N}^*}$  ensures that  $\mathbf{P}(B_x) = 1$ . Thus, the set

$$C := \bigcap_{n \in \mathbf{N}} B_n$$

has probability 1.

Now, we fix  $\omega \in C$ , and it suffices to prove that  $\omega \in A$ . Let  $\varepsilon > 0$ .

As  $X_i$ ’s have finite expectation, by dominated convergence theorem,  $g(a) \xrightarrow{a \rightarrow \infty} 0$ . There exists  $K \in \mathbf{N}$  such that  $g(K) < \varepsilon$ .  $\omega \in B_K$ , so there exists  $N \in \mathbf{N}^*$  such that for every  $n \geq N$ ,  $|f_{n,\omega}(K) - g(K)| < \varepsilon$ .

Then for every  $n \geq N$ ,  $f_{n,\omega}(K) < 2\varepsilon$ . Hence,  $\omega \in A$ . ■

#### D.3.2 $W_1$ convergence implies weak convergence

We now give another proof of the following result, used in appendix B.3.

**Proposition D.4.** *Let  $(\mu_n)_{n \in \mathbf{N}} \in \mathbf{M}^{\mathbf{N}}$  and  $\mu_\infty \in \mathbf{M}$ . If  $\mu_n \xrightarrow[n \rightarrow \infty]{W_1} \mu_\infty$ , then  $\mu_n \xrightarrow[n \rightarrow \infty]{\mathcal{W}} \mu_\infty$ .*

*Proof.* Apply the following lemma to the  $Q_{\mu_n}$ ’s and the characterization of weak convergence by quantiles (proposition B.13). ■

**Lemma D.5.** *Let  $I \subseteq \mathbf{R}$  an interval and consider the measured space  $(I, \mathfrak{B}, \mathcal{L})$ . Let  $E$  the set of nondecreasing, integrable functions  $E \rightarrow \mathbf{R}$ .*

*If  $\|f_\infty - f_n\|_1 \xrightarrow{n \rightarrow \infty} 0$ , then  $f_n \xrightarrow[n \rightarrow \infty]{} f$   $\mathcal{L}$ -almost everywhere.*

*Proof.* We shall first prove the following claim: for all  $f \in E$ ,  $x \in I$  such that  $f$  is continuous at  $x$  and  $\varepsilon > 0$ , there exists  $\delta > 0$  such that for all  $g \in E$ ,

$$|g(x) - f(x)| > \varepsilon \implies \|g - f\|_1 > \delta.$$

Indeed, assume that  $g(x) > f(x) + \varepsilon$ . Then:

- Either  $g(x) > f(x) + \varepsilon$ . If so, let  $\eta > 0$  such that  $f(x + \eta) < f(x) + \frac{\varepsilon}{2}$ . If  $t \in [x, x + \eta)$ , then

$$g(t) - f(t) > g(x) - f(x) - \frac{\varepsilon}{2} > \frac{\varepsilon}{2}.$$

Hence,  $\|g - f\|_{L^1} > \eta \frac{\varepsilon}{2}$ . By choosing  $\delta := \eta \frac{\varepsilon}{2}$ , the assertion holds.

- Or  $g(x) < f(x) - \varepsilon$ . Then, the same reasoning leads to the same conclusions.

Now, assume  $\|f_\infty - f_n\|_1 \xrightarrow{n \rightarrow \infty} 0$ . By the claim,  $f_n(x) \xrightarrow[n \rightarrow \infty]{} f_\infty(x)$  for all  $x \in I$  where  $f_\infty$  is continuous. Since  $f_\infty$  is nondecreasing, this is the case  $\mathcal{L}$ -almost everywhere. ■

# Contents

|           |                                                                                                                        |           |
|-----------|------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>I</b>  | <b>Alternate definitions of Lorenz curve, Gini index and Hoover index</b>                                              | <b>2</b>  |
| <b>1</b>  | <b>Alternate definitions of the Lorenz curve</b>                                                                       | <b>2</b>  |
| 1.1       | Integral of quantile definition and elementary properties                                                              | 2         |
| 1.2       | Some other intuitive definitions and how to deal with atoms                                                            | 3         |
| 1.2.1     | The pseudo-Lorenz function                                                                                             | 3         |
| 1.2.2     | Interpolling atomic Lorenz functions with nonatomic ones?                                                              | 4         |
| 1.2.3     | Kendall's parametric curve                                                                                             | 4         |
| <b>2</b>  | <b>The Lorenz curve and the mean fully characterize a distribution</b>                                                 | <b>6</b>  |
| 2.1       | Bijection between $\mathbf{M}$ and $\mathfrak{Q}$                                                                      | 6         |
| 2.2       | Bijection between $\mathbf{M}$ and $\mathfrak{L} \times \mathbf{R}_+^*$                                                | 6         |
| 2.3       | The quotient set $\mathbf{M}/\mathbf{R}_+^*$                                                                           | 9         |
| <b>3</b>  | <b>Alternate definitions of the Gini and Hoover indicators</b>                                                         | <b>9</b>  |
| 3.1       | Gini and Hoover as mean differences and mean deviation                                                                 | 9         |
| 3.2       | Gini coefficient as area between the Lorenz curve and the diagonal                                                     | 10        |
| 3.2.1     | First proof (via $\int_0^1 (1 - F(t))^2 dt$ )                                                                          | 10        |
| 3.2.2     | Second proof (direct computation) for nonatomic measures                                                               | 11        |
| 3.3       | Interpretations of the Hoover coefficient                                                                              | 11        |
| 3.3.1     | Preliminary computations                                                                                               | 11        |
| 3.3.2     | Hoover coefficient as Robin Hood index                                                                                 | 12        |
| 3.3.3     | Hoover coefficient as maximum vertical distance between the Lorenz curve and the diagonal                              | 12        |
| <b>4</b>  | <b>Some direct applications of these results</b>                                                                       | <b>13</b> |
| 4.1       | $G$ and $H$ are nondecreasing with respect to Lorenz-domination                                                        | 13        |
| 4.2       | Extreme values of Gini index under constraint on Hoover index                                                          | 13        |
| <b>II</b> | <b>Convergence properties of Lorenz curve, Gini index and Hoover index</b>                                             | <b>14</b> |
| <b>5</b>  | <b>Convergence of <math>L</math>, <math>G</math> and <math>H</math> when the measures converge in <math>W_1</math></b> | <b>16</b> |
| 5.1       | $W_1$ distance, weak convergence and uniform integrability                                                             | 16        |
| 5.2       | Necessary and sufficient conditions for convergence of Lorenz curves                                                   | 16        |
| 5.3       | Sufficient condition for Gini and Lorenz convergence                                                                   | 18        |
| 5.4       | Topological formulation: homeomorphism between $(\mathbf{M}, W_1)$ and $\mathfrak{L} \times \mathbf{R}_+^*$            | 18        |
| 5.4.1     | Basic result                                                                                                           | 18        |
| 5.4.2     | $W_1$ topology on $\mathbf{M}/\mathbf{R}_+^*$                                                                          | 18        |
| 5.4.3     | Embedding $\mathfrak{L}$ with the pointwise convergence topology                                                       | 19        |
| 5.4.4     | $G$ and $H$ as continuous applications                                                                                 | 20        |
| <b>6</b>  | <b>Applications of the <math>W_1</math> convergence</b>                                                                | <b>20</b> |
| 6.1       | Vanishing noise                                                                                                        | 20        |
| 6.2       | Sampling                                                                                                               | 21        |
| 6.3       | Quantile approximation                                                                                                 | 22        |
| 6.4       | Quantile-of-sample approximation                                                                                       | 22        |
| 6.5       | Kernel density estimation of a sample                                                                                  | 23        |

|          |                                                                                                            |           |
|----------|------------------------------------------------------------------------------------------------------------|-----------|
| <b>7</b> | <b>Weaker assumptions</b>                                                                                  | <b>25</b> |
| 7.1      | Weak convergence without $W_1$ convergence                                                                 | 25        |
| 7.1.1    | Weak convergence is not enough for convergence of $L$ , $G$ and $H$                                        | 25        |
| 7.1.2    | Topological properties of $\Phi : (\mathbf{M}, \mathscr{W}) \rightarrow \mathcal{L} \times \mathbf{R}_+^*$ | 25        |
| 7.2      | Convergence of means without weak convergence                                                              | 26        |
| 7.2.1    | Pointwise convergence of Lorenz function implies weak convergences of the measures                         | 26        |
| 7.2.2    | Uniform integrability and the infimum of the Lorenz curves                                                 | 27        |
| 7.2.3    | Conclusion                                                                                                 | 28        |
|          | <b>References</b>                                                                                          | <b>29</b> |
|          | <b>Appendices</b>                                                                                          | <b>31</b> |
| <b>A</b> | <b>Details of notations and basic definitions</b>                                                          | <b>31</b> |
| <b>B</b> | <b>Some elementary results and prerequisites</b>                                                           | <b>33</b> |
| B.1      | Quantile functions                                                                                         | 33        |
| B.1.1    | Definition and elementary properties                                                                       | 33        |
| B.1.2    | Galois inequalities and immediate applications                                                             | 33        |
| B.1.3    | Characterizing measures from quantile function                                                             | 34        |
| B.1.4    | Several properties that can be characterized through the quantile function                                 | 34        |
| B.1.5    | Extensions to $[0, 1]$ and $[0, \infty]$                                                                   | 35        |
| B.2      | Weak convergence of measures                                                                               | 35        |
| B.2.1    | Generalities                                                                                               | 35        |
| B.2.2    | The real line                                                                                              | 37        |
| B.3      | $W_1$ metric and $W_1$ convergence                                                                         | 38        |
| B.3.1    | The $W_1$ metric                                                                                           | 38        |
| B.3.2    | The $W_1$ convergence and its characterization                                                             | 39        |
| B.4      | Uniform integrability                                                                                      | 39        |
| B.4.1    | Definition                                                                                                 | 39        |
| B.4.2    | Some sufficient conditions for having uniform integrability                                                | 40        |
| B.4.3    | Operations on uniformly integrable collections                                                             | 41        |
| B.4.4    | Uniform integrability and $W_1$ convergence                                                                | 41        |
| <b>C</b> | <b>Motivation of the alternate definitions of the Gini index (or: Why Is The Section 3.2 Useful?)</b>      | <b>42</b> |
| <b>D</b> | <b>Alternate proofs of some results</b>                                                                    | <b>44</b> |
| D.1      | Two alternate proofs of the relation between Gini index and Lorenz curve                                   | 44        |
| D.1.1    | Direct computation dealing with atoms                                                                      | 44        |
| D.1.2    | Approximating atoms with diffuse measures 3.3                                                              | 45        |
| D.2      | Convergence of the Hoover index of a sample                                                                | 46        |
| D.3      | Miscellaneous tool theorems                                                                                | 47        |
| D.3.1    | Uniform integrability of samples                                                                           | 47        |
| D.3.2    | $W_1$ convergence implies weak convergence                                                                 | 47        |