Centrum voor Wiskunde en Informatica

REPORTRAPPORT

PNA

Probability, Networks and Algorithms

Probability, Networks and Algorithms

Strong laws for generalized absolute Lorenz curves when data are stationary and ergodic sequences

R. Helmers, R. Zitikis

REPORT PNA-R0407 SEPTEMBER 2004

P.O. Box 94079, 1090 GB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

ISSN 1386-3711

Strong laws for generalized absolute Lorenz curves when data are stationary and ergodic sequences

ABSTRACT

We consider generalized absolute Lorenz curves that include, as special cases, classical and generalized L - statistics as well as absolute or, in other words, generalized Lorenz curves. The curves are based on strictly stationary and ergodic sequences of random variables. Most of the previous results were obtained under the additional assumption that the sequences are weakly Bernoullian or, in other words, absolutely regular. We also argue that the latter assumption can be undesirable from the applications point of view.

2000 Mathematics Subject Classification: 60F15, 62G20 ,62G30

Keywords and Phrases: strong laws, stationary data, Lorenz curves, generalized L- statistics, ergodicity

Note: This work is carried out under project PNA 0 ' Mathematical Statistics'.

STRONG LAWS FOR GENERALIZED ABSOLUTE LORENZ CURVES WHEN DATA ARE STATIONARY AND ERGODIC SEQUENCES

ROELOF HELMERS
Centre for Mathematics and Computer Science (CWI),
Kruislaan 413, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands.
E-Mail: R.Helmers@cwi.nl

RIČARDAS ZITIKIS[1]
University of Western Ontario, Department of Statistical and Actuarial Sciences,
London, Ontario, Canada, N6A 5B7.
E-Mail: zitikis@stats.uwo.ca

We consider generalized absolute Lorenz curves that include, as special cases, classical and generalized L-statistics as well as absolute or, in other words, generalized Lorenz curves. The curves are based on strictly stationary and ergodic sequences of random variables. Most of the previous results were obtained under the additional assumption that the sequences are weakly Bernoullian or, in other words, absolutely regular. We also argue that latter assumption can be undesirable from the applications point of view.

1. INTRODUCTION AND MOTIVATION

Lorenz curves and their various functionals such as Gini indices and L-statistics have been used by econometricians to measure economic inequality for a century (cf., e.g., Lorenz, 1905, Gini, 1912). These objects (to be defined and discussed rigorously below) are usually defined in terms of the ordered values $X_{1:n}, \dots, X_{n:n}$ of some random variables $X_1, \dots, X_n$ that are frequently interpreted as incomes of n randomly selected individuals. Large sample properties of Lorenz curves, Gini indices, L-statistics, etc., have been thoroughly investigated when $X_1, X_2, \dots$ are independent and identically distributed random variables (cf., e.g., Serfling, 1980, Helmers, 1982, Shorack and Wellner, 1986, and references therein). A number of results in the area have also been obtained when the random variables $X_1, X_2, \dots$ form a stationary and ergodic sequence (cf., e.g., Aaronson, Burton, Dehling, Gilat, Hill, Weiss [ABDGHW], 1996, Gilat and Helmers, 1997, as well as the survey paper by Davydov and Zitikis, 2004b).

Serfling (1984) introduced a very general class of statistics, called generalized L-statistics. Various asymptotic properties of the generalized L-statistics were investigated by Helmers, Janssen, and Serfling [HJS] (1988), including their strong convergence when n tends to infinity. Those results were obtained when $X_1, X_2, \dots$ are independent and identically distributed random variables. This assumption was later relaxed by Gilat and Helmers (1997), assuming

[1]Partially supported by the Netherlands Organization for Scientific Research (NWO), as well as by a Discovery Research Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

that the sequence $(X_1, X_2, \dots)$ is strictly stationary and ergodic, and also that it is weakly Bernoullian or, in other words, absolutely regular (cf., e.g., Berbee, 1986, on the topic).

The goal of the present paper is to show that strong laws for the aforementioned curves and statistics can be obtained assuming only that the sequence $(X_1, X_2, \dots)$ is strictly stationary and ergodic, that is, *without* assuming that it is weakly Bernoullian (or absolutely regular). This is important since there are situations when it is easy to verify strict stationarity and ergodicity of the sequences of random variables but it might be difficult to verify, say, their weak Bernoullianity. Moreover, there are situations – and we have in mind so-called long-range dependent sequences – where weak Bernoullianity or other mixing assumptions might not even hold (cf., e.g., Rosenblatt, 1991, and references therein). We shall now introduce and discuss the main object of the present paper.

Let $(X_1, X_2, \dots)$ be a strictly stationary and ergodic sequence. Let $m \in \mathbf{N}$ be a fixed integer, and let $h : \mathbf{R}^m \to \mathbf{R}$ be a measurable function. Define

$$H_n(x) := \frac{1}{(n)_m} \sum_{*} \mathbf{I}\left\{h(X_{i_1}, \dots, X_{i_m}) \le x\right\}, \quad x \in \mathbf{R},$$

where the summation $\sum_*$ is taken over all different indices $1 \le i_1, \dots, i_m \le n$. The corresponding quantile function is defined by the formula:

$$H_n^{-1}(s) := \inf\left\{x \in \mathbf{R} : \ H_n(x) \ge s\right\}, \quad s \in (0,1).$$

We are now in the position to define the generalized absolute Lorenz curve

$$GALC_n(t) := \int_0^t H_n^{-1}(s) J(s) \mathrm{d}s, \quad t \in [0,1],$$

where $J : (0,1) \to \mathbf{R}$ is an integrable function on the interval $(0,1)$. We shall later assume some further assumptions on J, such as integrability of its certain power. We note at the outset that the (empirical) generalized absolute Lorenz curve $GALC_n$ includes, as special cases, the classical and generalized L-statistics (cf. Examples 1.1 and 1.3 below) and the absolute Lorenz curve (cf. Example 1.2 below). Several additional notes concerning $GALC_n$ follow.

In order to work out additional intuition on the just introduced generalized absolute Lorenz curve $GALC_n$, we proceed as follows. We first order all $(n)_m := n(n-1) \times \dots \times (n-m+1)$ random variables $h(X_{i_1}, \dots, X_{i_m})$ with different indices $1 \le i_1, \dots, i_m \le n$ in the non-decreasing order. Denote the ordered values by $Y_{1:(n)_m}, \dots, Y_{(n)_m:(n)_m}$. If we add the $t \times 100\%$ smallest ordered values and divide the sum by n, we get the value of $GALC_n(t)$ when the weight function $J(s)$ is equal to 1 for all $s \in [0,1]$. When $m = 1$, this coincides with the interpretation of the (classical) absolute Lorenz curve to be discussed in Example 1.2 below.

Instead of H_n we can also use the empirical distribution function D_n corresponding to the random variables $h(X_{i_1}, \dots, X_{i_m})$, $1 \le i_1, \dots, i_m \le n$, which is explicitly defined as follows:

$$D_n(x) := \frac{1}{n^m} \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} \mathbf{I}\left\{h(X_{i_1}, \dots, X_{i_m}) \le x\right\}, \quad x \in \mathbf{R}.$$

We shall show below (cf. Theorem 2.2) that, without making any additional assumption, the main result of this paper (cf. Theorem 2.1 below) also holds with D_n instead of H_n, and thus it will be shown in particular that the following version of the (empirical) generalized absolute Lorenz curve

$$GALC_n^D(t) := \int_0^t D_n^{-1}(s)J(s)\mathrm{d}s, \quad t \in [0,1],$$

converges uniformly in $t \in [0,1]$ and almost surely to a deterministic curve, $GALC_H$, to be defined later in the text.

We have already noted that the generalized absolute Lorenz curve $GALC_n$ covers several classical objects. We shall now define and discuss those objects along with related strong convergence results that are already available in the literature. We note in passing that in the current paper we do not consider other convergence modes such as weak convergence or convergence in distribution; these topics are beyond the scope of the present paper.

Example 1.1 (L-statistics)**.** Let $m = 1$ and $h(x) = 1$. Then $GALC_n(1)$ is the classical L-statistic. It can explicitly be written as

$$L_n := \sum_{i=1}^n c_{i,n} X_{i:n},$$

where $c_{i,n} := \int_{(i-1)/n}^{i/n} J(s)\mathrm{d}s$ and $X_{1:n} \leq \cdots \leq X_{n:n}$ are the order statistics corresponding to $X_1, \dots, X_n$. When $X_1, X_2, \dots$ are independent and identically distributed random variables, the L-statistic L_n converges almost surely to

$$L_F := \int_0^1 F^{-1}(s)J(s)\mathrm{d}s,$$

subject to some assumptions on F and J (cf., e.g., van Zwet, 1980, and references therein). When the sequence $(X_1, X_2, \dots)$ is strictly stationary and ergodic, the aforementioned almost sure convergence of L_n was proved by ABDGHW (1996). Gilat and Helmers (1997) noted that van Zwet's (1980) proof of the strong law for L-statistics in the case of i.i.d. observations remains valid for strictly stationary and ergodic sequences.

Example 1.2 (Absolute Lorenz curves)**.** Let $m = 1$, $h(x) = x$, and $J(s) = 1$. Then $GALC_n(t)$ is the (empirical) absolute Lorenz curve defined by the formula

$$ALC_n(t) := \int_0^t F_n^{-1}(s)\mathrm{d}s,$$

where F_n^{-1} is the quantile function corresponding to $X_1, \dots, X_n$. When $X_1, X_2, \dots$ are independent and identically distributed random variables, Goldie (1977) proved that, uniformly in $t \in [0,1]$, the (empirical) absolute Lorenz curve $ALC_n(t)$ converges almost surely to the (theoretical) absolute Lorenz curve

$$ALC_F(t) := \int_0^t F^{-1}(s)\mathrm{d}s,$$

where F^{-1} is the quantile function corresponding to the distribution function F of X_1. Davydov and Zitikis (2003) proved strong convergence under the assumption that the sequence

$(X_1, X_2, \dots)$ is strictly stationary and ergodic. For a more complete account of developments in the area, we refer to the survey paper by Davydov and Zitikis (2004b). We conclude this discussion with the note that Shorrocks (1983) and a number of subsequent authors use the term "generalized Lorenz curves" for the two curves defined above. Yitzhaki and Olkin (1991) and Shalit and Yitzhaki (1994) suggest using the term "absolute Lorenz curves," which we also use throughout the present paper, and thus use the acronyms ALC and $GALC$ when defining, respectively, absolute and generalized absolute Lorenz curves.

Example 1.3 (Generalized L-statistics)**.** The value of the curve $GALC_n(t)$ at the point $t = 1$ defines the generalized L-statistic GL_n. When $X_1, X_2, \dots$ are independent and identically distributed random variables, HJS (1988) proved that GL_n converges almost surely to

$$GL_H := \int_0^1 H^{-1}(s)J(s)\mathrm{d}s,$$

where H denotes the distribution function of the random variable $h(X_1, \dots, X_m)$, subject to some assumptions on F, J, and h. Gilat and Helmers (1997) proved the aforementioned strong convergence result assuming that the strictly stationary and ergodic sequence $(X_1, X_2, \dots)$ is weakly Bernoullian or, in other words, absolutely regular. In this case the distribution function H is that of the random variable $h(Y_1, \dots, Y_m)$, where $Y_1, \dots, Y_n$ are independent copies of X_1.

We conclude this section with the note that various generalizations and extensions of the aforementioned results are also available in the literature. For example, van Zwet (1980) and Gilat and Helmers (1997) consider (classical and generalized) L-statistics when the function J depends on the sample size n. Davydov and Zitikis (2002, 2004a) consider absolute Lorenz curves when the random variables $X_1, X_2, \dots$ are influenced by (additive and multiplicative) deterministic noises. To save the space, we do not consider such generalizations in the present paper, concentrating mainly on showing that strong convergence of the generalized absolute Lorenz curve $GALC_n$ holds for strictly stationary and ergodic sequences $(X_1, X_2, \dots)$ without imposing the assumption of absolute regularity. This situation is in agreement with the case $m = 1$ investigated thoroughly in the literature (cf., e.g., Davydov and Zitikis, 2004b, and referenced therein).

2. Main results

Let $(X_1, X_2, \dots)$ be a strictly stationary and ergodic sequence, and let F denote the distribution function of X_1. As we have noted above, our main goal is to prove that there exists a distribution function H such that, almost surely and uniformly in $t \in [0, 1]$,

$$GALC_n(t) \to GALC_H(t), \tag{2.1}$$

where the (theoretical) generalized absolute Lorenz curve $GALC_H$ is defined by the formula

$$GALC_H(t) := \int_0^t H^{-1}(s)J(s)\mathrm{d}s, \quad t \in [0, 1].$$

We shall prove below (cf. Theorem 2.1) that the function H is defined by the formula (cf. Gilat and Helmers, 1997):

$$H(x) := \mathbf{P}\left\{h(Y_1, \dots, Y_m) \le x\right\}, \tag{2.2}$$

where $Y_1, \dots, Y_n$ are independent copies of X_1.

Note that statement (2.1) follows if the integral $\int_0^1 |H_n^{-1}(s) - H^{-1}(s)|\,|J(s)| \mathrm{d}s$ converges to zero almost surely when n tends to infinity. In turn, the latter statement follows if the following two conditions are satisfied: first,

$$\int_0^1 \left|H_n^{-1}(s) - H^{-1}(s)\right|^p \mathrm{d}s \to_{a.s} 0, \quad n \to \infty, \tag{2.3}$$

and, second, $\int_0^1 |J(s)|^q \mathrm{d}s < \infty$ for the q such that $p^{-1} + q^{-1} = 1$ when $p > 1$, and $\sup_{s\in[0,1]} |J(s)| < \infty$ when $p = 1$. Naturally, we need to assume that the distribution function H has a finite pth absolute moment. This condition, however, is automatically satisfied since throughout the paper (unless explicitly stated otherwise) we assume the following, stronger one: the kernel h is such that, for all $y_1, \dots, y_m \in \mathbf{R}$,

$$|h(y_1, \dots, y_m)| \le \phi(y_1) \times \cdots \times \phi(y_m), \tag{2.4}$$

where $\phi : \mathbf{R} \to [0, \infty)$ is a function satisfying the condition $\int_{\mathbf{R}} |\phi(x)|^p \, \mathrm{d}F(x) < \infty$. We shall comment on the condition later, after our main result (cf. Theorem 2.1 below) has been formulated.

Theorem 2.1. *Let $X_1, X_2, \dots$ be a strictly stationary and ergodic sequence. Assume that h satisfies condition (2.4) and is such that the set of its discontinuity points has Q-measure zero, where Q is the m-fold product of the probability law of X_1. Then statement (2.3) holds with the distribution function H defined in (2.2). In particular, the (empirical) generalized absolute Lorenz curve $GALC_n$ converges uniformly and almost surely to the (theoretical) generalized absolute Lorenz curve $GALC_H$.*

We shall now comment on conditions of Theorem 2.1 and start with the note that condition (2.4) has already been used in the literature on U- and L-statistics (cf., e.g., ABDGHW, 1996, Gilat and Helmers, 1997) where we also find comments and examples indicating that the condition is a minor one from the practical point of view. Nevertheless, it is worth noting that – as our proofs below as well as results by, for example, Arcones (1998) and Borovkova, Burton, and Dehling [BBD] (1999) suggest – condition (2.4) can be removed at the expense of assuming that the sequence $(X_1, X_2, \dots)$ is weakly Bernoullian (or absolutely regular) in addition to being strictly stationary and ergodic. Weak Bernoullianity, however, is less desirable from the practical point of view than imposing condition (2.4) on the kernel h, since the kernel is chosen by the statistician. Furthermore, we note that Theorem U (ii) on p. 2849 in ABDGHW (1996) and our proof of Theorem 2.1 indicate that, under condition (2.4), the assumption that "the set of h discontinuity points has Q-measure zero" can be removed from Theorem 2.1 at the expense of assuming that the sequence $(X_1, X_2, \dots)$ is weakly Bernoullian. However, it is important to keep in mind that the kernel h is chosen by the statistician who, on the other hand, might not have control over the population distribution and/or the dependence structure between random variables. Hence, condition (2.4) and the aforementioned continuity condition on h should be preferred to assuming weak Bernoullianity of $(X_1, X_2, \dots)$. The final note – following BBD (1999) – concerning conditions of Theorem 2.1: If instead of almost sure convergence we are only interested in convergence in probability, then instead of condition (2.4) we can require that the family of random variables $|h(X_{i_1}, \dots, X_{i_m})|^p$, $i_1, \dots, i_m \geq 1$, is uniformly integrable. The latter condition holds if, for example, the supremum of $\mathbf{E}(|h(X_{i_1}, \dots, X_{i_m})|^{p+\delta})$ over all indices $i_1, \dots, i_m \geq 1$ is finite for some $\delta > 0$.

As we have already hinted at, a counterpart of Theorem 2.1 holds if we use the distribution function D_n instead of H_n. We now formulate this claim as the following theorem.

Theorem 2.2. *Under the assumptions of Theorem 2.1, we have that the statement*

$$\int_0^1 \left|D_n^{-1}(s) - H^{-1}(s)\right|^p \mathrm{d}s \to_{a.s} 0, \quad n \to \infty, \tag{2.5}$$

holds with the distribution function H in (2.2). In particular, the corresponding (empirical) generalized absolute Lorenz curve $GALC_n^D$ converges uniformly and almost surely to the (theoretical) generalized absolute Lorenz curve $GALC_H$.

In the following section we shall see that, at least notationally, it is easier to first prove Theorem 2.2 and then derive Theorem 2.1 using a few additional arguments.

3. PROOFS

Proof of Theorem 2.2. Bickel and Freedman (1981) showed that, for every $p \geq 1$, statement (2.5) holds if, almost surely,

$$D_n \Rightarrow H \tag{3.1}$$

and, almost surely,

$$\int_{\mathbf{R}} |x|^p \,\mathrm{d}D_n(x) \to \int_{\mathbf{R}} |x|^p \,\mathrm{d}H(x). \tag{3.2}$$

Assume for the time being that statement (3.1) holds on some subset $\Omega_0 \subseteq \Omega$ of probability one, that is, for every $\omega \in \Omega_0$, we have that $D_n^\omega \Rightarrow H$ with the obvious notation for D_n^ω. Hence,

$$\int_{\mathbf{R}} f(x)\mathrm{d}D_n^\omega(x) \to \int_{\mathbf{R}} f(x)\mathrm{d}H(x) \tag{3.3}$$

for every bounded and measurable function $f : \mathbf{R} \to \mathbf{R}$ whose set of discontinuity points has H-measure zero. We want to apply statement (3.3) in the special case when the function $f(x)$ is $|x|^p$; this would give us statement (3.2). The just noted function f is not bounded, and so we choose a positive number $K > 0$ such that both one-point sets $\{-K\}$ and $\{K\}$ have H-measure zero. We decompose the function as the sum $h_K + g_K$ of two functions $h_K(x) = |x|^p\mathbf{I}\{|x| \leq K\}$ and $g_K(x) = |x|^p\mathbf{I}\{|x| > K\}$. The function h_K is bounded, measurable, and its two discontinuity points have H-measure zero. Hence, we get from (3.3) that, for every $\omega \in \Omega_0$, the integral $\int_{\mathbf{R}} h_K(x)\mathrm{d}D_n^\omega(x)$ converges to $\int_{\mathbf{R}} h_K(x)\mathrm{d}H(x)$. In view of this fact we have that, for every $\omega \in \Omega_0$,

$$(3.4) \quad \limsup_{n\to\infty} \left| \int_{\mathbf{R}} |x|^p \,\mathrm{d}D_n^\omega(x) - \int_{\mathbf{R}} |x|^p \,\mathrm{d}H(x) \right| \le \limsup_{n\to\infty} \int_{\mathbf{R}} g_K(x)\mathrm{d}D_n^\omega(x) + \int_{\mathbf{R}} g_K(x)\mathrm{d}H(x).$$

Now we assume that there is a sequence of numbers $0 < K \to \infty$ such that each one-point set $\{K\}$ has H-measure zero. Since H has the finite pth moment, the integral $\int_{\mathbf{R}} g_K(x)\mathrm{d}H(x)$ can be made as small as desired. Next, we want to show that an analogous statement holds for $\limsup_{n\to\infty} \int_{\mathbf{R}} g_K(x)\mathrm{d}D_n^\omega(x)$ on a subset $\Omega_1 \subseteq \Omega$ that has probability one and does not depend on K. Certainly, it is enough to prove the statement for a subset $\Omega_{1,K}$ that possibly depends on K and has probability one, since the intersection of all such subsets $\Omega_{1,K}$ with respect to K has probability one. This we accomplish as follows. Using assumption (2.4), we obtain that

$$(3.5) \quad \int_{\mathbf{R}} g_K(x)\mathrm{d}D_n(x) \le \frac{1}{n^m} \sum_{i_1=1}^n \cdots \sum_{i_m=1}^n |\phi(X_{i_1})|^p \times \cdots \times |\phi(X_{i_m}))|^p \,\mathbf{I}\left\{ |\phi(X_{i_1})| \times \cdots \times |\phi(X_{i_m}))| > K \right\}.$$

The indicator on the right-hand side of (3.5) does not exceed the sum of the m indicators $\mathbf{I}\{|\phi(X_{i_1}))| > K^{1/m}\}, \dots, \mathbf{I}\{|\phi(X_{i_m}))| > K^{1/m}\}$. We now choose the first indicator and show that the quantity

$$\begin{aligned} \Delta_{n,K}(1) :=& \frac{1}{n^m} \sum_{i_1=1}^n \cdots \sum_{i_m=1}^n |\phi(X_{i_1})|^p \cdots |\phi(X_{i_m}))|^p \,\mathbf{I}\{|\phi(X_{i_1}))| > K^{1/m}\} \\ (3.6) \qquad =& \left(\frac{1}{n} \sum_{i=1}^n |\phi(X_i))|^p \right)^{m-1} \left(\frac{1}{n} \sum_{i=1}^n |\phi(X_i))|^p \,\mathbf{I}\{|\phi(X_i))| > K^{1/m}\} \right) \end{aligned}$$

can be made as small as desired by taking K sufficiently large. (Number 1 in the brackets in $\Delta_{n,K}(1)$ refers to the fact that we consider the first indicator.) We also want to specify a set of probability one on which the aforementioned statement holds. To start with, we note that, by the ergodic theorem, there is a subset $\Omega_2 \subseteq \Omega$ of probability one such that, for every $\omega \in \Omega_2$, the arithmetic mean $n^{-1} \sum_{i=1}^n |\phi(X_i^\omega))|^p$ converges to the expectation $\mathbf{E}(|\phi(X_1))|^p)$, which is finite by assumption. Furthermore, we have that there is a subset $\Omega_K(1) \subseteq \Omega$ of probability one such that, for every $\omega \in \Omega_K(1)$, the arithmetic mean $n^{-1} \sum_{i=1}^n |\phi(X_i^\omega))|^p \,\mathbf{I}\{|\phi(X_i^\omega))| > K^{1/m}\}$ converges to the expectation $\mathbf{E}(|\phi(X_1))|^p \mathbf{I}\{|\phi(X_1))| > K^{1/m}\})$, which in turn converges to zero when $K \to \infty$ since $\mathbf{E}(|\phi(X_1))|^p)$ is finite. Combining the statements above, we have that $\Omega_3 := \cap_{j=1}^m \cap_K (\Omega_2 \cap \Omega_K(j))$ is a set of probability one and such that, for every $\omega \in \Omega_3$,

$$(3.7) \qquad \lim_{K\to\infty} \limsup_{n\to\infty} \int_{\mathbf{R}} g_K(x)\mathrm{d}D_n^\omega(x) = 0.$$

From this result and also from (3.4) we have that, for every $\omega \in \Omega_0 \cap \Omega_3$, the integral $\int_{\mathbf{R}} |x|^p \, \mathrm{d}D_n^\omega(x)$ converges to $\int_{\mathbf{R}} |x|^p \, \mathrm{d}H(x)$. This completes the proof of statement (3.2).

We shall now prove statement (3.1). We start with the equations (cf., e.g., Davydov and Zitikis, 2003, for similar ideas):

$$D_n(x) = \frac{1}{n^m} \sum_{i_1=1}^n \cdots \sum_{i_m=1}^n \mathbf{I}\left\{(X_{i_1}, \ldots, X_{i_m}) \in h^{-1}(-\infty, x]\right\} = Q_n h^{-1}(-\infty, x], \tag{3.8}$$

where the probability measure Q_n is defined by the formula

$$Q_n := \frac{1}{n^m} \sum_{i_1=1}^n \cdots \sum_{i_m=1}^n \delta_{(X_{i_1}, \ldots, X_{i_m})}$$

with $\delta_{\mathbf{a}}(A)$ being 1 if $\mathbf{a} \in A$ and 0 otherwise. The program for finishing the proof of (3.1) is as follows. First, we check that $Q_n \Rightarrow Q$ a.s. for the probability measure Q specified in the formulation of Theorem 2.1. Second, we check that $Q_n f^{-1}(-\infty, x] \to Q f^{-1}(-\infty, x]$ a.s. under appropriate assumptions on f and x. Rigorous statements and proofs of these facts now follow.

We start with the note that statement $Q_n^\omega \Rightarrow Q$ follows if $\int f \mathrm{d}Q_n^\omega \to \int f \mathrm{d}Q$ for every function $f : \mathbf{R}^m \to \mathbf{R}$ which is continuous and has a compact support (cf., e.g., Billingsley, 1968, p. 41, problem 7). The class of these functions h can be restricted (cf., e.g., Davydov and Zitikis, 2004a, for more details) to a countable class of product functions

$$f = h_1 \otimes \cdots \otimes h_m, \tag{3.9}$$

where each h_i is an element of a countable class $\{\phi_l\}$ of continuous functions ϕ_l having compact supports. For the function f in (3.9) we have the equalities:

$$\begin{aligned} \int f \mathrm{d}Q_n^\omega &= \int h_1(x_1) \times \cdots \times h_m(x_m) \mathrm{d}Q_n^\omega(x_1, \ldots x_m) \\ &= \left\{ \frac{1}{n} \sum_{i=1}^n h_1(X_i^\omega) \right\} \times \cdots \times \left\{ \frac{1}{n} \sum_{i=1}^n h_m(X_i^\omega) \right\}. \end{aligned} \tag{3.10}$$

By the ergodic theorem, there is a subset $\Omega(h_k) \subseteq \Omega$ of probability one such that, for every $\omega \in \Omega(h_k)$ the arithmetic mean $n^{-1} \sum_{i=1}^n h_k(X_i^\omega))$ converges to the expectation $\mathbf{E}(h_k(X_1))$, which is finite since h_k is bounded. Since h_k is an element of $\{\phi_l\}$ and the latter class is countable, we have that the set $\Omega_4 := \cap_l \Omega(\phi_l)$ is of probability one and such that, for every $\omega \in \Omega_4$, the right-hand side of (3.10) converges to the product $\mathbf{E}(h_1(X_1)) \times \cdots \times \mathbf{E}(h_m(X_1))$, which can obviously be written as $\mathbf{E}(h_1(Y_1) \times \cdots \times h_m(Y_m))$ with $Y_1, \ldots, Y_m$ denoting independent copies of the random variable X_1. The equality

$$\mathbf{E}(h_1(Y_1) \times \cdots \times h_m(Y_m))) = \int f \mathrm{d}Q$$

implies that the measure Q is the m-fold product of the probability law $\mathcal{L}$ of X_1, as stated in the formulation of the theorem. Hence, we have verified that $Q_n^\omega \Rightarrow Q$ holds for every $\omega \in \Omega_4$. From this we in turn obtain that $Q_n^\omega f^{-1} \Rightarrow Q f^{-1}$ because the Q-measure of the discontinuity points of f is zero by assumption. This finishes the proof of Theorem 2.2. $\square$

Proof of Theorem 2.1. We shall closely follow the proof of Theorem 2.2 and make the necessary changes. Note at the outset that there were only two instances in the above proof of Theorem 2.2 where we used the special structure of the distribution function D_n. We shall now discuss these instances in detail.

The first instance occurred when we showed that $\int_{\mathbf{R}} g_K(x)\mathrm{d}D_n(x)$ can be made as small as desired by taking n and K sufficiently large. When proving Theorem 2.1, we now need to verify a similar claim for the quantity $\int_{\mathbf{R}} g_K(x)\mathrm{d}H_n(x)$. The quantity, however, does not exceed, up to a constant, $\int_{\mathbf{R}} g_K(x)\mathrm{d}D_n(x)$. Since we have already proved the smallness of the latter integral, the desired smallness of $\int_{\mathbf{R}} g_K(x)\mathrm{d}H_n(x)$ follows.

The second instance occurred in (3.10). We now need to prove an analogous statement for $\int h\mathrm{d}R_n^\omega$, where the probability measure R_n is defined by the formula

$$R_n := \frac{1}{(n)_m}\sum_{*}\delta_{(X_{i_1},\dots,X_{i_m})}.$$

In this case we have the equality

$$\int h\mathrm{d}R_n^\omega = \frac{1}{(n)_m}\sum_{*} h_1(X_{i_1}^\omega)\times\cdots\times h_m(X_{i_m}^\omega). \tag{3.11}$$

Note that the only difference between the right-hand side of (3.11) and the sum

$$\frac{1}{n^m}\sum_{i_1=1}^{n}\cdots\sum_{i_m=1}^{n} h_1(X_{i_1}^\omega)\times\cdots\times h_m(X_{i_m}^\omega) \tag{3.12}$$

is that the latter one contains terms with at least two indices equal. The sum of all such terms (which have at least two indices equal) is over-normalized by $1/n$, and hence converges to zero when n tends to infinity. Since we have already verified that the sum in (3.12) convergences to $\mathbf{E}(h_1(Y_1)\times\cdots\times h_m(Y_m))$, the proof of Theorem 2.1 is complete. □

ACKNOWLEDGMENTS

We are grateful to Youri Davydov for the techniques and ideas that we have learned from him and used in the present paper. We are also grateful to Shlomo Yitzhaki for explaining us some issues addressed in this paper, as well as for sending us several references. The work was done while the second author was visiting the Centre for Mathematics and Computer Science (CWI) in Amsterdam; the author thanks the CWI for hospitality.

REFERENCES

AARONSON, J., BURTON, R., DEHLING, H., GILAT, D., HILL, T., WEISS, B. (1996). Strong laws for L- and U-statistics. *Trans. Amer. Math. Soc.* 348, 2845–2866.

ARCONES, M. A. (1998). The law of large numbers for U-statistics under absolute regularity. *Electron. Comm. Probab.* 3, 13–19

BERBEE, H. (1986). Periodicity and absolute regularity. *Israel J. Math.* 55, 289–304.

BICKEL, P. J. AND FREEDMAN, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* 9, 1196–1217.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

BOROVKOVA, S., BURTON, R., AND DEHLING, H. (1999). Consistency of the Takens estimator for the correlation dimension. *Ann. Appl. Probab.* 9, 376–390.

DAVYDOV, Y. AND ZITIKIS, R. (2002). Convergence of generalized Lorenz curves based on stationary ergodic random sequences with deterministic noise. *Statist. Probab. Lett.* 59, 329–340.

DAVYDOV, Y. AND ZITIKIS, R. (2003). Generalized Lorenz curves and convexifications of stochastic processes. *J. Appl. Probab.* 40, 906–925.

DAVYDOV, Y. AND ZITIKIS, R. (2004a). The influence of deterministic noise on empirical measures generated by stationary processes. *Proc. Amer. Math. Soc.* 132, 1203–1210

DAVYDOV, Y. AND ZITIKIS, R. (2004b). Convex rearrangements of random elements. In: *Asymptotic Methods in Stochastics* (eds. L. Horváth and B. Szyszkowicz), pp. 141–171, Volume 44, Fields Institute Communications, American Mathematical Society, Providence, RI.

GILAT, D. AND HELMERS, R. (1997). On strong laws for generalized L-statistics with dependent data. *Comment. Math. Univ. Carolin.* 38, 187–192.

GINI, C. (1912). *Variabilitá e mutabilita.* Reprinted in: *Memorie di metodologia statistica*. (Ed. E. Pizetti and T. Salvemini; 1955) Libreria Eredi Virgilio Veschi, Rome.

GOLDIE, C. M. (1977). Convergence theorems for empirical Lorenz curves and their inverses. *Advances in Appl. Probability* 9, 765–791.

HELMERS, R. (1982). *Edgeworth Expansions for Linear Combinations of Order Statistics.* Mathematical Centre Tracts, 105. Mathematisch Centrum, Amsterdam.

HELMERS, R., JANSSEN, P. AND SERFLING, R. (1988). Glivenko-Cantelli properties of some generalized empirical DF's and strong convergence of generalized L-statistics. *Probab. Theory Related Fields* 79, 75–93.

LORENZ, M. O. (1905). Methods for measuring the concentration of wealth. *Amer. Stat. Assoc.* 9, 209–219.

ROSENBLATT, M. (1991). *Stochastic Curve Estimation.* Institute of Mathematical Statistics, Hayward, CA.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

SERFLING, R. J. (1984). Generalized L-, M-, and R-statistics. *Ann. Statist.* 12, 76–86.

SHALIT, H. AND S. YITZHAKI, S. (1994). Marginal conditional stochastic dominance. *Management Science* 40, 670-684.

SHORACK, G. R. AND WELLNER, J. A. (1986). *Empirical Processes With Applications to Statistics.* Wiley, New York.

SHORROCKS, A. J. (1983). Ranking income distributions. *Economica* 50, 3-17.

VAN ZWET, W. R. (1980). A strong law for linear functions of order statistics. *Ann. Probab.* 8, 986–990.

YITZHAKI, S. AND OLKIN, I. (1991). Concentration indices and concentration curves. In: *Stochastic Orders and Decision under Risk* (Hamburg, 1989), pp. 380–392, IMS Lecture Notes Monogr. Ser., 19, Inst. Math. Statist., Hayward, CA.