Limit theorems for L-processes

Evgeny Baklanov

Novosibirsk State University

26 August, 2020

Introduction

We consider L-processes including, as special cases, generalized Lorenz curves as well as classical L-statistics (linear combinations of functions of order statistics). These processes are based on weakly dependent random variables. Also we obtain the SLLN for bounded φ-mixing sequences of random variables and the Glivenko-Cantelli theorem for α- and φ-mixing sequences.

Introduction

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables with common distribution function F and let $X_{n:1} \leq \ldots \leq X_{n:n}$ be the order statistics based on the sample $\{X_i; i \leq n\}$. Let us consider the L-statistic

$$L_n = \frac{1}{n}\sum_{i=1}^{n} c_{n,i} H(X_{n:i}), \tag{1}$$

where $H\colon [0,1] \to \mathbb{R}$ is a measurable function, $c_{n,i}$, $i = 1, ..., n$, are some constants.

Let us consider a generalization of L-statistics (1). Namely we consider a random function on $[0,1]$ called the L-process and defined by the formula:

$$L_{n,H}(t) = \frac{1}{n}\sum_{i=1}^{n} c_{n,i}(t) H(X_{n:i}), \quad 0 \le t \le 1, \tag{2}$$

with the coefficients

$$c_{n,i}(t) = n \int_{(i-1)/n}^{i/n} J(s,t)\, ds.$$

L-processes $L_{n,H}(t)$ were considered by Greselin et al.[1] as a generalization of a wide class of statistics.

[1]Greselin, F., Puri, M. L. and Zitikis, R. *L-fuctions, proccesses, and statistics in measuring economic inequality and actuarial risks*, Statistics and Its Interface, **2** (2009), 227–245.

When t is fixed, or when the kernel $J(s,t)$ does not depend on t, then $L_{n,H}(t)$ is a classical L-statistic:

$$L_{n,H}(t) \equiv L_n = \frac{1}{n}\sum_{i=1}^{n} c_{n,i} H(X_{n:i}).$$

Strong laws for L-statistics based on the ergodic stationary and φ-mixing sequences were obtained by Baklanov[2].

[2] Baklanov, E. *Strong law of large numbers for L-statistics with dependent data*, Sib. Math. J., **47** (2006), 975–979.

Note also that if $J(s,t) = I(s \le t)$ and $H(t) = t$, then $L_{n,H}(t)$ is the (empirical) generalized Lorenz curve[3][4]:

$$L_{n,H}(t) \equiv L_n(t) = \int_0^t F_n^{-1}(s)\,ds, \quad t \in [0,1],$$

where F_n^{-1} denotes the empirical quantile function corresponding to the empirical distribution function F_n based on a sample $\{X_i; i \le n\}$.

[3] Davydov, Y. and Zitikis, R. *Generalized Lorenz curves and convexifications of stochastic processes*, J. Appl. Probab., **40** (2003), 906–925.

[4] Helmers, R. and Zitikis, R. *Strong laws for generalized absolute Lorenz curves when data are stationary and ergodic sequences*, Proc. Am. Math. Soc., **133** (2005), 3703–3712.

When $J(s,t) = (2s-1)$ and $H(t) = t$, then $L_{n,H}(t)$ is the Gini index:

$$L_{n,H}(t) = \int_0^1 (2s-1) F_n^{-1}(s)\,ds.$$

Introduction

Strong uniform convergence of $L_n(t)$ to the theoretical generalized Lorenz curve was proved for i.i.d. case by Goldie.[5]. Davydov and Zitikis[6] proved this strong convergence under the assumption that the sequence $\{X_n\}_{n\geq 1}$ is strictly stationary and ergodic.

---

[5]Goldie, C. M. *Convergence theorems for empirical Lorenz curves and their inverses*, Adv. Appl. Probab., **9** (1977), 765–791.

[6]Davydov, Y. and Zitikis, R. *Generalized Lorenz curves and convexifications of stochastic processes*, J. Appl. Probab., **40** (2003), 906–925.

Introduction

The aim of this talk is to prove strong laws for L-processes based on φ-mixing and α-mixing sequences. In our main result we give sufficient conditions for the following convergence:

$$\sup_{0\le t\le 1} |L_{n,H}(t) - L_H(t)| \to 0 \text{ a. s.},$$

where $L_H(t) = \int\limits_0^1 J(s,t)h(F^{-1}(y))\,dy$, $t \in [0,1]$.

This result immediately implies the SLLN for L-statistics and the uniform convergence of $L_n(t)$.

Introduction

Let $F^{-1}(t) = \inf\{x\colon F(x) \geq t\}$ be the quantile function corresponding to the distribution function F, and let $\{U_n\}_{n\geq 1}$ is a sequence of random variables with uniform distribution on $[0,1]$.
Since the joint distributions of the random vectors $(X_{n:1}, \ldots, X_{n:n})$ and $(F^{-1}(U_{n:1}), \ldots, F^{-1}(U_{n:n}))$ coincide, we obtain that

$$L_{n,h}(t) \stackrel{\mathrm{d}}{=} \frac{1}{n}\sum_{i=1}^{n} c_{n,i}(t) h(U_{n:i}),\ 0 \leq t \leq 1,$$

where $h(y) = H(F^{-1}(y))$ and $\stackrel{d}{=}$ denotes the equality in distribution.

Introduction

Let F_n and F_n^{-1} be the empirical and quantile distribution functions based on the sample $\{X_i;\, i \le n\}$:

$$\begin{aligned} F_n(x) &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\{X_i \le x\}, \quad x \in \mathbb{R}, \\ F_n^{-1}(t) &= \inf\{x : F_n(x) \ge t\}, \quad t \in (0,1). \end{aligned}$$

It is not difficult to see that in this case

$$L_{n,H}(t) = \int_0^1 J(s,t)h(F_n^{-1}(y))\,dy, \quad t \in [0,1].$$

Recall that

$$c_{n,i}(t) = n \int_{(i-1)/n}^{i/n} J(s,t)\,ds.$$

Define the limit process:

$$L_H(t) = \int\limits_0^1 J(s,t)h(y)dy.$$

SLLN and Glivenko-Cantelli theorem for mixing sequences

Let us define the mixing coefficients:

$$\alpha(n) = \sup_{k\geq 1} \sup\{|\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)| : A \in \mathcal{F}_1^k, B \in \mathcal{F}_{k+n}^\infty\},$$

$$\varphi(n) = \sup_{k\geq 1} \sup\{|\mathbf{P}(B|A) - \mathbf{P}(B)| : A \in \mathcal{F}_1^k, B \in \mathcal{F}_{k+n}^\infty, \mathbf{P}(A) > 0\},$$

where $\mathcal{F}_1^k$ and $\mathcal{F}_{k+n}^\infty$ denote the σ-fields generated by $\{X_i, 1 \leq i \leq k\}$ and $\{X_i, i \geq k+n\}$ respectively.
The sequence $\{X_n\}_{n\geq 1}$ is called *strong mixing* or α*-mixing* if $\alpha(n) \to 0$ as $n \to \infty$.
The sequence $\{X_n\}_{n\geq 1}$ is called *uniform mixing* or φ*-mixing* if $\varphi(n) \to 0$ as $n \to \infty$.
Note that $\alpha(n) \leq \varphi(n)$.

Theorem 1

Let $\{X_n\}_{n\geq 1}$ *be a uniformly mixing sequence of bounded random variables:* $|X_n| \leq C_n$. *Let*

$$\sum_{n=1}^{\infty} \frac{\mathsf{E}(X_n - \mathsf{E}X_n)^2}{n^2} < \infty$$

and

$$\sup_{n\geq 1} \frac{1}{n} \sum_{i=1}^{n} C_i < \infty.$$

Then

$$\frac{1}{n} \sum_{i=1}^{n} (X_i - \mathsf{E}X_i) \to 0 \quad \text{a. s.}$$

Corollary 1

Let $\{X_n\}_{n\geq 1}$ be a uniformly mixing sequence of bounded and identically distributed random variables. Then

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to \mathsf{E}X_1 \quad \text{a. s.}$$

From Corollary 1 we obtain the Glivenko–Cantelli theorem for φ-mixing sequences.

Corollary 2

Let $\{X_n\}_{n\geq 1}$ be a uniformly mixing sequence of identically distributed random variables. Then

$$\sup_{-\infty<x<\infty} |F_n(x) - F(x)| \to 0 \quad \text{a. s.}$$

The SLLN for α-mixing sequences was proved by Rio[7]:

Theorem 2

Let $\{X_n\}_{n\geq 1}$ be a strongly mixing sequence sequence of identically distributed random variables with

$$\sum_{n=1}^{\infty} \frac{\alpha(n)}{n} < \infty.$$

Then

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to \mathsf{E}X_1 \quad \textit{a. s.}$$

[7] Rio E. *A maximal inequality and dependent Marcinkiewicz-Zygmund strong laws*, Ann. Probab., **23** (1995), 918–937.

From theorem 2 we immediately obtain the Glivenko–Cantelli theorem for α-mixing sequences.

Corollary 3

Let $\{X_n\}_{n\geq 1}$ be a strongly mixing sequence sequence of identically distributed random variables with

$$\sum_{n=1}^{\infty} \frac{\alpha(n)}{n} < \infty.$$

Then

$$\sup_{-\infty<x<\infty} |F_n(x) - F(x)| \to 0 \quad \textit{a. s.}$$

Main result

Now we shall introduce the following notation:

$$C(q) = \begin{cases} \sup\limits_{0\le t\le 1} \int\limits_0^1 |J(s,t)|^q \, ds, \text{ if } 1 \le q < \infty, \\ \sup\limits_{\substack{0\le s\le 1, \\ 0\le t\le 1}} |J(s,t)|, \text{ if } q = \infty. \end{cases}$$

Now let us introduce the following conditions:

(A1) h is continious and $C(1) < \infty$;

(A2) $E|H(X_1)|^p < \infty$ and $C(q) < \infty$ for some $1 \le p < \infty, 1/p + 1/q = 1$.

Main result

And let us introduce the following conditions on strong mixing (α-mixing) coefficients:

(B1)
$$\sum_{n=1}^{\infty} \frac{\alpha(n)}{n} < \infty;$$

(B2)
$$\sum_{n=1}^{\infty} \alpha(n) \frac{(\log n)^{1/\delta}}{n} < \infty.$$

The latter condition holds if, for example, $\alpha(n) = O((\log n)^{-1-1/\delta}(\log\log n)^{-\theta})$ for some $\theta > 1$.

Main result

Theorem 3

Let $\{X_n\}_{n\geq 1}$ be a uniformly mixing sequence of identically distributed random variables. If either (A1) or (A2) hold and if

$$\sum_{n\geq 1} \varphi^{1/2}(2^n) < \infty$$

in case of (A2), then, as $n \to \infty$,

$$\sup_{0\leq t\leq 1} |L_{n,H}(t) - L_H(t)| \to 0 \quad \text{a. s.}$$

Main result

Theorem 4

Let $\{X_n\}_{n\geq 1}$ *be a strongly mixing sequence sequence of identically distributed random variables. If either* $(\mathrm{A1})$ *and* $(\mathrm{B1})$ *or* $(\mathrm{A2})$ *and* $(\mathrm{B2})$ *hold, then, as* $n \to \infty$,

$$\sup_{0\leq t\leq 1} |L_{n,H}(t) - L_H(t)| \to 0 \quad \textit{a. s.}$$

Greselin, F., Puri, M. L. and Zitikis, R. *L-fuctions, proccesses, and statistics in measuring economic inequality and actuarial risks*, Statistics and Its Interface, **2** (2009), 227–245.

Baklanov, E. *Strong law of large numbers for L-statistics with dependent data*, Sib. Math. J., **47** (2006), 975–979.

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

Goldie, C. M. *Convergence theorems for empirical Lorenz curves and their inverses*, Adv. Appl. Probab., **9** (1977), 765–791.

Helmers, R. and Zitikis, R. *Strong laws for generalized absolute Lorenz curves when data are stationary and ergodic sequences*, Proc. Am. Math. Soc., **133** (2005), 3703–3712.

Rio E. *A maximal inequality and dependent Marcinkiewicz-Zygmund strong laws*, Ann. Probab., **23** (1995), 918–937.