A uniform convergence result for weighted residual empirical processes

by
Hira L. Koul
Michigan State University

Abstract

This note establishes a uniform convergence result for a large class of weighted residual empirical processes. The underlying random variables appearing in these processes need neither be independent nor stationary.

1 Main result

Let η_i, $1 \le i \le n$ be r.v.'s, ξ_{ni}, $1 \le i \le n$ be an array of $q \times 1$ random vectors and $\gamma_{ni}, 1 \le i \le n$ be another array of r.v.'s. Define the weighted residual empirical processes

$$\begin{aligned} H_n(x,t,u) &:= n^{-1}\sum_{i=1}^n \gamma_{ni} I\Big(\eta_i \le x + n^{-1/2}xt + u'\xi_{ni}\Big), \quad t \in \mathbb{R}, u \in \mathbb{R}^q, \\ G_n(x) &:= n^{-1}\sum_{i=1}^n \gamma_{ni} I\big(\eta_i \le x\big), \quad \Gamma_n(x) := EG_n(x), \quad x \in \mathbb{R}. \end{aligned}$$

Consider the following assumptions, where all limits are taken as $n \to \infty$.

(1.1) $\quad \gamma_{ni} \ge 0, \quad \forall\, 1 \le i \le n, \qquad \sup_{n\ge1} n^{-1}\sum_{i=1}^n E\gamma_{ni} < \infty.$

(1.2) $\quad \max_{1\le i\le n} \|\xi_{ni}\| \to_p 0.$

(1.3) $\quad$ For every $x \in \mathbb{R}$, $\big|G_n(x) - \Gamma_n(x)\big| \to_p 0.$

(1.4) $\quad$ There exists a nondecreasing and continuous function $\Gamma(y)$, $y \in \mathbb{R}$, such that $0 < \Gamma(\infty) < \infty$, $\sup_{x\in\mathbb{R}} \big|\Gamma_n(x) - \Gamma(x)\big| \to 0$, and $\forall\, 0 < K < \infty$, $\sup_{x\in\mathbb{R}} \big|\Gamma(x + n^{-1/2}|x|K) - \Gamma(x)\big| \to 0$.

The following lemma gives the main result about the process H_n.

Lemma 1.1 *Under (1.1)-(1.4), the following holds. For every* $0 < K < \infty$,

(1.5) $$\mathcal{D}_n := \sup_{x\in\mathbb{R}, |t|\le K, \|u\|\le K} \big|H_n(x,t,u) - \Gamma(x)\big| \to_p 0.$$

The proof of this lemma is facilitated by the following preliminary lemma.

Lemma 1.2 *Suppose (1.1) and (1.3) hold and there exists a real valued nondecreasing right continuous function* $\Gamma(x), x \in \mathbb{R}$ *such that*

(1.6) $$0 < \Gamma(\infty) < \infty, \qquad \sup_{x\in\mathbb{R}} \big|\Gamma_n(x) - \Gamma(x)\big| \to 0.$$

Then

$$(1.7)\quad \text{(a)}\quad \Delta_n := \sup_{x\in\mathbb{R}} \big|G_n(x) - \Gamma_n(x)\big| \to_p 0, \qquad \text{(b)}\quad \sup_{x\in\mathbb{R}} |G_n(x) - \Gamma(x)| \to_p 0.$$

Proof. We prove only (1.7)(a) while (1.7)(b) follows trivially from (1.6). Let $\Delta_n(x) := G_n(x) - \Gamma_n(x)$. By (1.3), $|\Delta_n(x)| \to_p 0$, for every fixed $x \in \mathbb{R}$. To prove that this holds uniformly in $x \in \mathbb{R}$, let $k > 0$ be positive integer, $\delta > 0$ and $-\infty = x_0 < x_1 < \cdots < x_k < x_{k+1} = \infty$ be such that $\max_{0\le j\le k+1} \big(\Gamma(x_j) - \Gamma(x_{j-1})\big) \le \delta$. This is possible because $\Gamma(\cdot)/\Gamma(\infty)$ is a distribution function. Note that $\Delta_n = \max_{1\le j\le k+1} \sup_{x_{j-1}<x\le x_j} \big|\Delta_n(x)\big|$. Because $\gamma_{ni} \ge 0$, $G_n(y)$, $\Gamma_n(y)$ are nondecreasing functions of $y \in \mathbb{R}$. Hence, for $x_{j-1} < x \le x_j$,

$$\Delta_n(x_{j-1}) + \Gamma_n(x_{j-1}) - \Gamma_n(x_j) \le \Delta_n(x) \le \Delta_n(x_j) + \Gamma_n(x_j) - \Gamma_n(x_{j-1}).$$

In other words,

$$\Delta_n \le \max_{1\le j\le k+1} \big|\Delta_n(x_j)\big| + \max_{1\le j\le k+1} \big|\Delta_n(x_{j-1})\big| + 2 \max_{1\le j\le k+1} \big|\Gamma_n(x_j) - \Gamma_n(x_{j-1})\big|.$$

By (1.3), the first two terms in this bound tend to zero, in probability, for every fixed k. Moreover, by the triangle inequality, the third term is bounded from the above by

$$4 \sup_{x\in\mathbb{R}} \big|\Gamma_n(x) - \Gamma(x)\big| + 2 \max_{1\le j\le k+1} \big|\Gamma(x_j) - \Gamma(x_{j-1})\big| \le 4 \sup_{x\in\mathbb{R}} \big|\Gamma_n(x) - \Gamma(x)\big| + 2\delta.$$

By (1.6), the first term in this bound tends to zero. The proof of the Lemma 1.2 is completed upon letting $\delta \to 0$ in this bound.

Proof of Lemma 1.1. Fix a $0 < K < \infty$. Let $A_{n,\delta} := \big\{ \max_{1\le i\le n} \|\xi_{ni}\| \le \delta \big\}, \delta > 0$. By (1.2), for every $\delta > 0$, there exists $N \equiv N_\delta < \infty$ such that $P(A_{n,\delta}) \ge 1 - \delta$, for all $n \ge N$. Because the indicator function $I(\eta_i \le z)$ is nondecreasing in $z \in \mathbb{R}$, we obtain that on $A_{n,\delta}$ and for all $|t| \le K$, $\|u\| \le K$, $1 \le i \le n$,

$$I(\eta_i \le x - n^{-1/2}|x|K - K\delta) \le I\big(\eta_i \le x + n^{-1/2}xt + u'\xi_{ni}\big) \le I\big(\eta_i \le x + n^{-1/2}|x|K + K\delta\big).$$

Multiply all sides of these inequalities by $\gamma_{ni} \ge 0$, sum over $1 \le i \le n$ and divide all sides by n to obtain

$$G_n(x - n^{-1/2}|x|K - K\delta) \le H_n(x, t, u) \le G_n(x + n^{-1/2}|x|K + K\delta), \quad \forall\, x \in \mathbb{R}, n \ge 1.$$

Hence,

$$(1.8)\ \mathcal{D}_n \le \sup_{x\in\mathbb{R}} \Big|G_n(x + n^{-1/2}|x|K + K\delta) - \Gamma(x)\Big| + \sup_{x\in\mathbb{R}} \Big|G_n(x - n^{-1/2}|x|K - K\delta) - \Gamma(x)\Big|.$$

Consider the first term of the upper bound (1.8). Decompose

$$\begin{aligned} G_n(x + n^{-1/2}|x|K + K\delta) - \Gamma(x) = &\Big[G_n(x + n^{-1/2}|x|K + K\delta) - \Gamma_n(x + n^{-1/2}|x|K + K\delta)\Big] \\ &+ \Big[\Gamma_n(x + n^{-1/2}|x|K + K\delta) - \Gamma_n(x + n^{-1/2}|x|K)\Big] \\ &+ \Big[\Gamma_n(x + n^{-1/2}|x|K) - \Gamma(x)\Big], \end{aligned}$$

so that

$$
\begin{aligned}
&\sup_{x\in\mathbb{R}} \big|G_n(x + n^{-1/2}|x|K + k\delta) - \Gamma(x)\big| \\
\leq\ & \sup_{y\in\mathbb{R}} \big|G_n(y) - \Gamma_n(y)\big| + \sup_{y\in\mathbb{R}} \big|\Gamma_n(y + K\delta) - \Gamma_n(y)\big| + \sup_{x\in\mathbb{R}} \big|\Gamma_n(x + n^{-1/2}|x|K) - \Gamma(x)\big|.
\end{aligned}
$$

By (1.7), the first term in this bound tends to zero, in probability. Next, consider

$$
\begin{aligned}
&\sup_{x\in\mathbb{R}} \big|\Gamma_n(x + n^{-1/2}|x|K) - \Gamma(x)\big| \\
&\quad \leq\ \sup_{y\in\mathbb{R}} \big|\Gamma_n(y) - \Gamma(y)\big| + \sup_{x\in\mathbb{R}} \big|\Gamma(x + n^{-1/2}|x|K) - \Gamma(x)\big| \to 0,
\end{aligned}
$$

by (1.4).

Next,

$$
\sup_{y\in\mathbb{R}} \big|\Gamma_n(y + K\delta) - \Gamma_n(y)\big| \leq 2\sup_{y\in\mathbb{R}} \big|\Gamma_n(y) - \Gamma(y)\big| + \sup_{y\in\mathbb{R}} \big|\Gamma(y + K\delta) - \Gamma(y)\big|.
$$

The first term in this bound tends to zero by (1.4) as $n \to \infty$. By the uniform continuity of Γ, implied by (1.4), the second term tends to zero as $\delta \to 0$.

This then shows that the first term in the upper bound of (1.8) tends to zero, in probability. A similar arguments shows that the same thing holds for the second term in the upper bound of (1.8), thereby completing the proof of (1.5) and Lemma 1.1.

The above lemmas extend to the situation where the weights γ_{ni} are not nonnegative. More precisely, let ζ_{ni} real r.v.'s and define

$$
\begin{aligned}
U_n(x,t,u) &:= n^{-1}\sum_{i=1}^{n} \zeta_{ni} I\Big(\eta_i \leq x + n^{-1/2}xt + u'\xi_{ni}\Big), \quad t\in\mathbb{R}, u\in\mathbb{R}^q, \\
V_n(x) &:= n^{-1}\sum_{i=1}^{n} \zeta_{ni} I\big(\eta_i \leq x\big), \quad x\in\mathbb{R}.
\end{aligned}
$$

For $y\in\mathbb{R}$, let $y = \max(0,y), y^- = \max(0,-y)$ so that $y = y^+ - y^-$. Note that both $y^\pm \geq 0$. Write $\zeta_{ni} = \zeta_{ni}^+ - \zeta_{ni}^-$. Let $U_n^\pm$, $V_n^\pm$ denote U_n, V_n, where $\{\zeta_{ni}\}$ are replaced by $\{\zeta_{ni}^\pm\}$, respectively. Then $U_n(x,t,u) \equiv U_n^+(x,t,u) - U_n^-(x,t,u)$, $V_n(x) \equiv V_n^+(x) - V_n^-(x)$. Define $\nu_n^\pm(x) := EV_n^\pm(x)$, $\nu_n(x) \equiv \nu_n^+(x) - \nu_n^-(x)$. We are now ready to state the following lemma.

Lemma 1.3 *Suppose* (1.2) *and the following assumptions hold.*

(1.9) $\displaystyle\sup_{n\geq 1} n^{-1}\sum_{i=1}^{n} E\zeta_{ni}^\pm < \infty.$

(1.10) *For every* $x\in\mathbb{R}$, $\big|V_n^\pm(x) - \nu_n^\pm(x)\big| \to_p 0.$

(1.11) *There exists a nondecreasing and continuous function* $\nu^\pm(x)$, $x\in\mathbb{R}$, *such that* $0 < \nu^\pm(\infty) < \infty$, $\sup_{x\in\mathbb{R}} \big|\nu_n^\pm(x) - \nu^\pm(x)\big| \to 0$, *and* $\forall\, 0 < K < \infty$, $\sup_{x\in\mathbb{R}} \big|\nu^\pm(x + n^{-1/2}|x|K) - \nu^\pm(x)\big| \to 0.$

Then the following hold, where $\nu = \nu^+ - \nu^-$.

$$\sup_{x\in\mathbb{R}} \big|V_n(x) - \nu_n(x)\big| \to_p 0, \quad \sup_{x\in\mathbb{R}} \big|V_n(x) - \nu(x)\big| \to_p 0,$$
$$\sup_{x\in\mathbb{R}, |t|\le K, \|u\|\le K} \big|U_n(x,t,u) - \nu(x)\big| \to_p 0, \quad \forall\, 0 < K < \infty.$$

The proof follows from Lemmas 1.1 and 1.2 applied with $\gamma_{ni} \equiv \zeta_{ni}^{\pm}$ and the triangle inequality.

The above proofs show that the convergence in probability in the above lemmas can be replaced by the almost sure convergence, if (1.2) and (1.3) hold almost surely. Also note that there were no probabilistic assumptions like independence or stationarity, etc., made on the given r.v.'s.

The above uniform consistency results have roots in the literature. See for example Lemma 3.1 in Chang (1990), Lemma 3.4 in Stute, Thies and Zhu (1998), Lemma 4.2 in Koul and Stute (1999) and the monograph of Koul (2002).

References

Chang, N. M. (1990). Weak convergence of a self-consistent estimator of a survival function with doubly censored data. *Ann. Statist.*, **18**, 391–404.

Stute, W., Thies, S. and Zhu, Li-Xing. (1998). Model checks for regression: an innovation process approach. *Ann. Statist.*, **26**, 1926–1934.

Koul, H.L. and Stute, W. (1999). Nonparametric model checks for time series. *Ann. Statist.*, **27**, 204–236.

Koul, Hira L. (2002). *Weighted empirical processes in dynamic nonlinear models.* Second edition of *Weighted empiricals and linear models* [Inst. Math. Statist., Hayward, CA, 1992; MR1218395]. Lecture Notes in Statistics, **166**. Springer-Verlag, New York.