2.2. Ergodic Theorem and SLLN-SIP Theorem

- Ergodic Theorem

Theorem 2.1. *Assume that $\{X_t\}$ is a sequence of strictly stationary and ergodic time series with $E|X_t| < \infty$. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} X_t = EX_t \ \ a.s..$$

The ergodic theorem plays an essential role in the field of time series.

- SLLN Theorem

  When X_t's are i.i.d.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=-n}^{-1} X_t = EX_t \ \ a.s.$$

  and by the laws of the iterated logarithm,

$$\frac{1}{n} \sum_{t=1}^{n} X_t - EX_t = O[(\frac{\log\log n}{n})^{1/2}] \ \ a.s..$$

  Do we have these when $\{X_t\}$ is a sequence of strictly stationary and ergodic, but not i.i.d.?

  We now introduce a more general concept.

Let $\{\varepsilon_t\}$ be a series of independent random variables (or vectors) on the probability space $(\Omega, \mathcal{B}, P)$, $\mathcal{F}_t = \sigma\{\varepsilon_t, \varepsilon_{t-1}, \cdots\}$ and X_t be a $\mathcal{F}_t-$measureable $m \times 1$ random vector for $t = 0, \pm 1, \cdots$. We first introduce the following definition:

Definition 2.1. Let $\mathcal{F}_i(j)$ be the $\sigma-$field generated by $\{\varepsilon_j, \varepsilon_{j-1}, \cdots, \varepsilon_{j-i+1}\}$ with $i \geq 1$, and $\mathcal{F}_0(j) = \{\emptyset, \Omega\}$. $\{X_t\}$ is said to be $L^p(\nu)$ NED in terms of $\{\varepsilon_t\}$ if

$$\sup_{-\infty<t<\infty} \|X_t\|_p < \infty$$

and

$$\sup_{-\infty<t<\infty} \|X_t - E[X_t|\mathcal{F}_k(t)]\|_p = O(k^{-\nu}),$$

where $p \geq 1$ and $\nu > 0$.

This notion of NED sequence extends a concept introduced in Billingsley (1968). Some different versions appear in McLeish

(1975), Wooldridge and White (1988) and Pötscher and Prucha (1991). This NED $\{X_t\}$ implies that it is mixingale, i.e.,

$$\sup_{-\infty<t<\infty} \|EX_t - E(X_t|\mathcal{F}_{t-k})\|_p = O(k^{-\nu}).$$

Our SLLN is as follows.

Theorem 2.2. *Let $\{X_t : t = 0, \pm 1, \cdots\}$ be $L^{1+\iota}(\nu)$ NED and mean zero sequence in terms of $\{\varepsilon_t\}$ with $\iota > 0$ and $\nu > 0$. Then there exists a constant $\delta > 0$ such that*

$$(a) \qquad \frac{1}{k}\sum_{t=1}^{k} X_t = o(\frac{1}{k^\delta}) \ a.s.,$$

$$(b) \qquad \frac{1}{k}\sum_{t=-k}^{-1} X_t = o(\frac{1}{k^\delta}) \ a.s..$$

The moment condition in Theorem 2.2 is only slightly stronger than that in the ergodic theorem for the forward sums. But our SLLN includes a rate of convergence, while the ergodic theorem does not. We guess that this is the weakest moment condition for the NED sequence if a rate of convergence is asked.

The independence of $\{\varepsilon_t\}$ can be replaced by some mixing conditions. If we allow $\iota \geq 1$ and $\nu \geq 0.5$, then a sharper rate of convergence may be obtained, see e.g. Hall and Heyde (1980, p. 41). If we assume $\iota = 1$ and use the moment bound of Ing and Wei (2005), then a relationship between the rate of convergence and the series dependence can be given.

Our strong invariance principle (SIP) is as follows.

Theorem 2.3. *Let X_t be a martingale difference in terms of $\mathcal{F}_t$ with covariance matrix Ω and be $L^{2+\iota}(\nu)$ NED in terms of $\{\varepsilon_t\}$ with $\iota > 0$, where either $2\nu > 1$ or $2\nu = 1$ and there exist constants $\nu_1 > 0$ and $\iota_1 > 0$ with $2\nu_1 > 1$ such that*

$$\sup_{-\infty<t<\infty} \|E[X_t|\mathcal{F}_{k+1}(t)] - E[X_t|\mathcal{F}_k(t)]\|_{2+\iota_1} = O(k^{-\nu_1}). \quad (1)$$

Then, without changing its distribution, we can redefine $\{X_t\}$ on two richer probability spaces together, respectively, with two sequences of i.i.d. $m \times 1$ normal vectors with mean zero and covariance matrix Ω, $\{G_{1t} : t = 1, 2, \cdots\}$ and $\{G_{2t} : t = 1, 2, \cdots\}$, *such that, for some constant* $\delta > 0$, *we have*

$$(a) \qquad \sum_{t=1}^{k} X_t = \sum_{t=1}^{k} G_{1t} + O(k^{\frac{1}{2}-\delta}) \ a.s.,$$

$$(b) \qquad \sum_{t=-k}^{-1} X_t = \sum_{t=1}^{k} G_{2t} + O(k^{\frac{1}{2}-\delta}) \ a.s.$$

Two richer probability spaces may be different, for which we refer to Berkes and Philipp (1979) and Eberlein (1986). Theorems 2.2-2.3 don't require $\{X_t\}$ to be stationary and can be extended for triangular arrays as in Andrews (1988).