The Law of the Iterated Logarithm
For a Stationary Process

Michael Woodroofe
The University of Michigan
Joint work with Ou Zhao

Outline

Classical Limit Theorems.

Stationary Processes.

The Conditional Central Limit Theorem.

The Law of the Iterated Logarithm.

The Functional Version.

The Proof.

Questions.

Limit Theorems

Let $X_1, X_2, \cdots$ be independent and identically distributed random variables defined on a probability space $(\Omega, \mathcal{A}, P)$. Thus,

$$P[X_k \in B_k, k = 1, ..., n] = \prod_{k=1}^{n} P[X_1 \in B_k]$$

for Borel sets $B_k \subseteq I\!R$. Suppose that X_k have a finite mean and (not necessarily finite) variance

$$E(X_k) = \mu \qquad \text{and} \qquad \sigma^2 = E[(X_k - \mu)^2],$$

and let

$$S_n = X_1 + \cdots + X_n.$$

The Law of Large Numbers.

$$\lim_{n\to\infty} \frac{S_n}{n} = \mu \ w.p.1.$$

The Central Limit Theorem. If $0 < \sigma^2 < \infty$, then

$$\lim_{n\to\infty} P\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \le z\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}x^2} dx = \Phi(z),$$

The Law of the Iterated Logarithm. If $\sigma^2 < \infty$, then

$$\limsup_{n\to\infty} \frac{S_n - n\mu}{\sqrt{2n \log\log(n)}} = \sigma \ w.p.1.$$

Note: Physical interpretation of μ and σ.

Corollary For large n,

$$P\left[|\frac{S_n - n\mu}{\sigma\sqrt{n}}| \le 3\right] \approx \frac{1}{\sqrt{2\pi}}\int_{-3}^{3} e^{-\frac{1}{2}x^2}dx = .9974 \approx 1.$$

Corollary. If $0 < \sigma^2 < \infty$, then

$$P\left[\limsup_{n\to\infty} |\frac{S_n - n\mu}{\sigma\sqrt{n}}| > 3\right] = 1.$$

Remarks

- A single n versus for some n
- Statistical Paradox

A Statistical Paradox

Recall: $X_1, X_2, \cdots$ are i.i.d. with mean μ and variance σ^2.

An Hypothesis: $\mu = \mu_0$ and $\sigma = \sigma_0$.

A Test: Reject if

$$|\frac{S_n - n\mu_0}{\sigma_0\sqrt{n}}| > 3.$$

Then

$$P_0[\text{Reject}] < .01.$$

Optional Stopping. Now let N be the smallest $n \ge 1$ for which $|S_n - n\mu_0| > 3\sigma_0\sqrt{n}$. Then

$$N < \infty \ w.p.1 \qquad \text{and} \qquad P_0[\text{Reject}] = 1.$$

Example: ESP

Question: To Bayes or not to Bayes.

Stationary Processes

Def. A sequence $\cdots X_{-1}, X_0, X_1, \cdots$ is stationary iff

$$P[(X_{n+1}, \cdots, X_{n+m}) \in B]$$

is independent of n for all Borel sets $B \subseteq I\!R^m$ for all m.

Sequence Space: Let $\mathbf{X} = (\cdots, X_{-1}, X_0, X_1, \cdots)$, and

$$Q(B) = P[\mathbf{X} \in B]$$

for Borel sets $B \subseteq I\!R^Z$; and let θ be the shift operator

$$\theta(\cdots x_{-1}, x_0, x_1, \cdots) = (\cdots, x_0, x_1, x_2, \cdots).$$

Then the sequence is stationary iff $Q \circ \theta^{-1} = Q$.

Ergodicity: A stationary sequence is ergodic if $Q(A) = 0$ or 1 whenever $\theta^{-1}(A) = A$.

Examples

Let $\cdots X_{-1}, X_0, X_1, \cdots$ be independent and identically distributed; and let $\psi : R^Z \to I\!R$ be measurable. Then

$$Y_k = \psi(\cdots X_{k-1}, X_k, X_{k+1}, \cdots)$$

is stationary and ergodic.

Martingale Differences

Let $\cdots X_{-1}, X_0, X_1, \cdots$ be a sequence of random variables and $\mathcal{F}_k = \sigma\{\cdots X_{k-1}, X_k\}$. The random variables are martingale differences if

$$E|X_k| < \infty \qquad \text{and} \qquad E(X_k|\mathcal{F}_{k-1}) = 0 \ w.p.1$$

for all k in which case $M_n = X_1 + \cdots + X_n$ is called a martingale.

Example: If $\cdots X_{-1}, X_0, X_1, \cdots$ are independent and identically distributed (IID) with $E(X_k) = 0$, then they form martingale differences.

The CLT and LIL: The CLT and LIL are valid for stationary sequences of martingale differences in which case $\mu = 0$.

Notation: $\|Y\| = \sqrt{E(Y^2)}$.

A CCLT For Stationary Sequences

Let $\cdots X_{-1}, X_0, X_1, \cdots$ be a statonary, ergodic sequence with $E(X_k) = 0$ and $E(X_k^2) < \infty$; let $\mathcal{F}_k = \sigma\{\cdots X_{k-1}, X_k\}$; and let $S_n = X_1 + \cdots + X_n$. If

$$\sum_{n=1}^{\infty} n^{-\frac{3}{2}} \|E(S_n|\mathcal{F}_0)\| < \infty, \qquad (\dagger)$$

then

$$\sigma^2 = \lim_{n\to\infty} \frac{1}{n} E(S_n^2) \qquad (*)$$

exists (finite), and

$$\lim_{n\to\infty} P\left[\frac{S_n}{\sqrt{n}} \le z|\mathcal{F}_0\right] = \Phi(\frac{z}{\sigma}) \qquad (**)$$

in probability for all z. Conversely if (*) and (**), then

$$\|E(S_n|\mathcal{F}_0)\| = o(\sqrt{n}).$$

About the Condition: $\sum_{n=1}^{\infty} n^{-\frac{3}{2}} \|E(S_n|\mathcal{F}_0)\| < \infty$.

- Limits the dependence.
- Solely in terms of $\|E(S_n|\mathcal{F}_0)\|$.
- Best Possible.
- Compare with strong mixing.

Examples: Bernoulli shifts; Hasting Metroplis; linear proceses.

Refs: Maxwell and W (2000), *Ann. Prob.*

Peligrad and Utev (2005) *Ann. Prob.*

About the Proof

In the proof it is shown that

$$S_n = M_n + R_n,$$

where M_n and R_n are $\mathcal{F}_n$-measurable, M_n is the sum of a stationary sequence of martingale differences, and $\|R_n\| = o(\sqrt{n})$.

Connection to LIL: If

$$R_n = o[\sqrt{n \log\log(n)}] \ w.p.1,$$

then the LIL would hold for S_n.

Slow Variation

A function $\ell : (0,\infty) \to (0,\infty)$ varies slowly at ∞ iff

$$\lim_{t\to\infty} \frac{\ell(tx)}{\ell(t)} = 1$$

for all $0 < x < \infty$.

Notation: Let

$$\ell^*(n) = \sum_{j=1}^{n} \frac{1}{j\ell(j)}.$$

Example: If $\ell(n) = \log(n+1)$, then $\ell^*(n) \sim \log\log(n)$.

A Law of the Iterated Logarithm

Let $\cdots X_{-1}, X_0, X_1, \cdots$ be a stationary, ergodic sequence with $E(X_k) = 0$ and $E(X_k^2) < \infty$; let $\mathcal{F}_k = \sigma\{\cdots X_{k-1}, X_k\}$; and let $S_n = X_1 + \cdots + X_n$. If $\ell : (0,\infty) \to (0,\infty)$ is a non-decreasing, slowly varying function, for which

$$\sum_{n=1}^{\infty} n^{-\frac{3}{2}} \log(n)\sqrt{\ell(n)}\|E(S_n|\mathcal{F}_0)\| < \infty, \qquad (\dagger)$$

then

$$\lim_{n\to\infty} \frac{R_n}{\sqrt{n\ell^*(n)}} = 0 \ w.p.1.$$

Corollary 1. If $(\dagger)$ holds with $\ell(n) = \log(n+1)$,

$$\sum_{n=1}^{\infty} n^{-\frac{3}{2}} \log^{\frac{3}{2}}(n)\|E(S_n|\mathcal{F}_0)\| < \infty, \qquad (\dagger)$$

then

$$\lim_{n\to\infty} \frac{R_n}{\sqrt{n\log\log(n)}} = 0 \ w.p.1,$$

and

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\log\log(n)}} = \sigma \ w.p.1.$$

Corollary 2. If

$$\sum_{n=1}^{\infty} n^{-\frac{3}{2}} \log(n)\sqrt{\ell(n)}\|E(S_n|\mathcal{F}_0)\| < \infty, \qquad (\dagger)$$

for some ℓ for which $1/[n\ell(n)]$ is summable, then

$$\lim_{n\to\infty} \frac{R_n}{\sqrt{n}} = 0 \ w.p.1,$$

and there is convergence $w.p.1$ in the Conditional Central Limit Theorem. That is,

$$\lim_{n\to\infty} P\left[\frac{S_n}{\sqrt{n}} \le z|\mathcal{F}_0\right] = \Phi(\frac{z}{\sigma}) \ w.p.1. \qquad (*)$$

The Functional Version

Brownian Motion. Let $\mathbb{B}$ denote standard Brownian motion. Thus, $\mathbb{B}(t),\ 0 \le t \le 1$, is a stochastic process with

- independent increments;
- continuous sample paths
- $\mathbb{B}(0) = 0$
- $\mathbb{B}(t) - \mathbb{B}(s)$ is normally distributed with mean 0 and variance $t - s$.

Sample Versions. Let $X_1, X_2, \cdots$ be random variables, $S_n = X_1 + \cdots + X_n$, and let $\mathbb{B}_n$ be a continuous piecewise linear function for which

$$\mathbb{B}_n(\frac{k}{n}) = \frac{1}{\sqrt{n}} S_k,\ \ k = 0, 1, \cdots, n.$$

Strassen's LIL

Let $C[0,1]$ be the set of continuous functions $f : [0,1] \to \mathbb{R}$, and let $K \subseteq C[0,1]$ be the set of absolutely continuous f for which

$$\int_0^1 f'(t)^2 dt \le 1.$$

If $X_1, X_2, \cdots$ are i.i.d. with mean 0 and variance one, then the set of limit points of

$$\frac{\mathbb{B}_n}{\sqrt{2n \log\log(n)}}, n \ge 3,$$

is K $w.p.1.$.

Corollary 3. Let $\cdots X_{-1}, X_0, X_1, \cdots$ is be stationary process with mean 0 and finite variance. If

$$\sum_{n=1}^{\infty} n^{-\frac{3}{2}} \log^{\frac{3}{2}}(n) \|E(S_n|\mathcal{F}_0)\| < \infty, \qquad (\dagger)$$

then $w.p.1$ the set of limit points of $\mathbb{B}_n/\sqrt{2n \log\log(n)},\ n \ge 3$, is σK, where

$$\sigma^2 = \lim_{n\to\infty} \frac{1}{n} E(S_n^2).$$

Proof. Follow from Strassen's Theorem (for martingale differences) and $R_n = o[\sqrt{2n \log\log(n)}]$.

About the Proof
Markov Chains

There is no loss of generality in supposing that

$$X_k = g(W_k),$$

where $\cdots W_{-1}, W_0, W_1 \cdots$ is a stationary Markov chain with values in a measurable space $\mathcal{W}$, for we may let

$$W_k = (\cdots, X_{k-1}, X_k).$$

Let π and Q denote the stationary distribution and transition function of the chain,

$$\pi\{B\} = P[W_k \in B]$$

and

$$Q(w; B) = P[W_{n+1} \in B | W_n = w].$$

Some Operators

Let

$$Qf(w) = \int_{\mathcal{W}} f(z)Q(w;dz) = E[f(W_{n+1})|W_n = w] \ a.e.$$

for $f \in L^1(\pi)$. Then,

$$E(X_k|W_0) = Q^k g(W_0)$$

and

$$E(S_n|W_1) = \sum_{k=1}^{n} Q^{k-1} g(W_1) = V_n g(W_1).$$

More Operators

Next, let θ denote the shift operator, $W_n \circ \theta = W_{n+1}$ and

$$Tf = f \circ \theta.$$

Next, let

$$\beta_k = \frac{c}{k} \sum_{n=k}^{\infty} \frac{1}{\sqrt{n^3 \ell(n)}} \sim \frac{2c}{\sqrt{k^3 \ell(k)}}$$

where c is so chosen that $\beta_1 + \beta_2 + \cdots = 1$, and

$$B(z) = \sum_{k=1}^{\infty} \beta_k z^k.$$

Then $B(z)$ is continuous in $|z| \le 1$ and analytic in $|z| < 1$.

The Operator Norm. If $S : L^2(P) \to L^2(P)$ is a continuous linear transformation, let

$$\|S\|_{op} = \sup_{\|Y\| \le 1} \|SY\|.$$

The Operator $B(T)$. Since β_k are summable,

$$B(T) = \sum_{k=1}^{\infty} \beta_k T^k$$

converges in the operator norm. Let

$$A(z) = \frac{1}{1 - B(z)} = \sum_{k=0}^{\infty} \alpha_k z^k.$$

Then A is continuous in $|z| \le 1 \ne z$. A major problem is to make sense of $A(T)$.

About the Martingale

Recall that $X_k = g(W_k)$, where $g \in L^2(\pi)$ and $\int_{\mathcal{W}} g d\pi = 0$, and let

$$h_\epsilon = \sum_{k=1}^{\infty} \frac{Q^{k-1} g}{(1+\epsilon)^k}$$

and $H_\epsilon(w_0, w_1) = h_\epsilon(w_1) - Qh_\epsilon(w_0)$. Then, from MW(2000),

$$H = \lim_{\epsilon \to 0} H_\epsilon$$

exists in an L^2 space, $M_n = H(W_0, W_1) + \cdots + H(W_{n-1}, W_n)$, and

$$R_n = \sum_{k=1}^{n} \xi_0 \circ \theta^k.$$

with $\xi_0 = g(W_0) - H(W_{-1}, W_0)$.

About the Proof

First Step: Show $\xi_0 \in [I - B(T)]L^2(P)$; that is $\xi_0 = \eta_0 - B(T)\eta_0$ for some $\eta_0 \in L^2(P)$.

- Bound $\|R_n\|$.
- Fourier anlysis of $B(e^{it})$.

Second Step: Show that

$$\lim_{n\to\infty} \frac{1}{\sqrt{n\ell^*(n)}} \sum_{k=1}^{n} \xi \circ \theta^k = 0 \ w.p.1$$

for any $\xi \in [I - B(T)]L^2(P)$. Uses a version of the Dominated Ergodic Theorem.

Some Details

First Step. Recall

$$A(z) = \frac{1}{1 - B(z)} = \sum_{k=0}^{\infty} \alpha_k z^k.$$

To make sense of $A(T)\xi_0$, we need to bound α_k. From $\beta_k \sim 2c/\sqrt{k^3\ell(k)}$, we get

$$B(e^{it}) \sim \kappa_0 \sqrt{\frac{t}{\ell(t)}}$$

as $t \to 0$. Thus

$$A(e^{it}) \sim \frac{\sqrt{\ell(t)}}{\kappa_0\sqrt{t}}$$

and

$$\alpha_k - \alpha_{k+1} = O[\sqrt{\frac{\ell(n)}{n^3}}]$$

Second Step. Let Y_k be stationary and

$$Y^* = \sup_{n\geq 1} |\frac{Y_1 + \cdots + Y_n}{n}|.$$

Then the Dominated Ergodic Theorem states:

$$\|Y^*\|_p \leq \left(\frac{p}{p-1}\right) \|Y_1\|_p$$

for $p > 1$. Also

$$E[\sqrt{Y^*}] \leq 2E[\sqrt{|Y_1|}].$$

Questions: The Iceberg

To what extent does the the theory of random walks extend to stationary, ergodic sequences?

- Law of Large Numbers: Yes
- Central Limit Theorem: $S_n = M_n + R_n$.
- Law of the Iterated Logarithm: $S_n = M_n + R_n$.
- When is $S_n = M_n + R_n$?
- Local Limit Theorems, etc. ...: Bits and pieces,
- Renewal Theorem: Partially in Lalley (1985, PTRF).
- Spitzer's Identity for Ladder Heights: Operator versions.