STAT 811 Probability Theory II

Chapter 7: Laws of Large Numers and Sums of Independent Random Variables

Dr. Dewei Wang
Associate Professor
Department of Statistics
University of South Carolina
deweiwang@stat.sc.edu

7.1 Truncation and Equivalence

From the definition of ui, we see that dealing with random variables that are uniformly bounded or that have moments are often easier. When these desirable properties are absent, we often use truncation to induce their presence and then see whether the truncation makes any equivalence. For example, we often compare

$$\{X_n\} \text{ with } \{X_n'\} = \{X_n I_{|X_n| \leq n}\}$$

where the second one is a truncated version of the first.

7.1 Truncation and Equivalence

Definition.

Two sequences $\{X_n\}$ and $\{X_n'\}$ are tail equivalent if

$$\sum_n P[X_n \neq X_n'] < \infty.$$

Remark: From the Borel-Cantelli Lemma, we have that the above tail equivalence implies

$$P\left(\left[X_n \neq X_n'\right] \text{ i.o. }\right) = 0$$

or equivalently

$$P\left(\liminf_{n\to\infty}\left[X_n = X_n'\right]\right) = 1$$

Let $N = \liminf_{n\to\infty}[X_n = X_n'] = \cup_n \cap_{k\geq n}[X_k = X_k']$, we have $P(N) = 1$. For $\omega \in N$, it means when $k \geq K(\omega)$, $X_k(\omega) = X_k'(\omega)$. Thus for $\omega \in N$, $\sum_n (X_n(\omega) - X_n'(\omega))$ converges; i.e.,

$$\sum_n (X_n - X_n') \text{ converges a.s.}$$

7.1 Truncation and Equivalence

In addition, we have $\sum_{n=K(\omega)}^{\infty} X_n(\omega) = \sum_{n=K(\omega)}^{\infty} X_n'(\omega)$. And if $a_n \uparrow \infty$ and when $n \geq K(\omega)$,

$$a_n^{-1} \sum_{j=1}^{n} (X_j(\omega) - X_j'(\omega)) = a_n^{-1} \sum_{j=1}^{K(\omega)} (X_j(\omega) - X_j'(\omega)) \to 0.$$

Proposition 7.1.1 (Equivalence)

Suppose the two sequences $\{X_n\}$ and $\{X_n'\}$ are tail equivalent. Then

1. $\sum_n (X_n - X_n')$ converges a.s.
2. $\sum_n X_n$ conveges a.s. iff $\sum_n X_n'$ converge a.s.
3. If $a_n \uparrow \infty$ and if there exits a random variable X such that

$$a_n^{-1} \sum_{j=1}^{n} X_j \stackrel{a.s.}{\to} X, \text{ then also } a_n^{-1} \sum_{j=1}^{n} X_j' \stackrel{a.s.}{\to} X.$$

Theorem 7.2.1 (General weak law of large numbers)

Suppose $\{X_n\}$ are independent random variables and define $S_n = \sum_{j=1}^{n} X_j$. If

1. $\sum_{j=1}^{n} P[|X_j| > n] \to 0$,
2. $n^{-2} \sum_{j=1}^{n} E\left(X_j^2 I_{[|X_j \le n]}\right) \to 0$,

then if we define

$$a_n = \sum_{j=1}^{n} E\left(X_j I_{[|X_j \le n]}\right),$$

we get

$$\frac{S_n - a_n}{n} = \bar{X}_n - a_n/n \xrightarrow{P} 0.$$

Remark: No assumptions about moments of X_n's need to be made.

Case (a) WLLN with variances.

Suppose $\{X_n\}$ are iid with $E(X_n) = \mu$ and $E(X_n^2) < \infty$. Then as $n \to \infty$, $S_n/n \xrightarrow{P} \mu$.

Proof.

Case (b) Khintchin's WLLN under the first moment hypothesis.

Suppose $\{X_n\}$ are iid with $E(|X_n|) < \infty$ and $E(X_n) = \mu$. Then as $n \to \infty$, $S_n/n \xrightarrow{P} \mu$.

Proof.

7.2 A General Weak Law of Large Numbers (Special Cases)

Case (c) Feller's WLLN without a first moment assumption.

Suppose $\{X_n\}$ are iid with $\lim_{x \to \infty} x P[|X_1| > x] = 0$, then
$S_n/n - E(X_1 I_{[|X_1| \leq n]}) \xrightarrow{P} 0$.

Proof.

7.2 A General Weak Law of Large Numbers (Special Cases)

Proof of Theorem 7.2.1.

7.2 A General Weak Law of Large Numbers (Proof)

Example

Let us consider a CDF: $F(x) = 1 - \frac{e}{2x \log x}, x \geq e$. Suppose we have from it an iid sequence $\{X_n\}$. What is the mean of X? What does S_n/n converges in probability to?

Solution.

7.2 A General Weak Law of Large Numbers (Example)

Example

How about (standard) Cauchy? We have

$$F(x) = 0.5 + \pi^{-1} \arctan x.$$

Solution.

7.3 Almost Sure Convergence of Sums of Independent Random Variables

Proposition 7.3.0 (Komogorov's inequality: about tail probabilities of maxima of sums)

Suppose $\{X_n\}$ is an independent sequence of random variables and suppose $E(X_n) = 0$ and $\text{var}(X_n) < \infty$. Then for each $\alpha > 0$,

$$P\left[\sup_{j \le N} |S_j| > \alpha\right] \le \frac{1}{\alpha^2}\text{var}(S_N) = \frac{1}{\alpha^2}\sum_{j=1}^{N} E(X_j^2).$$

Proof.

7.3 Almost Sure Convergence of Sums of Independent Random Variables

Proposition 7.3.1 (Skorohod's inequality: about tail probabilities of maxima of sums)

Suppose $\{X_n\}$ is an independent sequence of random variables and suppose $\alpha > 0$ is fixed. For $n \geq 1$, set

$$c = \sup_{j \leq N} P[|S_N - S_j| > \alpha] \overset{\text{if iid}}{=} \sup_{j \leq N} P[|S_j| > \alpha].$$

Suppose $c < 1$, then

$$P\left[\sup_{j \leq N} |S_j| > 2\alpha\right] \leq \frac{1}{1-c} P[|S_N| > \alpha].$$

Remark. Compared to the famous Kolmogorov's inequality, one advantage of Skorohod's inequality is that it does not require moment assumptions.

Proof of Skorohod's inequality.

7.3 Almost Sure Convergence of Sums of Independent Random Variables

7.3 Almost Sure Convergence of Sums of Independent Random Variables

Based on Skorohod's inequality, we may now present a remarkable result which shows the equivalence of convergence in probability to almost sure convergence for sums of independent random variables.

Theorem 7.3.2 (Lévy's theorem)

If $\{X_n\}$ are independent, then

$$\sum_n X_n \text{ converges i.p. iff } \sum_n X_n \text{ converges a.s.};$$

i.e., letting $S_n = \sum_{i=1}^n X_i$, then the following are equivalent

1. $\{S_n\}$ is Cauchy in probability.
2. $\{S_n\}$ converges in probability.
3. $\{S_n\}$ converges almost surely.
4. $\{S_n\}$ is almost surely Cauchy.

Proof of Lévy's theorem.

7.3 Almost Sure Convergence of Sums of Independent Random Variables

7.3 Almost Sure Convergence of Sums of Independent Random Variables

Theorem 7.3.3 (Kolmogorov convergence criterion)

If $\{X_n\}$ are independent, and

$$\sum_{j=1}^{\infty} \mathrm{var}(X_j) < \infty,$$

then

$$\sum_{j=1}^{\infty} (X_j - E(X_j)) \text{ converges almost surely.}$$

Proof.

This section considers the problem of when sums of independent random variables properly scaled and centered converge almost surely. We will prove that sample averages converge to mathematical expectations when the sequence is iid and a mean exists.

We begin with a number theory result which is traditionally used in the development of the theory.

Lemma 7.4.1 (Kronecker's lemma)

Suppose we have two sequences $\{x_k\}$ and $\{a_k\}$ such that $x_k \in \mathbb{R}$ and $0 < a_n \uparrow \infty$. If

$$\sum_k \frac{x_k}{a_k} \text{ converges,}$$

then

$$\lim_{n\to\infty} a_n^{-1} \sum_{k=1}^{n} x_k = 0.$$

Proof of Kronecker's lemma.

7.4 Strong Laws of Large Numbers

Corollary 7.4.1 (A strong law)

Let $\{X_n\}$ be independent with $E(X_n^2) < \infty$. Suppose we have a monotone sequence $b_n \uparrow \infty$. If

$$\sum_k \text{var}\left(\frac{X_k}{b_k}\right) < \infty,$$

then

$$\frac{S_n - E(S_n)}{b_n} \stackrel{a.s.}{\to} 0.$$

Proof.

7.4.1 Example 1: Record counts

Suppose $\{X_n\}$ is iid with continuous cdf F. Define

$$\mu_N = \sum_{j=1}^{n} I_{[X_j \text{ is a record}]} \doteq \sum_{j=1}^{N} 1_j,$$

so μ_N is the number of recordes in the first N observations.

Proposition 7.4.1 (Logarithmic growth rate)

The number of records in an iid sequence grows logarithmically and we have the almost sure limit

$$\lim_{N\to\infty} \frac{\mu_N}{\log N} \to 1.$$

Proof.

7.4.1 Example 1: Record counts

7.4.1 Example 1: Record counts

7.4.2 Example 2: Explosions in the Pure Birth Process.

Let $\{X_n\}$ be independent with

$$P[X_n > x] = e^{-\lambda_n x}, x > 0$$

where $\lambda_n \geq 0$ are birth parameters. Let $S_0 = 0$, $S_n = \sum_{i=1}^{n} X_i$ and the population size process $\{X(t), t \geq 0\}$ of the pure birth process by

$$X(t) = n \text{ if } S_{n-1} \leq t < S_n.$$

Next define the event explosion by

$$[\text{explosion}] = [\sum_{n=1}^{\infty} X_n < \infty] = [X(t) = \infty \text{ for some finite } t].$$

Proposition 7.4.2

$$P[\text{explosion}] = I(\sum_n \lambda_n^{-1} < \infty).$$

Proof.

7.4.2 Example 2: Explosions in the Pure Birth Process.

7.4.2 Example 2: Explosions in the Pure Birth Process.

Lemma 7.5.1

Let $\{X_n\}$ be iid. The following are equivalent:

(a) $E|X_1| < \infty$

(b) $\lim_{n\to\infty} \left|\frac{X_n}{n}\right| = 0$ almost surely.

(c) For $\epsilon > 0$, $\sum_{n=1}^{\infty} P[|X_1| \geq \epsilon n] < \infty$.

Proof.

7.5 The Strong Law of Large Numbers for IID Sequences

Kolmogorov's SLLN

Let $\{X_n\}$ be iid and set $S_n = \sum_{i=1}^{n} X_i$. There exists $c \in \mathbb{R}$ such that

$$\bar{X}_n = S_n/n \stackrel{a.s.}{\to} c$$

iff $E(|X_1) < \infty$ in which case $c = \mu = E(X_1)$.

Corollary 7.5.1

If $\{X_n\}$ is iid, then $E(|X_1|) < \infty$ implies $\bar{X}_n \stackrel{a.s.}{\to} \mu$ and $EX_1^2 < \infty$ implies $\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 \stackrel{a.s.}{\to} \sigma^2 = \text{var}(X_1)$ and $\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 \stackrel{a.s.}{\to} \sigma^2 = \text{var}(X_1)$.

Proof of Kolmogorov's SLLN.

7.5 The Strong Law of Large Numbers for IID Sequences

7.5 The Strong Law of Large Numbers for IID Sequences

7.5 The Strong Law of Large Numbers for IID Sequences

7.5.1 Application 1: Renewal Theory

Suppose X_n's are iid and non-negative. with $E(X_n) = \mu \in (0, \infty)$. Then

$$S_n/n \stackrel{a.s.}{\to} \mu, \text{ and } S_n \stackrel{a.s.}{\to} \infty.$$

Let $S_0 = 0$ and define

$$N(t) = \sum_{j=0}^{\infty} I_{[S_j \le t]}.$$

We call $N(t)$ the number of *renewals* in $[0, t]$. Then

$$[N(t) \le n] = [S_n > t] \text{ and } S_{N(t)-1} \le t < S_{N(t)}.$$

We conjecture that

$$N(t)/t \stackrel{a.s.}{\to} \mu^{-1}.$$

7.5.1 Application 1: Renewal Theory

7.5.1 Application 1: Glivenko-Cantelli Theorem

Suppose X_n's are iid with common distribution F. We estimate F by the *empirical cumulative distribution function* (ecdf)

$$\hat{F}_n(x) = n^{-1} \sum_{j=1}^{n} I_{[X_j \leq x]}.$$

By the SLLN, $\hat{F}_n(x) \stackrel{a.s.}{\to} F(x)$ for each fixed x.

Glivenko-Cantelli Theorem

$$D_n = \sup_x |\hat{F}_n(x) - F(x)| \stackrel{a.s.}{\to} 0.$$

Proof.

7.5.1 Application 1: Glivenko-Cantelli Theorem

7.5.1 Application 1: Glivenko-Cantelli Theorem

7.5.1 Application 1: Glivenko-Cantelli Theorem

7.5.1 Application 1: Glivenko-Cantelli Theorem

7.6 The Kolmogorov Three Series Theorem

The Kolmogorov three series theorem provides necessary and sufficient conditions for a series of independent random variables to converge. The result is especially useful when the Kolmogorov convergence criterion may not be applicable, for example, when existence of variances is not guaranteed.

Theorem 7.6.1

Let $\{X_n\}$ be independent. in order for $\sum_n X_n$ converges a.s., it is necessary and sufficient that there exists $c > 0$ such that

(i) $\sum_n P[|X_n| > c] < \infty$

(ii) $\sum_n \text{var}(X_n I_{[|X_n| \le c]}) < \infty$

(iii) $\sum_n E(X_n I_{|X_n| \le c})$ converges.

If $\sum_n X_n$ converges a.s., then (i), (ii), (iii) hold for any $c > 0$. Thus if the three series converge for one value of $c > 0$, they converge for all $c > 0$.

Proof of Sufficiency.

We now examine a proof of the necessity of the Kolmogorov three series theorem. Two lemmas pave the way. The first is a partial converse to the Kolmogorov convergence criterion.

Lemma 7.6.1
Suppose $\{X_n\}$ are independent which are uniformly bounded, so that for some $\alpha > 0$ and all $\omega \in \Omega$ we have $|X_n(\omega)| \leq \alpha$. If $\sum_n (X_n - EX_n)$ converges almost surely, then $\sum_n \text{var}(X_n) < \infty$.

Proof.

7.6 The Kolmogorov Three Series Theorem

7.6 The Kolmogorov Three Series Theorem

7.6 The Kolmogorov Three Series Theorem

Lemma 7.6.2

Suppose $\{X_n\}$ are independent which are uniformly bounded, so that for some $\alpha > 0$ and all $\omega \in \Omega$ we have $|X_n(\omega)| \leq \alpha$. Then $\sum_n X_n$ converges almost surely implies that $\sum_n EX_n$ converges.

Proof.

7.6 The Kolmogorov Three Series Theorem

We now turn to the necessity of the Kolmogorov three series theorem.

Proof of necessity.