Lecture 3: Independence and Zero-One law

Tianpei Xie

Jul. 17th., 2015

Contents

1 Independence

1.1 Basic Definitions

- **Definition** (***Independence for Two Events***)
  Suppose $(\Omega, \mathscr{F}, \mathcal{P})$ is a fixed *probability space.* ***Events*** A, $B \in \mathscr{F}$ are ***independent*** if
  $$\mathcal{P}(A \cap B) = \mathcal{P}(A)\, \mathcal{P}(B).$$

- **Definition** (***Independence of a Finite Number of Events***)
  The events $A_1\,, \ldots, A_n$ are ***independent*** if
  $$\mathcal{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathcal{P}(A_i), \qquad \textbf{\textit{for all finite}}\ I \subseteq \{1\,, \ldots, n\}\,.$$

- **Remark** In order for n events to be independent, we need
  $$\sum_{k=2}^{n} \binom{n}{k} = 2^n - n - 1$$
  equations.

- **Remark** (***Alternative Definitions***)
  The events $A_1\,, \ldots, A_n$ are ***independent*** if
  $$\mathcal{P}\left(\bigcap_{i=1}^{n} B_i\right) = \prod_{i=1}^{n} \mathcal{P}(B_i)$$
  where for each $i = 1, \,, \ldots, n$,
  $$B_i = A \text{ or } \Omega.$$

- **Definition** (***Independent Classes***)
  Let $\mathscr{C}_i \subseteq \mathscr{F}$, $i = 1\,, \ldots, n$. The classes $\mathscr{C}_i$ are ***independent***, if for *any choice* $A_1\,, \ldots, A_n$, with $A_i \in \mathscr{C}_i$, $i = 1\,, \ldots, n$, we have the events $A_1\,, \ldots, A_n$ *independent events.*

- **Proposition 1.1** *(**Basic Criterion**) [Resnick, 2013]*
  If for each $i = 1\,, \ldots, n$, $\mathscr{C}_i$ is a non-empty class o fevents satisfying

  1. *$\mathscr{C}_i$ is a π-**system** (closure under finite intersection),*
  2. *$\mathscr{C}_i, i = 1\,, \ldots, n$ are **independent**,*

  then
  $$\sigma(\mathscr{C}_1)\,, \ldots, \sigma(\mathscr{C}_n)$$
  *are **independent**.*

- **Definition** (***Arbitrary Number of Independent Classes***)
  Let T be an arbitrary index set. The classes $\mathscr{C}_t, t \in T$ are ***independent families*** if for *each finite* I, $I \subset T$, $\{\mathscr{C}_t, t \in I\}$ is independent.

- **Corollary 1.2** *[Resnick, 2013]*
  If $\mathscr{C}_t, t \in T$ are non-empty π-systems that are ***independent****, then $\{\sigma(\mathscr{C}_t),\, t \in T\}$ are* ***independent****.*

1.2 Independent Random Variables

- **Definition** (***Independent Random Variables***)
  $\{X_t, t \in T\}$ is an ***independent family of random variables*** if $\{\sigma(X_t),\, t \in T\}$ are ***independent*** σ-algebras.

- **Remark** (***Indicator Random Variables***)
  Note that $\sigma(\mathbb{1}_A) = \{\emptyset, \Omega, A, A^c\}$.

$$\begin{aligned} &\{\mathbb{1}_{A_t}, t \in T\} \text{ are independent random variables} \\ \Leftrightarrow &\{A_t, t \in T\} \text{ are independent} \end{aligned}$$

- **Definition** (***Finite Dimensional Distribution Functions***)
  For a family of random variables $\{X_t, t \in T\}$ indexed by a set T, the ***finite dimensional distribution functions*** are *the family of* ***multivariate distribution functions***

$$F_J\left(x_t, t \in T\right) = \mathcal{P}\left[X_t \le x_t, \forall t \in J\right]$$

  for all *finite subsets* $J \subset T$.

- **Proposition 1.3** ***(Factorization Criterion)*** *[Resnick, 2013]*
  A family of random variables $\{X_t, t \in T\}$ indexed by a set T, is ***independent*** *if and only if for* ***all finite*** *$J \subset T$*

$$F_J\left(x_t, t \in T\right) = \prod_{t \in J} \mathcal{P}[X_t \le x_t], \quad \forall x_t \in \mathbb{R}.$$

- **Remark** A family of random variables $\{X_t, t \in T\}$ indexed by a set T above may contain *infinite number of random variables.*

- **Corollary 1.4** ***(Finite Dimensional Case)*** *[Resnick, 2013]*
  The finite collection of random variables $X_1, \ldots, X_k$ is ***independent*** *if and only if*

$$\mathcal{P}\left[X_1 \le x_1, \ldots, X_k \le x_k\right] = \prod_{i=1}^{k} \mathcal{P}[X_i \le x_i], \quad \forall x_i \in \mathbb{R}.$$

- **Corollary 1.5** ***(Finite Dimensional Discrete Case)*** *[Resnick, 2013]*
  The ***discrete*** *random variables $X_1, \ldots, X_k$ with* ***countable*** *range $\mathcal{R}$ are* ***independent*** *if and only if*

$$\mathcal{P}\left[X_i = x_i, i = 1, \ldots, k\right] = \prod_{i=1}^{k} \mathcal{P}[X_i = x_i], \quad \forall x_i \in \mathcal{R}.$$

1.3 Examples of Independence

1.4 Groupings

-

2 Independence, Zero-One Laws, Borel-Cantelli Lemma

- **Remark** There are several common zero-one laws which identify the possible range of a random variable to be trivial. There are also several zero-one laws which provide the basis for all proofs of *almost sure convergence.*

- **Remark** Note that ***almost surely convergence*** is the *pointwise convergence* ***outside a null set***. That is, ***the asymptotic behavior*** *is the same* for ***every possible outcome*** *besides* those with zero measure.

  1. ***The Borel-Cantelli Lemma*** provides a basic criterion for *almost sure convergence*, i.e. the ***total sum of probabilities*** *for* ***all events*** *is* ***convergent***

     $$\sum_{i=1}^{\infty} \mathcal{P}\left(A_i\right) < \infty.$$

     This condition *guarentees that* ***the measure of tail support converges to zero***. The drawback is that it only provides a ***sufficient condition*** for *the almost sure convergence*. In other word, it says that if the total proabilities of all event is ***unbounded***, then we ***cannot*** say we would not have almost sure convergence.

  2. With ***the independence assumption***, we have an <u>***almost deterministic criterion***</u> on whether or not *asymptotic events* will happen.

     (a) ***The Borel Zero-One Law*** directly comes from *the Borel-Cantelli Lemma*, which asserts that ***with independence assumption***, *the* ***convergence of total probabilities*** is *an almost deterministic criterion* for the almost sure convergence.

     (b) ***The Komogorov Zero-One Law*** even claims that ***all tail events*** follow the same *zero-one law*, i.e. it will *either* ***happen almost surely*** or ***not happen almost surely***.

- **Remark** (***Zero-One Law = Almost Deteriminstic Test on Asymptotic of Independent Variables***)
  The Komogorov zero-one Law provides a siginificant insight on *the* ***test*** *of* ***asymptotic behavior*** *of* ***indenpendent*** *random variables.* Note that *all* ***asymptotic statistics*** *are* ***tail random variables***, i.e. it relies on ***information collected in the future***.

  And the conclusion of the zero-one law is that <u>*the* ***test*** *on the asymptotic statistics*</u> will have a <u>***deterministic answer***</u> (1 or 0) for ***every possible outcome*** of the experiment ***excepts*** for outcomes with *zero probability.*

2.1 Borel-Cantelli Lemma

- **Theorem 2.1** *(**Borel-Cantelli Lemma**). [Resnick, 2013]*
  Let $\{A_n\}$ be any events. If

  $$\sum_n \mathcal{P}(A_n) < \infty,$$

  then

  $$\mathcal{P}\left\{\limsup_{n\to\infty} A_n\right\} \equiv \mathcal{P}\left([A_n \;\; i.o.]\right) = 0$$

  where i.o is infinitely often.

  Proof: We know that

  $$\begin{aligned}
  \mathcal{P}\left\{\limsup_{n\to\infty} A_n\right\} &= \mathcal{P}\left\{\bigcap_{k\ge 1}\bigcup_{n\ge k} A_n\right\} \\
  &= \lim_{k\to\infty} \mathcal{P}\left\{\bigcup_{n\ge k} A_n\right\} \quad \text{(by downward convergence)} \\
  &= \lim_{k\to\infty} \sum_{n=k}^{\infty} \mathcal{P}\left\{A_n\right\} \\
  &\le \limsup_{k\to\infty} \sum_{n=k}^{\infty} \mathcal{P}\left\{A_n\right\} = 0,
  \end{aligned}$$

  since $\sum_n \mathcal{P}(A_n) < \infty$ implies that $\sum_{n=k}^{\infty} \mathcal{P}\left\{A_n\right\} \to 0$ as $k \to \infty$. ∎

- **Remark** The *Borel-Cantelli Lemma* does not require any independence between events. It states that *almost every outcome ω in Ω is contained* ***at most finitely many*** *of events A_n* or equivalently, $\{n : \omega \in A_n\}$ is finite for every ω.

- **Remark** The *Borel-Cantelli Lemma* is used as the basis for all proofs of ***almost sure convergence***.

2.2 Borel Zero-One Law

- The *Borel-Cantelli Lemma* does not require ***independence***. The next result does.

  Theorem 2.2 *(**Borel Zero-One Law**) [Resnick, 2013]*
  If $\{A_n\}$ is a sequence of ***independent*** *events, then*

  $$\mathcal{P}\left\{\limsup_{n\to\infty} A_n\right\} = \mathcal{P}\left([A_n \;\; i.o.]\right) = \begin{cases} 0 & \text{if } \sum_n \mathcal{P}(A_n) < \infty \\ 1 & \text{if } \sum_n \mathcal{P}(A_n) = \infty \end{cases}$$

  Proof: From the *Borel-Cantelli lemma*, we see that if $\sum_n \mathcal{P}(A_n) < \infty$, $\mathcal{P}\left\{A_n \text{ i.o.}\right\} = 0$. For

$\sum_n \mathcal{P}(A_n) = \infty$, we see that

$$
\begin{aligned}
\mathcal{P}\left\{\limsup_{n\to\infty} A_n\right\} &= \mathcal{P}\left\{\bigcap_{k\geq 1}\bigcup_{n\geq k} A_n\right\} \\
&= 1 - \mathcal{P}\left\{\bigcup_{k\geq 1}\bigcap_{n\geq k} A_n^c\right\} \\
&= 1 - \lim_{k\to\infty} \mathcal{P}\left\{\bigcap_{n\geq k} A_n^c\right\} \quad \text{(by upward convergence)} \\
&= 1 - \lim_{k\to\infty} \mathcal{P}\left\{\lim_{m\to\infty} \downarrow \bigcap_{n=k}^{m} A_n^c\right\} \\
&= 1 - \lim_{k\to\infty}\lim_{m\to\infty} \mathcal{P}\left\{\bigcap_{n=k}^{m} A_n^c\right\} \quad \text{(by downward convergence)} \\
&= 1 - \lim_{k\to\infty}\lim_{m\to\infty} \left(\prod_{n=k}^{m} \mathcal{P}(A_n^c)\right) \\
&= 1 - \lim_{k\to\infty}\lim_{m\to\infty} \prod_{n=k}^{m} \left(1 - \mathcal{P}(A_n)\right).
\end{aligned}
$$

Thus it suffice to show that

$$
\lim_{k\to\infty}\lim_{m\to\infty} \prod_{n=k}^{m} \left(1 - \mathcal{P}(A_n)\right) = 0.
$$

We use the inequality

$$
1 - x \leq \exp(-x), \quad x \in (0,1)
$$

thus

$$
\begin{aligned}
\lim_{m\to\infty} \prod_{n=k}^{m} \left(1 - \mathcal{P}(A_n)\right) &\leq \lim_{m\to\infty} \prod_{n=k}^{m} \exp\left(-\mathcal{P}(A_n)\right) \\
&= \lim_{m\to\infty} \exp\left(-\sum_{n=k}^{m} \mathcal{P}(A_n)\right) \\
&= \exp(-\infty) \\
&= 0, \text{ for all } n \leq m,
\end{aligned}
$$

since $\sum_n \mathcal{P}(A_n) = \infty$. Therefore $\lim_{k\to\infty}\lim_{m\to\infty} \prod_{n=k}^{m} \left(1 - \mathcal{P}(A_n)\right) = 0$. ∎

- **Example** (***Behavior of Exponential Random Variables***) [Resnick, 2013]
We assume that $\{E_n, n \geq 1\}$ are *i.i.d.* ***unit exponential variables***; that is,

$$
\mathcal{P}\left[E_n > x\right] = e^{-x}, \quad x > 0.
$$

Then

$$
\mathcal{P}\left\{\limsup_{n\to\infty} \frac{E_n}{\log n} = 1\right\} = 1
$$

Proof: For any ω,

$$\limsup_{n\to\infty} \frac{E_n(\omega)}{\log n} = 1$$

means that $\forall \epsilon > 0$, there exists n such that

$$\frac{E_n(\omega)}{\log n} > 1 - \epsilon$$

and also for large n,

$$\frac{E_n(\omega)}{\log n} \leq 1 + \epsilon.$$

Therefore

$$\begin{aligned}
\left\{\limsup_{n\to\infty} \frac{E_n}{\log n} = 1\right\} &= \bigcap_s \left[\bigcup_{n\geq 1} \left\{\frac{E_n(\omega)}{\log n} > 1 - \epsilon_s\right\} \bigcap \bigcup_{k\geq 1} \bigcap_{n\geq k} \left\{\frac{E_n(\omega)}{\log n} \leq 1 + \epsilon_s\right\}\right] \\
&= \bigcap_s \left\{\left[\frac{E_n(\omega)}{\log n} > 1 - \epsilon_s\right], i.o.\right\} \bigcap \bigcap_s \left\{\liminf_{n\to\infty} \left[\frac{E_n(\omega)}{\log n} \leq 1 + \epsilon_s\right]\right\}
\end{aligned}$$

To show the RHS has probability 1, it then suffice to show that each bracket event occurs with probability 1.

For the event $\left\{\left[\frac{E_n(\omega)}{\log n} > 1 - \epsilon_s\right], n \geq 1\right\}$,

$$\sum_{n=1}^{\infty} \mathcal{P}\left\{\frac{E_n(\omega)}{\log n} > 1 - \epsilon_s\right\} = \sum_{n=1}^{\infty} \exp\left((1-\epsilon_s)\log n\right) = \sum_{n=1}^{\infty} \frac{1}{n^{(-1+\epsilon_s)}} = \infty,$$

so $\left\{\left[\frac{E_n(\omega)}{\log n} > 1 - \epsilon_s\right], i.o.\right\}$ occurs with probability 1 for all s.

$$\mathcal{P}\left\{\liminf_{n\to\infty} \left[\frac{E_n(\omega)}{\log n} \leq 1 + \epsilon_s\right]\right\} = 1 - \mathcal{P}\left\{\limsup_{n\to\infty} \left[\frac{E_n(\omega)}{\log n} \geq -1 - \epsilon_s\right]\right\}$$

Since

$$\begin{aligned}
\sum_{n=1}^{\infty} \mathcal{P}\left\{\frac{E_n(\omega)}{\log n} \geq -1 - \epsilon_s\right\} &= \sum_{n=1}^{\infty} \exp\left((-1-\epsilon_s)\log n\right) \\
&= \sum_{n=1}^{\infty} \frac{1}{n^{(1+\epsilon_s)}} < \infty,
\end{aligned}$$

we have

$$\begin{aligned}
\mathcal{P}\left\{\liminf_{n\to\infty} \left[\frac{E_n(\omega)}{\log n} \leq 1 + \epsilon_s\right]\right\} &= 1 - \mathcal{P}\left\{\limsup_{n\to\infty} \left[\frac{E_n(\omega)}{\log n} \geq -1 - \epsilon_s\right]\right\} \\
&= 1 - 0 = 1. \quad \blacksquare
\end{aligned}$$

- **Remark** (***Heavy Tail***)
  This result is sometimes considered ***surprising***. There is a (*mistaken*) tendency to think of i.i.d sequences as somehow roughly constant, and therefore the division by $\log n$ should send the ratio to 0.

  However, *every so often*, the sequence $\{E_n\}$ ***spits out a large value*** and the ***growth*** of these ***large values*** *approximately matches* that of $\{\log n, n \geq 1\}$.

- **Example** (***Behavior of Normal Random Variables***) [Resnick, 2013]
  We assume that $\{X_n, n \geq 1\}$ are *i.i.d.* ***standard normal variables*** $\mathcal{N}(0, 1)$. Then

$$\mathcal{P}\left\{\limsup_{n\to\infty} \frac{|X_n|}{\sqrt{\log n}} = \sqrt{2}\right\} = 1.$$

  Use the fact that

$$\lim_{x\to\infty} \frac{\mathcal{P}\left[X_n \geq x\right]}{n(x)/x} = 1$$

  where $n(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2)$ is standard normal density.

  Proof:

$$\left\{\limsup_{n\to\infty} \frac{|X_n|}{\sqrt{\log n}} = \sqrt{2}\right\} = \bigcap_{s}\left[\bigcup_{n\geq 1}\left\{\frac{|X_n|}{\sqrt{\log n}} > \sqrt{2} - \epsilon_s\right\}\bigcap \bigcup_{k\geq 1}\bigcap_{n\geq k}\left\{\frac{|X_n|}{\sqrt{\log n}} \leq \sqrt{2} + \epsilon_s\right\}\right]$$
$$= \bigcap_{s}\left\{\left[\frac{|X_n|}{\sqrt{\log n}} > \sqrt{2} - \epsilon_s\right], i.o.\right\}\bigcap\bigcap_{s}\left\{\liminf_{n\to\infty}\left[\frac{|X_n|}{\sqrt{\log n}} \leq \sqrt{2} + \epsilon_s\right]\right\}.$$

  1. Show that

$$\mathcal{P}\left\{\left[\frac{|X_n|}{\sqrt{\log n}} > \sqrt{2} - \epsilon_s\right], i.o.\right\} = 1 \quad \forall \epsilon_s > 0.$$

     From

$$\lim_{x\to\infty} \frac{\mathcal{P}\left[X_n \geq x\right]}{n(x)/x} = 1,$$

     we know that for any $\epsilon_s > 0$, for large x

$$\left|\frac{\mathcal{P}\left[X_n \geq x\right]}{n(x)/x} - 1\right| < \epsilon_s$$
$$\left|x\mathcal{P}\left[X_n \geq x\right] - \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)\right| < \epsilon_s$$
$$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right) - \epsilon_s < |x\mathcal{P}\left[X_n \geq x\right]| < \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right) + \epsilon_s$$

Thus

$$
\begin{aligned}
\left|\left(\sqrt{\log n}(\sqrt{2}-\epsilon_s\right)\mathcal{P}\left[X_n \geq \sqrt{\log n}(\sqrt{2}-\epsilon_s')\right]\right| &< \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\sqrt{\log n}(\sqrt{2}-\epsilon_s'\right)^2\right)+\epsilon_s \\
&= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left((\sqrt{2}-\epsilon_s'\right)^2\log n\right)+\epsilon_s \\
&= \frac{1}{\sqrt{2\pi}}\frac{1}{n^{c_{\epsilon'}}}+\epsilon_s \\
&\text{where } c_{\epsilon'} = \left(1-\frac{\epsilon_s'}{\sqrt{2}}\right)^2 < 1
\end{aligned}
$$

$$
\begin{aligned}
&\text{for } X_n \geq 0 \\
\mathcal{P}\left[X_n \geq \sqrt{\log n}(\sqrt{2}-\epsilon_s')\right] &\leq \frac{1}{2\sqrt{\pi}}\frac{1}{n^{c_{\epsilon'}}\sqrt{\log n}}+\frac{B\epsilon_s}{\sqrt{2\log n}} \quad 0 < c_{\epsilon'} < 1 \\
\text{for } X_n \leq 0,\ \text{let } X_n' = -X_n \sim \mathcal{N}(0,1) \\
\mathcal{P}\left[X_n' \geq \sqrt{\log n}(\sqrt{2}-\epsilon_s')\right] &\leq C_{\epsilon_s'}\frac{1}{n^{c_{\epsilon'}}\sqrt{\log n}}+\epsilon_s \\
\mathcal{P}\left[X_n \leq \sqrt{\log n}(-\sqrt{2}+\epsilon_s')\right] &\leq C_{\epsilon_s'}\frac{1}{n^{c_{\epsilon'}}\sqrt{\log n}}+\epsilon_s \quad 0 < c_{\epsilon'} < 1
\end{aligned}
$$

Similarly

$$
\begin{aligned}
\mathcal{P}\left[X_n \geq \sqrt{\log n}(\sqrt{2}-\epsilon_s')\right] &\geq C_{\epsilon_s'}\frac{1}{n^{c_{\epsilon'}}\sqrt{\log n}}-\epsilon_s \quad \text{for } X_n \geq 0 \\
\mathcal{P}\left[X_n \leq \sqrt{\log n}(-\sqrt{2}+\epsilon_s')\right] &\geq C_{\epsilon_s'}\frac{1}{n^{c_{\epsilon'}}\sqrt{\log n}}-\epsilon_s \quad \text{for } X_n \leq 0
\end{aligned}
$$

Therefore consider

$$
\begin{aligned}
\mathcal{P}\left\{\frac{|X_n|}{\sqrt{\log n}} > \sqrt{2}-\epsilon_s'\right\} &= \mathcal{P}\left\{\left(\frac{X_n}{\sqrt{\log n}} > \sqrt{2}-\epsilon_s'\right)\cup\left(\frac{X_n}{\sqrt{\log n}} < -\sqrt{2}+\epsilon_s'\right)\right\} \\
&\geq C_{\epsilon_s'}\frac{1}{n^{c_{\epsilon'}}\sqrt{\log n}}-\epsilon_s \quad \text{(by monotonicity)} \\
\sum_{n=1}^{\infty}\mathcal{P}\left\{\frac{|X_n|}{\sqrt{\log n}} > \sqrt{2}-\epsilon_s'\right\} &= \infty
\end{aligned}
$$

as $\sum_{n=1}^{\infty}\frac{1}{n^{c_{\epsilon'}}\sqrt{\log n}}$ diverges for $0 < c_{\epsilon'} < 1$ and we see that by Borel-Cantelli lemma

$$
\mathcal{P}\left\{\left[\frac{|X_n|}{\sqrt{\log n}} > \sqrt{2}-\epsilon_s\right], i.o.\right\} = 1 \quad \forall \epsilon_s > 0.
$$

2. Show that

$$
\begin{aligned}
&\mathcal{P}\left\{\liminf_{n\to\infty}\left[\frac{|X_n|}{\sqrt{\log n}} \leq \sqrt{2}+\epsilon_s\right]\right\} = 1 \\
\Leftrightarrow\ &\mathcal{P}\left\{\limsup_{n\to\infty}\left[\frac{|X_n|}{\sqrt{\log n}} \geq -\sqrt{2}-\epsilon_s\right]\right\} = 0
\end{aligned}
$$

See that

$$\mathcal{P}\left\{\left[\frac{|X_n|}{\sqrt{\log n}} \geq -\sqrt{2}-\epsilon_s'\right]\right\} \leq \mathcal{P}\left\{\left(\frac{X_n}{\sqrt{\log n}} \geq -\sqrt{2}-\epsilon_s'\right)\right\} + \mathcal{P}\left\{\left(\frac{X_n}{\sqrt{\log n}} \leq \sqrt{2}+\epsilon_s'\right)\right\}$$

We consider

$$\mathcal{P}\left\{X_n \geq (\sqrt{2}+\epsilon_s')\sqrt{\log n}\right\} \geq C_{\epsilon_s'}\frac{1}{n^{c'_{\epsilon'}}\sqrt{\log n}} - \epsilon_s, \text{where } c'_{\epsilon'} = \left(1+\frac{\epsilon'}{\sqrt{2}}\right)^2 \geq 1$$

so

$$\begin{aligned}
\sum_{n=1}^{\infty}\mathcal{P}\left\{\left(\frac{X_n}{\sqrt{\log n}} \geq -\sqrt{2}-\epsilon_s'\right)\right\} &= \mathcal{P}\left\{X_n \geq -(\sqrt{2}+\epsilon_s')\sqrt{\log n}\right\} \\
&= 1-\sum_{n=1}^{\infty}\mathcal{P}\left\{X_n \geq (\sqrt{2}+\epsilon_s')\sqrt{\log n}\right\} \quad \text{(by symmetry of } \mathcal{N}(0,1)) \\
&\leq 1-\sum_{n=1}^{\infty}\frac{1}{n^{c'_{\epsilon'}}\sqrt{\log n}} < \infty
\end{aligned}$$

since $\sum_{n=1}^{\infty}\frac{1}{n^{c'_{\epsilon'}}\sqrt{\log n}} < \infty$ for $c'_{\epsilon'} \geq 1$. Similarly,

$$\sum_{n=1}^{\infty}\mathcal{P}\left\{\left(\frac{X_n}{\sqrt{\log n}} \leq \sqrt{2}+\epsilon_s'\right)\right\} < \infty$$

By Borel-Cantelli lemma,

$$\begin{aligned}
\mathcal{P}\left\{\limsup_{n\to\infty}\left[\frac{|X_n|}{\sqrt{\log n}} \geq -\sqrt{2}-\epsilon_s\right]\right\} &= 0, \\
\Rightarrow \quad \mathcal{P}\left\{\liminf_{n\to\infty}\left[\frac{|X_n|}{\sqrt{\log n}} \leq \sqrt{2}+\epsilon_s\right]\right\} &= 1,
\end{aligned}$$

which completes our proof. ∎

2.3 Tail σ-Algebra and Komogorov Zero-One Law

- **Definition** (***Tail σ-Algebra***)
  Let $\{X_n\}$ be a sequence of random variables and define

  $$\mathscr{F}_n' = \sigma\left(X_{n+1},\cdots\right), \quad n \geq 1,$$

  which is the smallest σ-algebra containing random variables X_k for $k > n$. Define the ***tail σ-algebra*** as

  $$\mathscr{T} = \bigcap_{n\geq 1}\mathscr{F}_n' = \lim_{n\to\infty}\downarrow\sigma\left(X_n, X_{n+1},\cdots\right).$$

  These are events which depend on *the* ***tail*** *of the* $\{X_n\}$ *sequence*. If $A \in \mathscr{T}$, we will call A a ***tail event***.

  A ***random variable measurable*** with respect to $\mathscr{T}$ is called a ***tail random variable***.

- **Example** (***Examples of Tail Events***)

  1. The event

     $$\left\{\omega : \sum_{n=1}^{\infty} X_n(\omega) \text{ converges}\right\} \in \mathscr{T}$$

     To see this note that, for any m, the sum $\sum_{n=1}^{\infty} X_n(\omega)$ *converges* if and only if $\sum_{n=m}^{\infty} X_n(\omega)$ *converges*. So

     $$\left\{\omega : \sum_{n=1}^{\infty} X_n(\omega) \text{ converges}\right\} = \left\{\omega : \sum_{n=m+1}^{\infty} X_n(\omega) \text{ converges}\right\} \in \mathscr{F}'_m$$

     This holds for all m and after *intersecting* over m. $\blacksquare$

  2. The event

     $$\left\{\omega : \lim_{n\to\infty} X_n(\omega) \text{ exists}\right\} \in \mathscr{T}$$

     Note that both $\limsup\limits_{n\to\infty} X_n$ and $\liminf\limits_{n\to\infty} X_n$ are the same as $\lim\limits_{m\to\infty} \sup_{n\geq m} X_n$ and $\lim\limits_{m\to\infty} \inf_{n\geq m} X_n$ $\blacksquare$.

  3. Let $S_n = \sum_{i=1}^{n} X_i$, the event

     $$\left\{\omega : \lim_{n\to\infty} \frac{S_n(\omega)}{n} = 0\right\} = \left\{\omega : \lim_{n\to\infty} \frac{\sum_{k=1}^{n} X_k(\omega)}{n} = 0\right\} \in \mathscr{T},$$

     This is because for any m,

     $$\lim_{n\to\infty} \frac{S_n(\omega)}{n} = \lim_{n\to\infty} \frac{\sum_{k=1}^{n} X_k(\omega)}{n} = \lim_{n\to\infty} \frac{\sum_{k=m+1}^{n} X_k(\omega)}{n}$$

     and so for any m,

     $$\lim_{n\to\infty} \frac{S_n(\omega)}{n} \text{ is } \mathscr{F}'_m \text{ measurable.}$$

- **Example** (***Examples of Tail Random Variables***)

  1. $\sum_{i=1}^{\infty} X_i$ is a ***tail random variable***.

  2. $\limsup\limits_{n\to\infty} X_n$ and $\liminf\limits_{n\to\infty} X_n$ are both ***tail random variables***.

     This is true since the $\limsup$ of the sequence $\{X_1, X_2, \ldots\}$ is the same as the lim sup of the sequence $\{X_m, X_{m+1}, \ldots\}$ for all m. $\blacksquare$

  3. Let $S_n = \sum_{i=1}^{n} X_i$, then

     $$\lim_{n\to\infty} \frac{S_n}{n} = \lim_{n\to\infty} \frac{\sum_{k=1}^{n} X_k}{n}$$

     is a ***tail random variable***.

- **Definition** (***Almost Trivial σ-Algebra***)
  For a σ-algebra $\mathscr{F}$, if

$$\mathcal{P}(A) \in \{0, 1\}, \quad \forall A \in \mathscr{F}$$

  then $\mathscr{F}$ is ***almost trivial***.

- **Remark** A trivial σ-algebra $\mathscr{F} = \{\emptyset, \Omega\}$ is almost trivial, of course.

  The Komogorov Zero-One Law below confirms that ***the tail σ-algebra $\mathscr{T}$ for a set of independent random variables is almost trivial***. This provide the basis for all proofs of ***almost sure convergence*** under the independence assumption.

- **Lemma 2.3** *(**Almost Trivial σ-Algebra**) [Resnick, 2013]*
  *Let $\mathcal{G}$ be an **almost trivial** σ-algebra and let X be a random variable measureable w.r.t. $\mathcal{G}$. Then there exists c such that $\mathcal{P}\{X = c\} = 1$.*

  Proof: Let $F(x) = \mathcal{P}\{X \le x\}$. Then F is non-decreasing and since $\{\omega : X(\omega) \le x\} \in \sigma(X) \subset \mathcal{G}$, $F(x) = 0$ or $F(x) = 1$ for any $x \in \mathbb{R}$.

  Let $c = \sup\{x : F(x) = 0\}$. The distribution function must have a jump of size 1 at $x = c$ and thus $\mathcal{P}\{X = c\} = 1$. $\blacksquare$

- **Theorem 2.4** *(**Komogorov Zero-One Law**) [Resnick, 2013]*
  *If $\{X_n\}$ are **independent** random variables with **tail σ-algebra** $\mathscr{T}$, then $\Lambda \in \mathscr{T}$ implies $\mathcal{P}(\Lambda) = 0$ or $\mathcal{P}(\Lambda) = 1$ so that **σ-algebra $\mathscr{T}$ is almost trivial**.*

  Proof: Suppose $\Lambda \in \mathscr{T}$. It suffice to show that Λ ***is independent to itself*** so that $\mathcal{P}(\Lambda) = \mathcal{P}(\Lambda \cap \Lambda) = \mathcal{P}(\Lambda)^2$. It only occurs if $\mathcal{P}(\Lambda) = 0$ or 1.

  See that

$$\mathscr{F}_n = \sigma(X_1, \cdots, X_n) = \bigvee_{k=1}^{n} \sigma(X_k), n \ge 1.$$

  is the smallest σ-algebra contains all $\sigma(X_k), 1 \le k \le n$. Here $\mathcal{C} = \{A, B\}$ and $\mathcal{D} = \{C, D\}$, then $\mathcal{C} \vee \mathcal{D} = \{A \cap C, A \cap D, B \cap C, B \cap D\}$ is union of collection via elementwise intersection. Therefore, $\mathscr{F}_n \subset \mathscr{F}_{n+1}$ with

$$\mathscr{F}_\infty = \sigma(X_1, X_2, \ldots) = \bigvee_{n=1}^{\infty} \mathscr{F}_n = \bigvee_{n=1}^{\infty} \sigma(X_n).$$

  Note that

$$\Lambda \in \mathscr{T} \subset \mathscr{F}_n' = \sigma(X_{n+1}, \cdots) \subset \mathscr{F}_\infty = \sigma(X_1, X_2, \ldots).$$

  So since for all n, $\Lambda \in \mathscr{F}_n'$ and $\mathscr{F}_n' \perp\!\!\!\perp \mathscr{F}_n$, we have

$$\Lambda \perp\!\!\!\perp \mathscr{F}_n, \text{ for all } n.$$

  Hence

$$\Lambda \perp\!\!\!\perp \bigcup_{n=1}^{\infty} \mathscr{F}_n.$$

Let $\mathscr{C}_1 = \{\Lambda\}$, and $\mathscr{C}_2 = \bigcup_n \mathscr{F}_n$. Then $\mathscr{C}_i$ is a π-system, $i = 1, 2$, $\mathscr{C}_1 \perp\!\!\!\perp \mathscr{C}_2$ and therefore *the Basic Criterion* implies

$$\sigma(\{\Lambda\}) = \{\emptyset, \Omega, \Lambda, \Lambda^c\} \quad \text{and} \quad \sigma\left(\bigcup_{n=1}^{\infty} \mathscr{F}_n\right) = \bigvee_{n=1}^{\infty} \mathscr{F}_n = \mathscr{F}_\infty \quad \text{are } \textbf{\textit{independent}}.$$

Now $\Lambda \in \sigma(\mathscr{C}_1)$ and $\Lambda \in \mathscr{F}_\infty = \sigma(\mathscr{C}_2)$ therefore Λ is independent to itself. ∎

- **Corollary 2.5** *(**Corollaries of the Kolmogorov Zero-One Laws**) [Resnick, 2013]*
  *Let $\{X_n\}$ be **independent** random variables. Then*

  1. *The event*

  $$\left\{\omega : \sum_{n=1}^{\infty} X_n(\omega) \text{ converges}\right\} \in \mathscr{T}$$

  has probability 0 *or* 1.

  2. *The **tail** random variables* $\limsup_{n\to\infty} X_n$ *and* $\liminf_{n\to\infty} X_n$ *are **constant** with probability* 0 *or* 1.

  3. *The event*

  $$\left\{\omega : \lim_{n\to\infty} \frac{S_n(\omega)}{n} = 0\right\} = \left\{\omega : \lim_{n\to\infty} \frac{\sum_{k=1}^{n} X_k(\omega)}{n} = 0\right\} \in \mathscr{T},$$

  has probability 0 *or* 1.

- **Remark** (***Independence Assumption is Preferred***)
  The Komogorov zero-one law reveals the reason why ***the independence assumption is preferred*** in a lot of *statistical analysis* esp. concerning ***the consistency of statistics***.

  The theorem reveals that the tail σ-algebra $\mathscr{T} = \bigcap_{n\geq 1} \sigma(X_{n+1}, \ldots)$ for ***independent variables*** are <u>***almost trivial***</u>, thus the test run on the $\mathscr{T}$ is <u>***almost deterministic***</u>. This is because ***all tail events*** either ***form a null set*** or occupy ***the entire space outside a null set***.

References

Sidney I Resnick. *A probability path*. Springer Science & Business Media, 2013.