Lecture 7: 0-1 Laws

0-1 laws
\{X_i\}_{i \geq 1} - indpt random variables, \(\sigma(\{X_i\}_{i \geq 1})\) - cylindrical \(\sigma\)-algebras

**Definition 1.** \(A \in \sigma(\{X_i\}_{i \geq 1})\) - tail event if \(A \in \sigma(\{X_i\}_{i \geq n}) \forall n\)

\(A_i \in \sigma(X_i),\) then \(\{A_i\ i.o.\} = \bigcap_{n \geq 1} \bigcup_{i \geq n} A_i\) - tail event

**Theorem (Kolmogorov’s 0-1 law).** \(A\) - tail event, then \(\Pr(A) = 0\) or 1.

**Proof.**

\(\{X_i \geq i\}\) i.o. has only probability 0 or 1, cannot be split between half of the realizations.

\(A \in \sigma(\{X_i\}_{i \geq 1})\) - \(\sigma\)-algebra generated by algebra

\(\{X_{nk} \in B_k : B_k \text{ - Borel set in } \mathbb{R}, k = 1, ..., m\}\)

By approximation theory, can approximate any algebra by a smaller set that generates it.

\(\Delta\): symmetric difference

\(\forall \epsilon > 0, \exists A' \in \text{algebra} \text{ such that } \Pr(A \Delta A') \leq \epsilon, A' \in \sigma(X_1, ..., X_n)\)

\(|\Pr(A) - \Pr(A')| \leq \epsilon, |\Pr(A) - \Pr(AA')| \leq \epsilon, A \in \sigma(\{X_i\}_{i \geq n+1})\)

\(A, A'\ - \text{indpt, } \Pr(AA') = \Pr(A)\Pr(A'), \Pr(A) \approx \Pr(AA') = \Pr(A)\Pr(A') \approx \Pr(A)\Pr(A)\)

\(\epsilon \to 0, \Pr(A) = \Pr(A)^2 \Rightarrow 0\) or 1

**Example.**

1. \(\left\{ \sum_{i \geq 1} X_i \ - \text{converges} \right\}\) - tail event \(\Rightarrow 0\) or 1

2. \(\sum_{i \geq 1} X_iz^i\), \(z\)-complex, radius of convergence \(r = \limsup_{i \to \infty} |X_i|^{-\frac{1}{i}}\)

\(\{r \geq x\}\) - tail event \(\Rightarrow \Pr = 0\) or 1.

\(r = r_0\) - constant with \(p = 1\).

\[
\Pr(r \leq x)
\]

\begin{tabular}{c}
\hline
\midrule
\end{tabular}

Savage-Hewitt 0-1 law

\(\{X_i\}_{i \geq 1}\) iid, \((\mathbb{R}, \mathcal{B}, \Pr), X_i(x) = x: \mathbb{R} \to \mathbb{R}\) (identity on \(\mathbb{R}\))

\((\mathbb{R}^N, \mathcal{B}^\infty, \Pr), \{X_i\}_{i \geq 1} \in \mathbb{R}^N\) by Kolmogorov extension theory

**Definition 2.** \(A \in \mathcal{B}^\infty = \sigma(\{X_i\}_{i \geq 1})\) - exchangeable/symmetric

if \(\forall n(x_1, x_2, ..., x_n, x_{n+1}, ...) \in A \Rightarrow (x_n, x_2, ..., x_{n-1}, x_1, x_{n+1}, ...) \in A\)

**Theorem 1.** \(A\) - symmetric \(\Rightarrow \Pr(A) = 0\) or 1

Symmetric events are more general than tail events, but not all symmetric events are tail events

**Proof.**
\[ A : x_1 \ x_2 \ \ldots \ x_n \ \underbrace{x_{n+1} \ \ldots \ x_{2n} \ x_{2n+1} \ \ldots}^B \]

Can exchange \( A \) and \( B \) because of symmetry.

Define generator \( \Gamma x = (x_{n+1}, \ldots, x_{2n}, x_1, \ldots, x_n, x_{2n+1}, \ldots) \)
\( \Gamma A = \{ \Gamma x : x \in A \} \), \( A \) - symmetric \( \Rightarrow \Gamma A = A \)

By approximation theorem, \( \exists A_n \in \sigma(X_1, \ldots, X_n) \) and \( \mathbb{P}(A_n \Delta A) \leq \epsilon \)
\( B_n = \Gamma A_n \in \sigma(X_{n+1}, \ldots, X_{2n}) \)
\( \mathbb{P}(B_n \Delta A) = \mathbb{P}(\Gamma A_n \Delta \Gamma A) \overset{\text{id}}{=} \mathbb{P}(A_n \Delta A) \leq \epsilon \)
\( \mathbb{P}(A_n B_n \Delta A) \leq 2\epsilon, \mathbb{P}(A_n B_n) = \mathbb{P}(A_n) \mathbb{P}(B_n) \approx \mathbb{P}(A) \)
\( \mathbb{P}(A_n B_n) \approx \mathbb{P}(A), \epsilon \to 0, \mathbb{P}(A) = \mathbb{P}(A)^2 \Rightarrow 0, 1 \)

**Example.**

1. \( r = \limsup_{n \to \infty} \frac{S_n - a_n}{b_n} \overset{a.s.}{=} r_0, \{ r \leq x \} \) - symmetric event. \( \Rightarrow \mathbb{P}(r \leq x) = 0 \) or 1

2. Probability space: circle of length 1

\[
X_k = 1 \left( x \in \left[ 1 + \frac{1}{2} + \cdots + \frac{1}{k}, 1 + \cdots + \frac{1}{k+1} \right] \mod 1 \right)
\]

\( X_k \to 0 \) in probability, \( \mathbb{P}(|X_k - 0| \geq \epsilon) = \frac{1}{k+1} \to 0 \)

Does not converge almost surely because the sum continues on.

**Theorem (Kolmogorov’s inequality).** \( \mathbb{P}(|S_n - S_j| \geq a) \leq p < 1 \)
\[
\mathbb{P} \left( \max_{1 \leq j \leq n} |S_j| \geq x \right) \leq \frac{1}{1 - p} \mathbb{P}(|S_n| > x - a), x > a
\]

\[
\tau = \begin{cases} 
\min\{j \leq n : |S_j| \geq x\}, & \mathbb{P}(|S_n| \geq x - a) \geq \sum_{j=1}^{n} \mathbb{P}(|S_n| \geq x - a, \tau = j) \\
(1) + 1, & \text{otherwise}
\end{cases}
\]

\( \{\tau = j\} \cap \{\tau = i\} = \emptyset, i \neq j \leq n \)
\( \tau = j, |S_j| \geq x, |S_n - S_j| \leq a \Rightarrow |S_n| \geq x - a \)
\[
\geq \sum_{j=1}^{n} \mathbb{P}(\tau = j, |S_n - S_j| \leq a) = \sum_{j=1}^{n} \mathbb{P}(\tau = j) \mathbb{P}(|S_n - S_j| \geq a) \\
\{\tau = j\} = \sigma(X_1, \ldots, X_j), \{\sum_{j=1}^{n} X_j | \leq a\} \in \sigma(X_{j+1}, \ldots, X_n) \\
\geq \sum_{j=1}^{n} \mathbb{P}(\tau = j)(1 - p) = (1 - p) \mathbb{P}(\tau \leq n) = (1 - p) \mathbb{P} \left( \max_{1 \leq n} |S_j| \geq x \right)
\]

**Theorem (Kolmogorov).** If \( \sum_{i \geq 1} X_i \) converges in probability \( \Rightarrow \) converges a.s.
\( S_n \Rightarrow S \) in probability
Proof.
\[ \mathbb{P}(|S_n - S| \geq \epsilon) \leq \epsilon, \text{ for } n \geq n_0(\epsilon) \]
\[ \mathbb{P}(|S_n - S_k| \geq 2\epsilon) \leq \mathbb{P}(|S_n - S| \geq \epsilon) \leq \epsilon \text{ or } \mathbb{P}(|S_k - S| \geq \epsilon) \leq 2\epsilon \]
\[ n, k \geq n_0(\epsilon), k > n \]

Same proof as Kolmogorov, but start partial sum at \( n \).

\[ \mathbb{P} \left( \max_{n \leq j \leq k} |S_j - S_n| \geq 4\epsilon \right) \leq \frac{1}{1 - 2\epsilon} \mathbb{P}(|S_k - S_n| \geq 2\epsilon) \leq \frac{2\epsilon}{1 - 2\epsilon} \epsilon \text{ small } \leq 2\epsilon \]

Let \( k \to \infty \), sets are increasing.

\[ \mathbb{P} \left( \max_{n \leq j} |S_j - S_n| \geq 4\epsilon \right) \leq 3\epsilon \]

\[ \mathbb{P} \left( \max_{n \leq j} |S_j - S| \geq 5\epsilon \right) \leq 4\epsilon \text{ because } \mathbb{P}(|S_n - S| \geq \epsilon) \leq \epsilon \]

Take \( \epsilon = \frac{1}{m^2} \)

\[ A_m = \left\{ \max_{n = n(\epsilon) = n(m) \leq j} |S_j - S| \geq \frac{5}{m^2} \right\}, \sum \mathbb{P}(A_m) \leq \sum \frac{4}{m^2} < \infty \]

Apply Borel-Cantelli:
\[ \mathbb{P}(A_m \text{ i.o.}) = 0 \text{ with probability } 1 \]

\[ \max_{j \geq n(m)} |S_j - S| \leq \frac{5}{m^2} \Rightarrow S_j \to S \text{ with probability } 1 \]

To prove almost surely, always use Borel-Cantelli.

(Only used the fact that random variables are independent, did not use iid property).