

1 Independence

Let (Ω, F, P) be a probability space. Two sets $A, B \in F$ are independent if

$$P(A \cap B) = P(A)P(B).$$

Suppose g_1, \dots, g_n are σ -algebra such that $g_i \subset F$. Then g_1, \dots, g_n are independent if for any $A_j \in g_j$,

$$P(\cap A_j) = \prod_j P(A_j).$$

Suppose X_1, \dots, X_n are random variable on (Ω, F, P) , then they are independent if

$$\sigma(X_1), \dots, \sigma(X_n)$$

are independent, where

$$\sigma(X) = \{X^{-1}(B) | B \in \mathcal{B}\}.$$

We say a random variable X is independent of a σ algebra \mathcal{g} if $\sigma(X)$ is independent of \mathcal{g} .

Moreover, we define a set of random variables $\{X_i\}_{i \in I}$ are independent each other if any finite collection of them are independent.

Exercise

$X \perp\!\!\!\perp Y$ iff $P(X \leq x)P(Y \leq y) = P(X \leq x, Y \leq y)$.

Proof

\Rightarrow : Fix any $X, Y \in R$, then two sets

$$\{w | X(w) \leq x\} \in \sigma(X), \{w | Y(w) \leq y\} \in \sigma(Y)$$

are independent by b.

\Leftarrow If for any x, y $P(X \leq x)P(Y \leq y) = P(X \leq x, Y \leq y)$, then

$$\{w | X(w) \leq x\}, \{w | Y(w) \leq y\}$$

are independent of each other but they generate $\sigma(X), \sigma(Y)$.

Exercise

Proof by machine

1

If $X = 1_A, Y = 1_B$, then

$$E1_A 1_B = E1_{A \cap B} = P(A \cap B) = P(A)P(B) = E1_A E1_B = EXY.$$

2

If $X = \sum a_i 1_{A_i}, Y = \sum b_j 1_{B_j}$

$$\begin{aligned}
EXY &= E \sum a_i 1_{A_i} \sum b_j 1_{B_j} \\
&= E \sum_{i,j} a_i b_j 1_{A_i \cap B_j} \\
&= \sum_{i,j} a_i b_j P(A_i \cap B_j) \\
&= \sum_i a_i P(A_i) \sum_j b_j P(B_j) \\
&= EXEY.
\end{aligned}$$

Theorem

If $f: X \rightarrow [0, \infty]$ is measurable, then there exists real-valued simple functions s_1, s_2, \dots on X such that $0 \leq s_1 \leq s_2 \leq \dots \leq f, s_n(x) \rightarrow f(x)$ pointwise and uniformly on any set for which f is bounded.

3

By above theorem, if $X > 0, Y > 0$ a.e., then we can choose a sequence $x_n \uparrow x, y_n \uparrow y$

By definition of integration of non-negative measurable functions, $EX = \lim EX_n, EY = \lim EY_n$

$$EXY = \lim EX_n Y_n - \lim EX_n EY_n = EXEY.$$

4

For general integrable X and Y , since $|X|, |Y|$ are integrable,

$$EX^+ X^- < \infty, EY^+ + Y^- < \infty,$$

$$E|XY| = E|X|E|Y| \quad (\text{by 3})$$

which implies the result.

Corollary

If X_1, \dots, X_n are independent,

$$E(\prod X_i) = \prod E(X_i).$$

Stochastic process

Given a stochastic process $\{X_i\}_{i=1}^\infty, \Omega \rightarrow R^n$,

$$\omega \rightarrow X_1(\omega), X_2(\omega), \dots$$

2 Borel-cantelli lemma

2.1 Borel cantelli lemma 1

Suppose $\{A_n\}_{n=1}^\infty$ is a sequence of events and $\sum_{n=1}^\infty P(A_n) < \infty$, then

$$\begin{aligned} P(A_n \text{ infinitely occurs (denote i.o.)}) \\ &= P(\{w | w \text{ belongs to infinitely many } A_n\}) \\ &= P(\{w | \sum_{i=1}^\infty 1_{A_n}(w) = \infty\}) = 0. \end{aligned}$$

Proof:

$$\begin{aligned} P(A_n \text{ i.o.}) &= P(\cap_{n=1}^\infty \cup_{k \geq n} A_k) \\ &\leq P(\cup_{k \geq n} A_k) \\ &\leq \sum_{k=n}^\infty P(A_k). \end{aligned}$$

$n \rightarrow \infty$, we have that

$$P(A_n \text{ i.o.}) = 0.$$

2.2 Example

$\{X_n\}_{n=1}^\infty$ is a sequence of independent random variable's with

$$P(X_n = 0) = \frac{1}{n^2} \text{ then}$$

$$\sum_{n=1}^\infty P(X_n = 0) = \sum_{n=1}^\infty \frac{1}{n^2} < \infty$$

then by B-C-1, $P(X_n = 0 \text{ i.o.}) = 0$.

Foureir Series

Below is the fourier series, sine-cosine form.

$$f(x) = x^2 = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos((2\pi nx)/P) + b_n \sin((2\pi nx)/P)),$$

,where P is a function's period, so in this case $\cos nx$ where $b_n = 0$ if f is even.
If f be defined as x^2 on $[-\pi, \pi]$ Computing fourier series coefficients,

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx,$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx,$$

$$a_0 = \frac{2\pi^2}{3}, \quad a_n = \frac{(-1)^n 4}{n}, \quad b_n = 0.$$

$$\pi^2 = \frac{1}{3}\pi^2 + \sum_n \frac{4}{n^2}, \quad \sum_n \frac{4}{n^2} = \frac{2}{3}\pi^2$$

hence

$$\sum \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Example

Inverse of BC-I is not true. Consider $([0, 1], B, \text{unif})$.

$$A_n = [0, \frac{1}{n}], \quad \sum P(A_n) = \infty,$$

however, $P(A_n \text{ i.o.}) = p(\{0\}) = 0$.

3 Borel cantelli lemma -2

If $\{A_n\}$ independent and $\sum_{n=1}^{\infty} P(A_n) = \infty$ then

$$P(A_n \text{ i.o.}) = 1.$$

proof

Get used to the first step!!

$$\begin{aligned} P(A_n \text{ i.o.}) &= P(\cap \cup A_k) \\ &= \lim P(\cup_{k \geq n} A_k) \\ &= \lim (1 - P(\cap A_k^c)) \\ &= \lim [1 - \prod_{k=n}^{\infty} (1 - P(A_k))] \\ &\geq \lim (1 - e^{-\sum_{k=n}^{\infty} (1 - P(A_k))}) \quad (\text{since } 1 - x \leq e^{-x}) \\ &= 1. \end{aligned}$$

Example: simple 0-1 law

Let $\{A_n\}$ be a sequence of independent events then $P(A_n \text{ i.o.}) = 0$ or 1 .

$$\sum P(A_n) < \infty \iff 0$$

$$\sum P(A_n) = \infty \iff 1.$$

3.1 The kolmogorov 0-1 law

Let X_1, \dots be a sequence of independent r.v.'s. Define the tail σ -field be

$$\mathcal{T} = \cap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \dots)$$

Then if $A \in \mathcal{T}$, then $P(A) = 0$ or 1 . 증명 생략함 week4 note 볼 것

3.2 Theorem: If $A \in \mathcal{T}$ then $P(A)=0$ or 1 .

Take any such $A \in \mathcal{T}$. I claim \mathcal{A} is independent to $\sigma(X_1, \dots, X_{k-1})$ for any $k \geq 2$. Therefore

$$A \perp\!\!\!\perp \cup_{k=2}^{\infty} \sigma(X_1, \dots, X_{k-1}) = \mathcal{A}.$$

Since

$$\begin{aligned} \sigma(\mathcal{A}) &\supset \sigma(X_1, \dots), \\ \sigma(A) &= \sigma(X_1, \dots, X_n), \end{aligned}$$

therefore

$$A \perp\!\!\!\perp \sigma(X_1, \dots, X_n, \dots).$$

In particular $A \perp\!\!\!\perp A$, hence

$$P(A \cap A) = p(A) = P(A^2)$$

so $P(A)=0$ or 1 .

Application lattice bond percolation.

Notice that kolmogorov 0-1 law works for independent random variables. We can prove stronger results with iid random variables.

4 Theorem

Let $\{x_I\}_{I \in I}$ be a countably infinite set of random variables. Suppose $X = \{x_I\}_{I \in I}$ are iid distributed. Further, there is a function $f : R^n \rightarrow R$ and a collection Γ of one to one maps from I to I such that

1. f is symmetric or invariant under the action of $v \in \Gamma$

$$f(w^n) = f(w) \text{ for every } v \in \Gamma, w \in R^I.$$

2. *The set is large enough so that we can preserve the distance without touching the intersection.* for any finite subset $J \subseteq I$, there exists some $v \in \Gamma$ such that

$$V(J) \cap J = \phi.$$

Then $f(x)$ is a constant almost surely.

4.1 Example

$I = Z^d$ is countable and let $\Gamma = \{\text{all the translations}\}$. V is a translation mean

$$v : z^d \rightarrow z^d, v \rightarrow v + v_0, v_0 \in z^d.$$

Suppose J is a finite set of z^d , then we can find a large enough constant c such that

$$J \subseteq [-c, c]x[-c, c]x \dots x[-c, c].$$

If we choose a translation

$$z^d \rightarrow z^d$$

$$v \rightarrow v + \begin{pmatrix} 100c \\ \vdots \\ 100c \end{pmatrix}$$

Then we can see $v(J) \cap J = \phi$

5 Corollary

Let z^d be the d-dim integer lattice. E be the set of edges between neighbors. Also let Γ be the set of all the translations of E . For any function f which is translation invariant, i.e.

$$f(\{x_e\}_{e \in E}) = f(\{x_{v(e)}\}_{e \in E})$$

for $v \in \Gamma$ must be a constant a.s, where $p(x_e = 1) = p$ independently, $f(x_e)$ is a constant a.s.

5.0.1 Example

For example, let N be the number of infinite clusters, then N equals $\{0, 1, 2, \dots\} \cup \{\infty\}$ almost surely. Let $f(\{x_e\}) = \#$ of infinite clusters, then it is almost surely a constant. We don't know the exact number but know it is not random.

6 Corollary: Hewitt-savage 0-1 law

Let X_1, \dots, X_n are iid random variables. Suppose $f : R^{N^+} \rightarrow R$ satisfies

$$f(X) = f(X_\sigma),$$

where σ is a permutation on integers that permutes finitely many numbers. Then f is constant a.s.

6.0.1 Proof:

$I = N^+$, $\Gamma = \{\sigma : \sum_{i=1}^{\infty} 1_{\sigma(i) \neq i} < \infty\}$ for any $J \subset N^+$. We can find M such that $J \subseteq \{1, 2, \dots, M\}$ we can define σ :

$$1 \rightarrow M+1$$

$$2 \rightarrow M+2$$

$$\vdots$$

$$M \rightarrow 2M$$

$$M+1 \rightarrow 1$$

$$\vdots$$

$$2M \rightarrow M$$

Then $\sigma(J) \subseteq \{M+1, \dots, 2M\}$ and $\sigma(J) \cap J = \emptyset$.

7 Exercise 1

Suppose X_1, \dots, X_n iid r.v.s $x_i \in R^d$. Let $S_0 = 0, S_n = \sum_{i=1}^n X_i$ like a random walk, then consider

$$f(\{x_i\}) = 1_{\{s_n=0, i.o.\}}$$

function of path of random walk. f is invariant under permutations of finite elements. There exists N such that for every $k \geq N$,

$$\tilde{S}_k = S_k$$

so S_k hits 0 i.o. if and only if \tilde{S}_k does, so $p(\{s_n = 0, i.o.\}) = 0$ or 1 .

8 exercices 2: box problem

Suppose M boxes at each time t, throw a ball uniformly. At each time t, with probability 1, each box has the maximum number of balls infinitely open.

9 Convergence of random variables

1. $X_n \rightarrow x$ a.s.

$$P(\lim(X_n = X)) = P(\{w : X_n(w) \rightarrow X(w)\}) = 1.$$

2. $X_n \rightarrow x$ in p

$$\lim P(|X_n - x| > \epsilon) = 0.$$

3. $X_n \rightarrow x$ in L^r

$$E|X_n - X|^r \rightarrow 0$$

4. $X_n \rightarrow x$ in D

$$F_{X_n}(t) \rightarrow F_X(t)$$

for continuity point t of F_X

Note that a.s. implies p implies D and L^r implies p and p implies D.

9.1 Theorem 1

Show almost sure convergence implies convergence in probability.

$$A = \{w | X_n(w) \rightarrow X(w)\},$$

by assumption $P(A)=1$. For any $w \in A$, and for all $\epsilon > 0$, there exists $N(\epsilon)$ such that if $n > N$, then

$$|X_n(w) - X(w)| < \epsilon$$

\iff For any $w \in A$, for any ϵ , $w \in \{\cup_{n=1}^{\infty} \cap_{N \geq n} |X_n - X| < \epsilon\}$ Then it can be seen that

$$\begin{aligned} 1 &= P(\cup_{n=1}^{\infty} \cap_{N \geq n} |X_n - X| < \epsilon) \\ &= P(\liminf |X_n - X| < \epsilon) \\ &\leq \liminf P(|X_n - X| < \epsilon) \\ &\leq \limsup P(|X_n - X| < \epsilon) \leq 1 \end{aligned}$$

so we have convergence in p.

9.2 Thm 2

If $X_n \rightarrow X$ in L^r , then $X_n \rightarrow X$ in p.

$$P(|X_n - X| > \epsilon) = P(|X_n - X|^r > \epsilon^r) \leq \frac{E|X_n - X|^r}{\epsilon^r}$$

9.3 Thm 3

Convergence in P implies convergence in D.

Let t be a continuous point of F_X . We want to show

$$P(X_n \leq t) \rightarrow F_X(t) = P(X \leq t).$$

$$\begin{aligned} P(X_n \leq t) &= P(X_n \leq t, X > t + \epsilon) + P(X_n \leq t, X \leq t + \epsilon) \\ &\leq P(|X_n - X| > \epsilon) + P(X \leq t + \epsilon) \end{aligned}$$

복잡하게 생각하지 말고 포함관계로 생각하면 명확함. Then if $n \rightarrow \infty$, first term diminish,

$$\limsup F_n(t) \leq F_X(t + \epsilon)$$

take $\epsilon \rightarrow 0$,

$$\limsup F_n(t) \leq F_X(t). \quad (1)$$

Conversely,

$$\begin{aligned} P(X_n > t) &= P(X_n > t, X \leq t - \epsilon) + P(X_n > t, X > t - \epsilon) \\ &\leq P(|X_n - X| > \epsilon) + 1 - F_X(t - \epsilon) \end{aligned}$$

Then if $n \rightarrow \infty$, first term diminish,

$$\limsup F_n(t) \leq 1 - F_X(t - \epsilon)$$

ake $\epsilon \rightarrow 0$,

$$\limsup P(X_n > t) \leq 1 - F_X(t),$$

$$\limsup 1 - F_n(t) \leq 1 - F_X(t),$$

$$1 - \liminf F_n(t) \leq 1 - F_X(t), \quad (\text{I was not used to this.})$$

$$\liminf F_n(t) \geq F_X(t). \quad (2)$$

By (1) and (2), we have the convergence in D.