Probability Theory

III WLLN $\sim$ Independence

Seongho, Joo

August 7, 2025

Seoul National University

L^2 Space

Given a measure space $(\Omega, \mathcal{F}, \mu)$,

$$L^2(\Omega, \mathcal{F}, \mu) = \{f : \Omega \to \mathbb{R} \text{ measurable s.t } \int_\Omega f^2 \, \mathrm{d}\mu < \infty\}$$

Note: for any real number f, g,

$$0 \leq (|f| - |g|)^2 = f^2 - 2|fg| + g^2 \implies |fg| \leq \frac{1}{2}(f^2 + g^2)$$

Thus, if $f, g \in L^2$, then $\int |fg| \, \mathrm{d}\mu \leq \int \frac{1}{2}(f^2 + g^2) \, \mathrm{d}\mu = \frac{1}{2} \int f^2 \, \mathrm{d}\mu + \frac{1}{2} \int g^2 \, \mathrm{d}\mu < \infty$
This implies if $f, g, \in L^2$, then $fg \in L^1$
As a corollary, if μ is a finite measure, then $L^2(\mu) \subseteq L^1(\mu)$. (Take $g = 1 \in L^2$)

L^2 Space

For $f, g \in L^2(\Omega, \mathcal{F}, \mu)$, define:

$$||f||_{L^2} := (\int_\Omega f^2 \,\mathrm{d}\mu)^2, \quad \langle f, g \rangle := \int_\Omega fg \,\mathrm{d}\mu$$

Theorem 1 (Cauchy-Schwarz)

For $f, g \in L^2(\Omega, \mathcal{F}, \mu)$, then

$$|\langle f, g \rangle_{L^2}| \leq \int fg \,\mathrm{d}\mu \leq ||f||_{L^2} ||g||_{L^2}$$

Proof. For $t \in \mathbb{R}$, $p(t) = \int (|f| - t|g|)^2 \,\mathrm{d}\mu \geq 0$

L^2 Space

For $f, g \in L^2(\Omega, \mathcal{F}, \mu)$, define:

$$||f||_{L^2} := (\int_\Omega f^2 \,\mathrm{d}\mu)^2, \quad \langle f, g \rangle := \int_\Omega fg \,\mathrm{d}\mu$$

Theorem 1 (Cauchy-Schwarz)

For $f, g \in L^2(\Omega, \mathcal{F}, \mu)$, then

$$|\langle f, g \rangle_{L^2}| \leq \int fg \,\mathrm{d}\mu \leq ||f||_{L^2} ||g||_{L^2}$$

Proof. For $t \in \mathbb{R}$, $p(t) = \int (|f| - t|g|)^2 \,\mathrm{d}\mu \geq 0$

Corollary 1

$L^2(\Omega, \mathcal{F}, \mu)$ is a vector space, and $||f||_{L^2}^2 = \sqrt{\langle f, f \rangle_{L^2}}$ is a norm on it.

Fact: L^2 is actually a *Hilbert space*: it is Cauchy complete. If $f_n \in L^2$ s.t. $||f_n - f_m||_{L^2} \to 0$ as $n, m \to \infty$, then $\exists! f \in L^2$ such that $||f_n - f||_{L^2} \to 0$

Covariance

Definition 1

For $X, Y \in L^2$, let $\mathring{X} = X - \mathbb{E}[X], \mathring{Y} = Y - \mathbb{E}[Y]$. Their *covariance* is

$$\mathsf{Cov}(X, Y) := \mathbb{E}[\mathring{X}\mathring{Y}] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

For $X \in L^2$, its *variance* is

$$\mathsf{Var}(X) := \mathsf{Cov}(X, X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \leq 0$$

Lemma 1

If $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathsf{Var}(X) = 0$, then $X =$ Const a.s.

Example 1. $N = \mathsf{Pois}(\alpha)$, $\mathbb{E}[N] = \alpha$

$$\mathsf{Var}(N) = \mathbb{E}[(N - \alpha)^2] = \sum_{k=0}^{\infty}(k - \alpha)^2 e^{-\alpha}\frac{\alpha^k}{k!} = \alpha$$

Example 2. $X = \mathcal{N}(\alpha, t)$, $X = \sqrt{t}Z + \alpha$, $Z = \mathcal{N}(0, I)$

$$\mathsf{Var}(X) = \mathbb{E}[(X - \alpha)^2] = \mathbb{E}[(\sqrt{t}Z)^2] = t\mathbb{E}[Z^2] = t$$

Correlation

Definition 2

$X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ are *uncorrelated* if $\mathrm{Cov}(X, Y) = 0$ In general, their *correlation* is $\mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$

Proposition 1

$\mathrm{Cov}(X + \alpha, Y) = \mathrm{Cov}(X, Y + \alpha) = \mathrm{Cov}(X, Y)$ for any $\alpha \in \mathbb{R}$. As a result, if $X_1, \ldots, X_n$ are all (pairwise) uncorrelated , then

$$\mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}(X_1) + \ldots \mathrm{Var}(X_n)$$

Proof.

Chebyshev's inequality

Recall Markov inequality: if $f \geq 0, \varepsilon, p > 0$, then

$$\mu\{f \geq \varepsilon\} \leq \frac{1}{\varepsilon^p} \int f^p \, \mathrm{d}\mu$$

suppose μ is a probability measure, $X \in L^2$. Set $p = 2$, and apply Markov's inequality to $f = |\mathring{X}| = |X - \mathbb{E}[X]|$

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \varepsilon) = \mathbb{P}(|\mathring{X}| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}\mathbb{E}[|\mathring{X}|^2] = \frac{1}{\varepsilon^2}\mathsf{Var}(X)$$

The Weak Law of Large Numbers

Theorem 2

Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of L^2 random variables on a probability space, that are pairwise uncorrelated: $\text{Cov}(X_n, X_m) = 0$ if $n! = m$ and all with the same mean and variance: $\mathbb{E}[X_n] = \alpha,\ \text{Var}(X_n) = t, \quad \forall n$
Let $S_n = X_1 + \cdots + X_n$. Then for any $\varepsilon > 0$,

$$\mathbb{P}(|\frac{S_n}{n} - \alpha| \geq \varepsilon) \leq \mathcal{O}(\frac{1}{n})$$

It means that $\frac{S_n}{n}$ is *asymptotically* concentrated at α. But does it mean $\frac{S_n}{n} \to \alpha$ a.s.? Take X_n such that $\mathbb{P}(X_n = 1) = \frac{1}{n}, \mathbb{P}(X_n = 0) = 1 - \frac{1}{n}$. By Borel-Cantelli II: $\mathbb{P}(X_n = 1 \ \text{ i.o}) = 1$

Convergence in Measure

Definition 3

Let $(\Omega, \mathcal{F}, \mu)$ be a measurable space. Given a measurable functions $f_n, f : \Omega \to \mathbb{R}$, we say $f_n \underset{\mu}{\to} f$ if $\forall \varepsilon > 0, \lim_{n\to\infty} \mu\{|f_n - f| \geq \varepsilon\} = 0$

Theorem 3

Let $f_n, g_n, f, g, \in L^0(\Omega, \mathcal{F}, \mu)$.

1. (Uniqueness of limits) If $f_n \underset{\mu}{\to} f$ and $f_n \underset{\mu}{\to} g$, then $f = g$ μ-a.s.
2. If $\alpha, \beta \in \mathbb{R}, f_n \underset{\mu}{\to} f,$ and $g_n \underset{\mu}{\to} g$, then $\alpha f_n + \beta g_n \underset{\mu}{\to} \alpha f + \beta g$
3. If $f_n \underset{\mu}{\to} f$, then $\{f_n\}$ is cauchy in measure.

Convergence in Measure

Theorem 4

If $\{f_n\}$ is a L^0-cauchy sequence, then $\exists \in L^1$ such that subsequence $f_{n_k} \to f$ a.s. Moreover, $f_n \underset{\mu}{\to} f$.

Proof.

Almost sure convergence implies convergence in measure

Theorem 5

If $f_n \to f$ μ-a.s., then $f_n \underset{\mu}{\to} f$. f

Proof. For any $\varepsilon > 0$, $\mu\{|f_n - f| \geq \varepsilon i.o\} = 0$ Let $A_n = \{|f_n - f| \geq \varepsilon\}$.
$0 = \mu\{A_n \text{ i.o }\} = \mu(\cap_{k=1}^{\infty} \cap_{n \geq k} A_n) = \lim_{k\to\infty} \mu(\cup_{n\geq k} A_n)$
$\implies \mu(A_k) \geq \mu(\cup_{n\geq k} A_n) \to 0$
$\implies \mu\{|f_k - f| \geq \varepsilon\} \geq \mu\{\cup_{n\geq k} A_n) \to 0$

L^p convergence implies convergence in measure

For $1 \leq p < \infty$, $||f||_{L^p} := (\int_\Omega |f|^p \, \mathrm{d}\mu)^{1/p}$ defines a norm on $L^p(\Omega, \mathcal{F}, \mu)$. In particular,

$$||f + g||_{L^p} \geq ||f||_{L^p} + ||g||_{L^p}$$

for $f, g \in L^p$. Thus L^p is a normed vector space.

Lemma 2

Let $f_n, f \in L^p(\Omega, \mathcal{F}, \mu), (1 \geq p < \infty)$ with $||f_n - f||_{L^p} \to 0$, then $f_n \underset{\mu}{\to} f$.

Proof. Followed by Markov's inequality.

The converse is false. Take $f_n = n \cdot \mathbf{1}_{[0,1/n]}$

L^p space is complete

Theorem 6

For $1 \leq p < \infty$, $f_n \in L^p$, $||f_n - f_m||_{L^p} \to 0$ as $n, m \to \infty$
$\implies \exists f \in L^p$ such that $||f_n - f||_{L^p} \to 0$

Proof. Since f_n is L^0-Cauchy. There exists (n_k) such that $f_{n_k} \to f$ a.s, $f \in L^0$.

$$||f_{n_k} - f||_{L^p}^p = \int \lim_{j\to\infty} |f_{n_k} - f_{n_j}|^p \, \mathrm{d}\mu \overset{\dagger}{\leq} \liminf_{j\to\infty} \int |f_{n_k} - f_{n_j}|^p \, \mathrm{d}\mu = \liminf_{j\to\infty} ||f_{n_k} - f_{n_j}||_{L^p}^p$$

†: holds by Fatou's lemma. Take $k \to \infty$, then $||f_{n_K} - f||_{L^p} \to 0$

$$||f_n - f||_{L^p} \leq ||f_n - f_{n_k}|| + ||f_{n_k} - f||_{L^p} \to 0$$

Multiplicative System

A set $\mathbb{H}$ of $\mathbb{R}$-valued functions on Ω is *closed under bounded convergence* if

$f_n \in \mathbb{H}, \exists M < \infty$ s.t. $|f_n(w)| \leq M \; \forall n \in \mathbb{N}, w \in \Omega, \; \lim_{n\to\infty} f_n(w) = f(w) \in \mathbb{R} \; \forall w \in \Omega$

$\implies f \in \mathbb{H}$

Remark. $C_c(\mathbb{R}), C_b(\mathbb{R})$ are *not* closed under bounded convergence.

Multiplicative System

A set $\mathbb{H}$ of $\mathbb{R}$-valued functions on Ω is *closed under bounded convergence* if
$f_n \in \mathbb{H}, \exists M < \infty$ s.t. $|f_n(w)| \leq M \ \forall n \in \mathbb{N}, w \in \Omega, \ \lim_{n \to \infty} f_n(w) = f(w) \in \mathbb{R} \ \forall w \in \Omega$
$\implies f \in \mathbb{H}$

Remark. $C_c(\mathbb{R}), C_b(\mathbb{R})$ are *not* closed under bounded convergence.

Notation: Given a collection $\mathbb{M}$ of $\mathbb{R}$-valued bounded functions on Ω, let

$\mathbb{H}(\mathbb{M}) :=$ the smallest subspace of $\mathbb{B}(\Omega)$ containing $\mathbb{M} \cup \{1\}$, and closed under bounded convergence.

Multiplicative System Theorem

Theorem 7

Let $\mathbb{H} \subseteq \mathbb{B}(\Omega)$ be a subspace, containing 1, and closed under bounded convergence.
Let $\mathbb{M} \subseteq \mathbb{H}$ be a *multiplicative system*: $f, g \in \mathbb{M} \to f \cdot g \in \mathbb{M}$
Then $\mathbb{H}$ contains all bounded $\sigma(\mathbb{M})$-measurable functions: $\mathbb{B}(\Omega, \sigma(\mathbb{M})) \subseteq \mathbb{H}$. In fact, $\mathbb{B}(\Omega, \sigma(\mathbb{M}) = \mathbb{H}(\mathbb{M})$

Multiplicative System Theorem

Theorem 7

Let $\mathbb{H} \subseteq \mathbb{B}(\Omega)$ be a subspace, containing 1, and closed under bounded convergence. Let $\mathbb{M} \subseteq \mathbb{H}$ be a *multiplicative system*: $f, g \in \mathbb{M} \to f \cdot g \in \mathbb{M}$
Then $\mathbb{H}$ contains all bounded $\sigma(\mathbb{M})$-measurable functions: $\mathbb{B}(\Omega, \sigma(\mathbb{M})) \subseteq \mathbb{H}$. In fact, $\mathbb{B}(\Omega, \sigma(\mathbb{M}) = \mathbb{H}(\mathbb{M})$

Corollary 2

$\mathbb{H}(C_c(\mathbb{R})) = \mathbb{B}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ i.e. the bounded convergence closure of the compactly-supported continuous functions is all bounded Borel measurable functions.

Corollary 3

Suppose ν, μ are Borel probability measures on $\mathbb{R}$, and

$$\int_{\mathbb{R}} f \, \mathrm{d}\mu = \int_{\mathbb{R}} f \, \mathrm{d}\nu \quad f \in C_c(\mathbb{R})$$

Then $\mu = \nu$

Product Measure

Definition 4

Given two measurable spaces $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2)$,

$$\mathcal{F}_1 \otimes \mathcal{F}_2 := \sigma(\mathcal{F}_1 \times \mathcal{F}_2) = \sigma(A_1 \times A_2 : A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2)$$

By induction, larger products are

$$\bigotimes_{j=1}^{d} \mathcal{F}_j = \sigma\{\prod_{j=1}^{d} B_j : B_j \in \mathcal{F}_j, 1 \leq j \leq d\}$$

Fact: $\bigotimes_{j=1}^{d} \mathcal{F}_j = \sigma\{\pi_k : 1 \leq k \leq d\}$, where π_k be the standard projection.

Product Measure

Definition 4

Given two measurable spaces $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2)$,

$$\mathcal{F}_1 \otimes \mathcal{F}_2 := \sigma(\mathcal{F}_1 \times \mathcal{F}_2) = \sigma(A_1 \times A_2 : A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2)$$

By induction, larger products are

$$\bigotimes_{j=1}^{d} \mathcal{F}_j = \sigma\{\prod_{j=1}^{d} B_j : B_j \in \mathcal{F}_j, 1 \leq j \leq d\}$$

Fact: $\bigotimes_{j=1}^{d} \mathcal{F}_j = \sigma\{\pi_k : 1 \leq k \leq d\}$, where π_k be the standard projection.

Lemma 3 (Product Measurability)

Let $(\Omega_j, \mathcal{F}_j)_{j \in J}$ and $(\Upsilon, \mathcal{B})$ be measurable spaces.
Then, $f : \Upsilon \to \prod_{j \in J} \Omega_j$ is $\mathcal{B}/\bigotimes_{j \in J} \mathcal{F}_j$-measurable. if and only if $\pi_k \circ f : \Upsilon \to \Omega_k$ is $\mathcal{B}/\mathcal{F}_k$-measurable $\forall k \in J$

Product Measure

Theorem 8

Let $(\Omega_j, \mathcal{F}_j, \mu_j)$, $j = 1, 2$ be σ-finite measure spaces. Let $f : \Omega_1 \times \Omega_2 \to [0, \infty)$ be a non-negative $\mathcal{F}_1 \otimes \mathcal{F}_2 / \mathcal{B}(\mathbb{R})$-measurable function. Then,

1. $w_1 \mapsto f(w_1, w_2)$ is $\mathcal{F}_1/\mathcal{B}(\mathbb{R})$-measurable $\forall w_2 \in \Omega_2$
   $w_2 \mapsto f(w_1, w_2)$ is $\mathcal{F}_2/\mathcal{B}(\mathbb{R})$-measurable $\forall w_1 \in \Omega_1$
2. $w_1 \mapsto \int_{\Omega_2} f(w_1, w_2)\, \mu_2(\mathrm{d}w_2)$ is $\mathcal{F}_1/\mathcal{B}(\bar{\mathbb{R}})$-measurable
   $w_2 \mapsto \int_{\Omega_1} f(w_1, w_2)\, \mu_1(\mathrm{d}w_1)$ is $\mathcal{F}_2/\mathcal{B}(\bar{\mathbb{R}})$-measurable
3. $\int_{\Omega_1} \left(\int_{\Omega_2} f(w_1, w_2)\, \mu_2(\mathrm{d}w_2) \mu_1(\mathrm{d}w_1) \right) =$
   $\int_{\Omega_2} \left(\int_{\Omega_1} f(w_1, w_2)\, \mu_1(\mathrm{d}w_2) \mu_2(\mathrm{d}w_1) \right)$

Proof.
Step 1. Verify that 1-3 hold for $f = f_1 \otimes f_2, f_j \in \mathbb{B}(\Omega_j, \mathcal{F}_j)$
Step 2. Using Dynkin's Multiplicative Systems Theorem, show that
$\mathbb{B}(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2) \subseteq \mathbb{H} = \{ f \in \mathbb{B} :$ 1-3 hold $\}$
Step 3. Show that it holds for non-negative, measurable $\forall f$

Tonelli-Fubini

Theorem 9 (Fubini)

Let $f \in L^0(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mu_1 \otimes \mu_2)$. TFAE:

1. $f \in L^1(\Omega_1 \times \Omega_2, \mathcal{F}_1, \otimes \mathcal{F}_2, \mu_1 \otimes \mu_2)$
2. $\int_{\Omega_1} \left(\int_{\Omega_2} |f(w_1, w_2)| \mu_2(\mathrm{d}w_2) \right) \mu_1(\mathrm{d}w_1) < \infty$
3. $\int_{\Omega_2} \left(\int_{\Omega_1} |f(w_1, w_2)| \mu_1(\mathrm{d}w_1) \right) \mu_2(\mathrm{d}w_2) < \infty$

In this case,

$$
\begin{gathered}
w_1 \mapsto f(w_1, w_2) \in L^1(\Omega, \mathcal{F}_1, \mu_1) \text{ for } \mu_2\text{-a.e. } w_2 \\
w_2 \mapsto f(w_1, w_2) \in L^1(\Omega_2, \mathcal{F}_2, \mu_2) \text{ for } \mu_1\text{-a.e. } w_1 \\
w_2 \mapsto \int f(w_1, w_2) \mu_1(\mathrm{d}w_1) \in L^1(\Omega_2, \mathcal{F}_2, \mu_2), \\
w_1 \mapsto \int f(w_1, w_2) \mu_2(\mathrm{d}w_2) \in L^1(\Omega_1, \mathcal{F}_1, \mu_1)
\end{gathered}
$$

and $\int_{\Omega_1 \times \Omega_2} f \, \mathrm{d}(\mu_1 \otimes \mu_2) = \int_{\Omega_1} \left(\int_{\Omega_2} f(w_1, w_2) \mu_2(\mathrm{d}w_2) \right) \mu_1(\mathrm{d}w_1)$, the integration order can be changed.

Independence

Let $\mathcal{C}_1, \ldots \mathcal{C}_n \subseteq \mathcal{F}$ be collections of events.

Definition 5

$\mathcal{C}_1, \ldots, \mathcal{C}_n \subseteq \mathcal{F}$ are *independent* if for $I \subseteq [n] = \{1, \ldots, n\}$

$$\mathbb{P}(\bigcap_{i \in I} A_i) = \prod_{i \in I} \mathbb{P}(A_i), \quad \forall A_i \in \mathcal{C}_i, i \in I$$

For infinite case, let $\{\mathcal{C}_t\}_{t \in T}$ be any collection of subsets of $\mathcal{F}$. We call them independent if and only if, for all finite subsets $J \subset T, \{\mathcal{C}_j\}_{j \in J}$ is independent.

Observation: If $\mathcal{C}_1, \ldots, \mathcal{C}_n$ are independent, so are $\mathcal{C}_1 \bigcup \Omega, \ldots \mathcal{C}_n \bigcup \Omega$. This makes the notation so much easier.

Independence

Let $\mathcal{C}_1, \dots \mathcal{C}_n \subseteq \mathcal{F}$ be collections of events.

Definition 5

$\mathcal{C}_1, \dots, \mathcal{C}_n \subseteq \mathcal{F}$ are *independent* if for $I \subseteq [n] = \{1, \dots, n\}$

$$\mathbb{P}(\bigcap_{i \in I} A_i) = \prod_{i \in I} \mathbb{P}(A_i), \quad \forall A_i \in \mathcal{C}_i, i \in I$$

For infinite case, let $\{\mathcal{C}_t\}_{t \in T}$ be any collection of subsets of $\mathcal{F}$. We call them independent if and only if, for all finite subsets $J \subset T, \{\mathcal{C}_j\}_{j \in J}$ is independent.

Observation: If $\mathcal{C}_1, \dots, \mathcal{C}_n$ are independent, so are $\mathcal{C}_1 \bigcup \Omega, \dots \mathcal{C}_n \bigcup \Omega$. This makes the notation so much easier.

Lemma 4

If $\mathcal{C}_1, \dots, \mathcal{C}_n \subseteq \mathcal{F}$ and $\Omega \in \mathcal{C}_j$ for all $j \in [n]$ then $\mathcal{C}_1, \dots, \mathcal{C}_n$ are independent if and only if

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \prod_{i=1}^{n} \mathbb{P}(A_i), \quad \forall A_i \in \mathcal{C}_i \cup \Omega$$

Independence

A collection $\mathcal{C} \subseteq \mathcal{F}$ is a π-**system** if it is closed under finite intersections.
$A, B \in \mathcal{C} \implies A \cap B \in \mathcal{C}$

Theorem 10

If $\mathcal{C}_1, \ldots \mathcal{C}_n \subseteq \mathcal{F}$ are independent π-systems, then $\sigma(\mathcal{C}_1), \ldots, \sigma(\mathcal{C}_n)$ are independent.

Proof. Use Lemma **5**, prove the case for $n = 2$.
In general, the independence of collections of events does *not* implies the σ-field independence.

Lemma 5

If $\mathcal{C} \subseteq \mathcal{F}$ is π-system, and μ, ν are probability measure on $\mathcal{F}$ such that $\mu = \nu$ on $\mathcal{C}$, then $\mu = \nu$ on $\sigma(\mathcal{C})$

Proof. Take $\mathbb{M} = \{\mathbf{1}_B : B \in \mathcal{C}\} \subseteq \mathbb{B}(\mathcal{F})$ which is a multiplicative system. Note that $\sigma(\mathcal{C}) \subseteq \sigma(\mathbb{M})$.

Borel-Cantelli Lemma II

Lemma 6

Let $\{A_n\}_{n=1}^{\infty}$ be an infinite sequence of independent events. Then

$$\mathbb{P}(\bigcap_{n=1}^{\infty} A_n) = \prod_{n=1}^{\infty} \mathbb{P}(A_n) := \lim_{M\to\infty} \prod_{n=1}^{M} \mathbb{P}(A_n)$$

Lemma 7 (Borel-Cantelli Lemma II)

Let $\{A_n\}_{n=1}^{\infty}$ be independent events. If $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$, then $\mathbb{P}(\{A_n \text{ i.o}\}) = 1$

Proof. $\{A_n \text{ i.o }\} = \bigcap_{k=1}^{\infty} \bigcup_{n\geq k} A_n$

Independent Random Variables

$X_i : (\Omega, \mathcal{F}, \mathbb{P}) \to (S_i, \mathcal{B}_i) = (\mathbb{R}^{d_i}, \mathcal{B}(\mathbb{R}^{d_i}))$
$\sigma(X_i) =$ minimal σ-field s.t X_i is $\mathcal{F}/\mathcal{B}_i$-measurable.

Definition 6

Random variables $\{X_i\}_{i \in I}$ are *independent* if the σ-field $\{\sigma(X_i)\}_{i \in I}$ are independent.

i.e. $\forall B_i \in \mathcal{B}_i$, $\{X_i^{-1}(B_i)\}_{i \in I}$ are independent.

$$\mathbb{P}(X_1 \in B_1, \ldots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \ldots \mathbb{P}(X_n \in B_n)$$

Independent Random Variables

$X_i : (\Omega, \mathcal{F}, \mathbb{P}) \to (S_i, \mathcal{B}_i) = (\mathbb{R}^{d_i}, \mathcal{B}(\mathbb{R}^{d_i}))$
$\sigma(X_i) =$ minimal σ-field s.t X_i is $\mathcal{F}/\mathcal{B}_i$-measurable.

Definition 6

Random variables $\{X_i\}_{i \in I}$ are *independent* if the σ-field $\{\sigma(X_i)\}_{i \in I}$ are independent.

i.e. $\forall B_i \in \mathcal{B}_i$, $\{X_i^{-1}(B_i)\}_{i \in I}$ are independent.

$$\mathbb{P}(X_1 \in B_1, \ldots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \ldots \mathbb{P}(X_n \in B_n)$$

Lemma 8

Given random variables $X_i : (\Omega, \mathcal{F}, \mathbb{P}) \to (S_i, \mathcal{B}_i)$, if $\mathcal{E}_i \subseteq \mathcal{B}_i$ are π-systems s.t $\sigma(\mathcal{E}_i) = \mathcal{B}_i$, then $\{X_i\}_{i \in I}$ are independent if and only if $\{X_i^{-1}(E_i)\}_{i \in I}$ are independent $\forall E_i \in \mathcal{E}_i$

- As a result, the X_i's are independent if and only if

$$\mathbb{P}(X_1 \le t_1, \ldots, X_n \le t_n) = \mathbb{P}(X_1 \le t_1) \ldots \mathbb{P}(X_n \le t_n) = F_{X_1}(t_1) \ldots F_{X_n}(t_n)$$

Independent Random Variables

Given $\bar{X} = (X_1, \dots X_n),\ X_i : (\Omega, \mathcal{F}, \mathbb{P}) \to (S_i, \mathcal{B}_i)$, their *joint law* $\mu_{\bar{X}}$ is the probability measure on $\mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_n$ defined by $\mu_{\bar{X}} := \mathbb{P} \circ \bar{X}^{-1}$

Theorem 11

$X_1, \dots, X_n$ are independent if and only if $\mu_{X_1,\dots,X_n} = \mu_{X_1} \otimes \dots \mu_{X_n}$

Proof.

Independent Random Variables

Theorem 12 (Independence conditions)

Let $X_i : (\Omega, \mathcal{F}, \mathbb{P}) \to (S_i, \mathcal{B}_i)$ be random variables, $i \in [n]$. set $\bar{X} = (X_1, \dots X_n)$. TFAE:

1. $X_1, \dots, X_n$ are independent.
2. $\mu_{\bar{X}} = \mu_{X_1} \otimes \cdots \otimes \mu_{X_n}$.
3. $\mathbb{E}[f_1(X_1) \dots f_n(X_n)] = \mathbb{E}[f_1(X_1)] \cdots \mathbb{E}[f_n(X_n)] \quad \forall f_i \in \mathbb{B}(S_i, \mathcal{B}_i) \cdots (\dagger)$
   Moreover, if each $(S_i, \mathcal{B}_i) = (\mathbb{R}^{d_i}, \mathcal{B}(\mathbb{R}^{d_i}))$, we also have
4. $\dagger$ holds $\forall f_i \in C_c(\mathbb{R}^{d_i})$
5. $\dagger$ holds $\forall f_i$ of the form $f_i = \mathbf{1}_{(-\infty, t_1] \times \cdots \times (-\infty, t_{d_i}]}$

Proof. $3 \implies 1$, $f_i = \mathbf{1}_{B_i}$, $B_i \in \mathcal{B}_i$

Grouping and Functions

Lemma 9

If $\mathcal{F}_1, \dots, \mathcal{F}_n$ are independent, σ-fields and $n = n_1 + n_2 + \dots + n_k$, then

$$\mathcal{G}_1 = \sigma(\mathcal{F}_1 \cup \dots \cup \mathcal{F}_{n_1}), \dots, \mathcal{G}_k = \sigma(\mathcal{F}_{n_1 + \dots n_{k-1} + 1} \cup \dots \cup \mathcal{F}_n)$$

are independent σ-fields

Corollary 4

Let $X_i : (\Omega, \mathcal{F}, \mathbb{P}) \to (S_i, \mathcal{B}_i)$ be independent, $i \in [n]$. Let $n = n_1 + n_2 + \dots + n_k$. Let $f_j : S_{n_1 + \dots + n_{j-1} + 1} \times \dots \times S_{n_1 + \dots + n_j} \to \mathbb{R}$ be measurable, $j \in [k]$. Then $Y_j = f_j(X_{n_1 + \dots n_{j-1} + 1}, \dots X_{n_1 + \dots + n_j})$ are independent, $j \in [k]$.

Example. If X_1, X_2, X_3, X_4, X_5 are independent, so are $X_1 + X_2, X_3 X_4, e^{X_5}$

Method of Moments for testing independence

Proposition 2

Let $X_i : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})$ be *bounded* random variables. Then, $X_1, \ldots X_n$ are independent if and only if

$$\mathbb{E}[X_1^{k_1} \ldots X_n^{k_n}] = \mathbb{E}[X_1^{k_1}] \ldots \mathbb{E}[X_n^{k_n}] \quad \forall k_1, \ldots, k_n \in \mathbb{N}$$