Independence

Sayan Mukherjee

Independence

We start with a collection of finite (2) events and then extend to events of σ algebras.

Definition 0.0.1 *Suppose $(\Omega, \mathcal{F}, \mathbf{P})$ is a fixed probability space. Events $A, B \in \mathcal{F}$ are independent if*

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B).$$

We now extend to finite numbers of events.

Definition 0.0.2 *The events $A_1, ..., A_n$ are independent if*

$$\mathbf{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbf{P}(A_i), \quad \text{for all finite } I \subset \{1, ..., n\}.$$

The index set I corresponds to

$$\sum_{k=2}^{n} \binom{n}{k} = 2^n - n - 1,$$

equations.

We want to extend this to cases where A_i are elements in a σ algebra. We first look at classes (collections) of events.

Definition 0.0.3 *Classes of events $\mathcal{C}_i \subset \mathcal{F}$ are independent if for any choice of $A_i, ..., A_n$ with $A_i \in \mathcal{C}_i, i = 1, ..., n$ the events A_i are independent.*

We now provide a basic criterion for independence of elements of a σ algebra.

Theorem 0.0.1 *If for each $i = 1, ..., n$ $\mathcal{C}_i$ is a non-empty class satisfying*

1. *$\mathcal{C}_i$ is a π-system*
2. *$\mathcal{C}_i, 1 = 1, ..., n$ are independent*

then

$$\sigma(\mathcal{C}_1), ..., \sigma(\mathcal{C}_n),$$

are independent.

Proof. We prove the result for $n = 2$ and then use induction.

Fix $A_2 \in \mathcal{C}_2$ and let

$$\mathcal{L} = \{A \in \mathcal{F} : \mathbf{P}(A \cap A_2) = \mathbf{P}(A)\mathbf{P}(A_2)\}.$$

It holds that $\mathcal{L}$ is a λ system since we can verify

1. $\Omega \in \mathcal{L}$ since
$$\mathbf{P}(\Omega \cap A_2) = \mathbf{P}(A_2) = \mathbf{P}(\Omega)\mathbf{P}(A_2).$$
2. If $A \in \mathcal{L}$ then $A^c \in \mathcal{L}$ since
$$\begin{aligned} \mathbf{P}(A^c \cap A_2) &= = \mathbf{P}((\Omega \setminus A) \cap A_2) = \mathbf{P}(A_2 \setminus (A \cap A_2)) \\ &= \mathbf{P}(A_2) = P(A \cap A_2) = \mathbf{P}(A_2) - \mathbf{P}(A)\mathbf{P}(A_2) \\ &= \mathbf{P}(A_2)(1 - \mathbf{P}(A)) = \mathbf{P}(A^c)\mathbf{P}(A_2). \end{aligned}$$

3. If $B_n \in \mathcal{L}$ are disjoint for $(n \geq 1)$ then $\sum_{n=1}^{\infty} B_n \in \mathcal{L}$ since

$$\begin{aligned} \mathbf{P}\left(\left[\bigcup_{n=1}^{\infty} B_n\right] \cap A_2\right) &= \mathbf{P}\left(\bigcup_{n=1}^{\infty} (B_n \cap A_2)\right) = \sum_{i=1}^{\infty} \mathbf{P}(B_n \cap A_2) \\ &= \sum_{i=1}^{\infty} \mathbf{P}(B_n)\mathbf{P}(A_2) = \mathbf{P}\left(\bigcup_{n=1}^{\infty} B_n\right)\mathbf{P}(A_2). \end{aligned}$$

Since $\mathcal{L}$ is a λ system and a π system and $\mathcal{C}_1 \subset \mathcal{L}$ by Dynkin's theorem $\sigma(\mathcal{C}_1) \subset \mathcal{L}$. This gives us $\sigma(\mathcal{C}_1) \perp\!\!\!\perp \mathcal{C}_2$. We can extend this argument to show $\sigma(\mathcal{C}_1) \perp\!\!\!\perp \sigma(\mathcal{C}_2)$ and then extend this $n > 2$ by induction. □

We now develop the idea of independence of random variables. A collection of random variables $\{X_i\}$ on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ are called independent if

$$\mathbf{P}\left(\bigcap_{i \in I} X_i \in B_i]\right) = \prod_{i \in I} \mathbf{P}[X_i \in B_1]$$

for each finite set I of indices and each collection of Borel sets $\{B_i \in \mathcal{B}(\mathbb{R})\}$ This can be reworded as the σ-algebras $\mathcal{F}_1 := \sigma(X_i) = X_i^{-1}(\mathcal{B})$ be independent. It is enough to check that the joint CDF factors

$$\mathbf{P}\left(\bigcap_{i=1}^{I} [X_i \leq x_i]\right) = \prod_{i \in I} F(X_i)$$

for each $x \in \mathbb{R}^I$.

Zero-One Laws

The idea of zero-one laws is to ask about events that either happen almost surely or don't happen at all. This is really a section on how to get strong law of large numbers results. Often the event is the convergence of sums or averages of independent random variables.

The following are a few examples: $\xi_1, ..., \xi_n \overset{iid}{\sim} \text{Rad}(p)$ where $\text{Rad}(p)$ is a rademacher rv with $\mathbf{P}(\xi = 1) = p = 1/2$ and $\mathbf{P}(\xi = -1) = 1 - p = 1/2$. The set

$$A_i = \left\{\omega : \frac{\sum_{i=1}^{n} \xi_i}{n} \text{ converges}\right\}$$

is the set of sample points that converge, what is $\mathbf{P}(A_i)$?

The standard way of thinking of this set is as $\limsup_{m \to \infty} A_n$. What we will show is that often $\mathbf{P}(\limsup_{m \to \infty} A_n)$ either is 0 or 1.

An important idea in this analysis is the concept or a tail algebra. Let $\mathcal{F}_n^{\infty} = \sigma(\xi_n, \xi_{n+1}, ...)$ be the σ-algebra generated by the sequence $\xi_n, \xi_{n+1}, ...$ then the tail algebra is

$$\mathcal{T} = \bigcap_{n=1}^{\infty} \mathcal{F}_n^{\infty}$$

this is a σ-algebra since the intersection of σ-algebras is a σ-algebra. It is called a tail algebra because for any $A \in \mathcal{T}$, $A \perp\!\!\!\perp \{\xi_1, ..., \xi_n\}$ for every finite n. The following are a few examples of tail algebras

$$\begin{aligned} A_1 &= \left\{\omega : \frac{\sum_{i=1}^{n} \xi_i}{n} < \infty\right\} \\ A_2 &= \left\{\omega : \frac{\sum_{i=1}^{n} \xi_i}{n} < C\right\} \\ A_3 &= \left\{\omega : \frac{\sum_{i=1}^{n} \xi_i}{\sqrt{2n \log n}} = 1\right\} \end{aligned}$$

Lemma 0.0.1 (Borel-Cantelli) *Let $\{A_n\}$ be events on $(\Omega, \mathcal{F}, \mathbf{P})$ s.t.*

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty,$$

then

$$\mathbf{P}\left(\limsup_{n\to\infty} A_n\right) = 0.$$

Proof.

$$\mathbf{P}\left(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_n\right) \leq \mathbf{P}\left(\bigcup_{m=n}^{\infty} A_n\right) \leq \sum_{m=n}^{\infty} \mathbf{P}(A_n) \to 0. \ \square$$

Lemma 0.0.2 (Borel Zero-One Law) *Let $\{A_n\}$ be independent events on $(\Omega, \mathcal{F}, \mathbf{P})$ s.t.*

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty,$$

then

$$\mathbf{P}\left(\limsup_{n\to\infty} A_n\right) = 1.$$

This means that the probability of $\{A_n\}$ i.o. is one.

Proof. Recall that $1 + x \leq e^x$. For $1 \leq n \leq N < \infty$

$$\begin{aligned}
\mathbf{P}\left(\bigcap_{m=n}^{\infty} A_n^c\right) &= \prod_{m=n}^{N}(1 - \mathbf{P}(A_m)) \\
&\leq \prod_{m=n}^{\infty} e^{-P(A_m)} = \exp\left(-\sum_{m=n}^{\infty} \mathbf{P}(A_m)\right) \\
&\to \exp\left(-\sum_{m=n}^{\infty} \mathbf{P}(A_m)\right) = e^{-\infty} = 0.
\end{aligned}$$

This implies that $\cap_{m=n}^{\infty} A_m^c = \emptyset$ and

$$\begin{aligned}
\mathbf{P}\left(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_n\right) &= 1 - \mathbf{P}\left(\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_n^c\right) \\
&\geq 1 - \mathbf{P}\left(\bigcap_{m=n}^{\infty} A_n^c\right) = 1. \ \square
\end{aligned}$$

The above results combine to give the Borel zero-one law based on sums of probabilities of independent events. The next result provides a zero-one law based on tail events.

Theorem 0.0.2 (Kolmogorov Zero-One Law) *Let $\{\xi_1, ..., \xi_n\}$ be a sequence of independent random variables and $A \in \mathcal{T}$ be a tail event*

$$\mathbf{P}(A) = 1 \quad \textit{or} \quad \mathbf{P}(A) = 0.$$

Proof.

$$\mathcal{F}_n = \sigma\{\xi_i : i \leq n\}, \quad \mathcal{P} = \bigcup_{n=1}^{\infty} \mathcal{F}_n.$$

Let $B \in \mathcal{P}$ and $A \in \mathcal{T}$. For some $n \in \mathbb{N}$, $B \in \mathcal{F}_n$, $A \in \sigma\{\xi_i : i \geq n+1\}$, $B \perp\!\!\!\perp A$. So $\mathcal{P}$ and $\mathcal{T}$ are independent and $\mathcal{P}$ is a π-system so $\sigma(\mathcal{P}) \perp\!\!\!\perp \mathcal{T}$.

Each ξ_n is $\sigma(\mathcal{P})$ measurable so $\mathcal{T} \subset \sigma(\mathcal{P})$ and each $A \in \mathcal{T}$ must also be in $\sigma(\mathcal{P})$, also $\sigma(\mathcal{P}) \perp\!\!\!\perp \mathcal{T}$ so

$$\mathbf{P}(A) = \mathbf{P}(A \cap A) = \mathbf{P}(A)\mathbf{P}(A) = \mathbf{P}^2(A),$$

since every tail event is independent of itself. The above implies

$$\mathbf{P}(A)(1 - \mathbf{P}(A)) = 0. \ \square.$$

The last zero-one law we consider extends results from tail events to symmetric functions. It allows us to make states about things like infinite crossings of zero of a random walk. The following setup illustrates the ideas we will explore. Consider $\xi_1, ..., \xi_n \overset{iid}{\sim} \text{Rad}(p)$ and $S_n = \sum_{i=1}^n \xi_i$, we are interested in the event

$$B = \{S_n = 0 \text{ i.o.}\},$$

B is not a tail event (why ?) yet it does have a zero one law.

We now define symmetric events. A permutation π is a map

$$\pi : (\pi_1, ..., \pi_n) \to (1, ..., n),$$

and $\pi(\xi) = \{\xi_{\pi_1}, ..., \xi_{\pi_n})$. A symmetric event is one for which if

$$A = \{\xi \in B\}, \quad B \in \mathcal{B}(\mathbb{R})$$

then

$$\pi(A) = \{\pi(\xi) \in B\}, \quad B \in \mathcal{B}(\mathbb{R})$$

for all permutations π. An even A is symmetric if $\pi(A) = A$ for all finite permutations π.