Class Notes| Week 2

MSO: Introduction to Probability Theory
Fall 2024

Contents

2 Conditional Probability and Functions on Sample Spaces

2.1 Bayes Theorem

Definition 1 (Conditional Probability)**.** Let $(\Omega, \{, \mathbb{P})$ be the probability space and let A be an event such that $\mathbb{P}(A) > 0$. For any event B, we define:

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \bigcap B)}{\mathbb{P}(A)} \tag{1}$$

to be the conditional probability of B given A has already occured.

Proposition 2.1. $(\Omega, \mathcal{F}, \mathbb{P}(\cdot\,|A))$ is a Probability Space

Proof.

$$\mathbb{P}(\Omega|A) = \frac{\mathbb{P}(\Omega \bigcap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A)}{\mathbb{P}(A)} = 1$$

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(B \bigcap A)}{\mathbb{P}(A)} > 0$$

Now we establish countable additivity.
If $\{E_n\}_n$ is a sequence of mutually exclusive events, then so are $\{E_n \bigcap A\}_n$. Now apply the countable additivity of mutually exclusive events on these events to get the result. QED

Proposition 2.2 (**Multiplication Rule**)**.** Let $(E_1, E_2, E_3, ..., E_n)$ be events in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{P}(E_1 \cap E_2 \cap E_3 \cap ... \cap E_n) \neq 0$, then:

$$\mathbb{P}(E_1 \cap E_2 \cap E_3 \cap ... \cap E_n) = \mathbb{P}(E_1)\cdot \mathbb{P}(E_2|E_1)\cdot \mathbb{P}(E_3|E_1 \cap E_2)\cdot ...\cdot \mathbb{P}(E_n|\bigcap_{i=1}^{n-1} E_i) \tag{2}$$

Proof.

$$\begin{aligned}
\mathbb{P}(E_1 \cap E_2 \cap E_3 \cap ... \cap E_n) &= \mathbb{P}(E_n|\bigcap_{i=1}^{n-1} E_i)\cdot \mathbb{P}(\bigcap_{i=1}^{n-1} E_i) \\
&= \mathbb{P}(E_n|\bigcap_{i=1}^{n-1} E_i)\mathbb{P}(E_{n-1}|\bigcap_{i=1}^{n-2} E_i)\cdot \mathbb{P}(\bigcap_{i=1}^{n-2} E_i) \\
&= ... \\
&= \mathbb{P}(E_1)\cdot \mathbb{P}(E_2|E_1)\cdot \mathbb{P}(E_3|E_1 \cap E_2)\cdot ...\cdot \mathbb{P}(E_n|\bigcap_{i=1}^{n-1} E_i)
\end{aligned}$$

QED

Definition 2 (**Exhaustive Events**)**.** Let $\mathcal{I}$ be an Indexing Set. The collection of events $E_i|i \in \mathcal{I}$ is said to be exhaustive if $\bigcup_{i\in\mathcal{I}} = \Omega$

Theorem 2.1 (**Theorem of Total Probability**)**.** Let $\mathcal{I}$ be a finite or countably infinite indexing set. Let the events $E_i|i \in \mathcal{I}$ be mutually exclusive and exhaustive. Then for any event E:

$$\mathbb{P}(E) = \sum_{i\in\mathcal{I}} \mathbb{P}(E \cap E_i) = \sum_{i\in\mathcal{I}} \mathbb{P}(E|E_i)\cdot \mathbb{P}(E_i)$$

Theorem 2.2 (**Bayes Theorem**)**.** Let $\mathcal{I}$ be a finite or countable infinite Indexing Set. Let $E_i|i \in \mathcal{I}$ be a collection of mutually exclusiev exhaustive events in a probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$, such that $\mathbb{P}(E_i) > 0$, $\forall i$. Then for any event $E \in \mathcal{F}$ with $\mathbb{P}(E) > 0$, we have:

$$\mathbb{P}(E_j|E) = \frac{\mathbb{P}(E_j)\cdot \mathbb{P}(E|E_j)}{\sum_{i\in\mathcal{I}} \mathbb{P}(E|E_i)\cdot \mathbb{P}(E_i)}$$

Definition 3 (Prior and Posterior Probabilities)**.** In context of Bayes Theorem, $\mathbb{P}(E_i)$ shall be referred to as prior probabilities and $\mathbb{P}(E_i|E)$ as posterior probabilities.

2.2 Independence of Events

Definition 4 (Independence of Two Events)**.** Two events belonging to the same probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$ are said to be independent if:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\cdot \mathbb{P}(A)$$

Note 1 (*Independence is different from Mutual Exclusiveness*)**.** If A and B are disjoint (or mutually exclusive), then $\mathbb{P}(A \cap B) = 0$, but independence implies $\mathbb{P}(A \cap B) = \mathbb{P}(A)\cdot \mathbb{P}(B)$

Definition 5 (Mutual Independence of a Collection of Events)**. (a)** Let $\{E_1, E_2, ..., E_n\}$ be a finite collection of events in the probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$. We say that the collection of events are mutually exclusive, if for all $k \in \{1, 2..., n\}$ and indices $1 \leq i_1 < i_2 < ... < i_k \leq n$, we have:

$$\mathbb{P}(\bigcap_{j=1}^{k} E_{i_j}) = \prod_{j=1}^{k} \mathbb{P}(E_{i_j})$$

(b) Let $\mathcal{I}$ be an indexing set and let $\{E_i | i \in \mathcal{I}$ be a collection of events in a probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$. We say that this collection of events is mutually independent or equivalently, the events are mutually independent, if all finite subcollections $\{E_{i_1}, E_{i_2}, ..., E_{i_k}\}$ are mutually independent.

Note 2. (a) To check if a collection of n events are mutually independent, we need to check for $\binom{n}{1} + \binom{n}{2} + ... + \binom{n}{n} = 2^n - n - 1$ conditions.
(b) To check if a collection of events are pairwise independent, we need to check for $\binom{n}{2}$ conditions.

2.3 Functions defined on Sample Spaces

Definition 6 (Functions defined on Sample Spaces)**.** Sometimes, we assign numerical value to each event in a sample space Ω of some random experiment $\mathcal{E}$. The domain of this function X is the sample space Ω and the range is a continuous or discrete set of numbers.

Definition 7 (Pre-Image of a set under a function)**.** The Ω be some non-empty set and let $X : \Omega \to \mathbb{R}$ be a function on this set. Given any subset A of $\mathbb{R}$, we consider the subset $X^{-1}(A)$ of Ω defined by:

$$X^{-1}(A) := \{\omega \in \Omega : X(\omega) \in A\}$$

Then the set $X^{-1}(A)$ shall be referred to as the pre-image of A under the function X.

This means, if A is the range of the function X, then the set defined by $X^{-1}(A)$ is the pre-image of A under the function.