Summary for Elementary Probability

SEUNGWOO HAN

Contents

Chapter 1

Basic Concepts

1.1 Events and Probability

Definition 1.1.1: Probability Space

A *probability space* contains of a triple $(\Omega, \mathcal{F}, P)$ where

- Ω is the sample space,
- $\mathcal{F} \subseteq 2^{\Omega}$ (each $A \in \mathcal{F}$ is called an *event*), and
- $P : \mathcal{F} \to [0, 1]$ maps each event $A \in \mathcal{F}$ to the *probability* of A

which satisfies the following conditions:

Axioms Relative to the Events The family $\mathcal{F}$ of events must be a σ-field on Ω:

(1) $\Omega \in \mathcal{F}$;
(2) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$ (where A^c is the complement of A);
(3) If $\langle A_n \rangle_{n \in \mathbb{Z}_+}$ is a sequence on $\mathcal{F}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Axioms Relative to the Probability The function P must satisfy the following conditions:

(1) $P(\Omega) = 1$;
(2) σ-additivity holds: if $\langle A_n \rangle_{n \in \mathbb{Z}_+}$ is a sequence of pairwise disjoint events, then

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

Note

Here are immediate properties of probability:

- $P(A^c) = 1 - P(A)$;
- $\varnothing = \Omega^c \in \mathcal{F}$ and $P(\varnothing) = 0$;
- If $\langle A_n \rangle_{n \in \mathbb{Z}_+}$ is a sequence of events, then $\bigcap_{n=1}^{\infty} A_n$ is also an event;
- $A, B \in \mathcal{F}$ and $A \subseteq B$ implies $P(A) \le P(B)$.

Lemma 1.1.2 sub-σ-additivity

If $\langle A_n \rangle_{n\in\mathbb{Z}_+}$ is a sequence of events, then

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) \le \sum_{n=1}^{\infty} P(A_n).$$

Proof. Let $B_n = A_n \setminus \bigcup_{i=1}^{n-1} A_i$ for each $n \ge 1$ and use σ-additivity. ☐

Lemma 1.1.3 Inclusion-Exclusion Principle

If $A_1, \cdots, A_n$ are events, then

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{\varnothing \neq I \subseteq [n]} (-1)^{|I|-1} P\left(\bigcap_{i\in I} A_i\right).$$

Proof. Classic. ☐

Theorem 1.1.4 Sequential Continuity of Probability

(1) Let $\langle B_n \rangle_{n\in\mathbb{Z}_+}$ be a sequence of events such that $B_n \subseteq B_{n+1}$ for all $n \ge 1$. Then,

$$P\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} P(B_n).$$

(2) Let $\langle C_n \rangle_{n\in\mathbb{Z}_+}$ be a sequence of events such that $C_n \supseteq C_{n+1}$ for all $n \ge 1$. Then,

$$P\left(\bigcap_{n=1}^{\infty} C_n\right) = \lim_{n\to\infty} P(C_n).$$

Proof.

(1) Let $B'_n := B_n \setminus B_{n-1}$ for each $n \ge 2$ and $B'_1 := B_1$. so that $B_m = \bigcup_{n=1}^{m} B'_n$ and B'_i's are pairwise disjoint. Hence, by σ-additivity, we have

$$P\left(\bigcup_{n=1}^{\infty} B_n\right) = P\left(\bigcup_{n=1}^{\infty} B'_n\right) = \sum_{n=1}^{\infty} P(B'_n) = P(B_1) + \sum_{n=1}^{\infty}\big(P(B_n) - P(B_{n-1})\big) = \lim_{n\to\infty} P(B_n).$$

(2) Let $C'_n := C_n{}^c$ for each $n \ge 1$ so that $C'_n \subseteq C'_{n+1}$ for all n. Hence, by (1), we have $P\left(\bigcup_{n=1}^{\infty} C'_n\right) = \lim_{n\to\infty} P(C'_n)$. The result follows from the fact that $\bigcup_{n=1}^{\infty} C'_n = \Omega \setminus \bigcap_{n=1}^{\infty} C_n$. ☐

1.2 Random Variables and Their Distributions

Definition 1.2.1: Random Variable

A *random variable* on $(\Omega, \mathcal{F})$ is any mapping $X : \Omega \to \overline{\mathbb{R}}$ such that for all $a \in \mathbb{R}$, $\{X \le a\} \triangleq \{\omega \in \Omega \mid X(\omega) \le a\} \in \mathcal{F}$. Here, $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$.

- If X only takes finite values, X is called a *real random variable*.
- If X only takes only a countable set of values $\{a_n\}_{n\in\mathbb{Z}_{\ge 0}}$, X is called a *discrete random variable*.

Definition 1.2.2: Cumulative Distribution Function

The *cumulative distribution function* (CDF) of a random variable X is the function $F: \mathbb{R} \to [0,1]$ defined by

$$F(x) = P(X \leq x) \triangleq P(\{X \leq x\}).$$

Lemma 1.2.3

Let F be a cumulative distribution function of a random variable X.

(1) F is monotone increasing.

(2) F is right-continuous.

(3) If we define $F(\infty) := \lim_{x\to\infty} F(x)$ and $F(-\infty) = \lim_{x\to-\infty} F(x)$, then $1 - F(\infty) = P(X = \infty)$ and $F(-\infty) = P(X = -\infty)$.

Proof.

(1) Take any $x, y \in \mathbb{R}$ with $x \leq y$. Then, $\{X \leq x\} \subseteq \{X \leq y\}$. Hence, $F(x) = P(X \leq x) \leq P(X \leq y) \leq F(y)$.

(2) Take any decreasing nonnegative sequence $\langle \varepsilon_n \rangle_{n\in\mathbb{Z}_+}$ of real numbers converging to zero and a real number x. Let $C_n := \{X \leq x + \varepsilon_n\}$ so that $\langle C_n \rangle_{n\in\mathbb{Z}_+}$ is a decreasing sequence of events. Note also that $\{X \leq x\} = \bigcap_{n=1}^{\infty} C_n$ Then, by Theorem 1.1.4 (2),

$$F(x) = P(X \leq x) = \lim_{n\to\infty} P(X \leq x + \varepsilon_n) = \lim_{n\to\infty} F(x + \varepsilon_n).$$

(3) Let $B_n := \{X \leq n\}$ for each $n \in \mathbb{Z}_+$ so that $\bigcup_{n=1}^{\infty} B_n = \{X < \infty\}$ and $\langle B_n \rangle_{n\in\mathbb{Z}_+}$ is an increasing sequence of events. By Theorem 1.1.4 (1),

$$1 - P(X = \infty) = P(X < \infty) = P\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} P(B_n) = \lim_{n\to\infty} F(n) = F(\infty).$$

The last equality is due to (1). $\square$

Definition 1.2.4: Probability Density

If a real random variable X admits a cumulative distribution function F such that

$$F(x) = \int_{-\infty}^{x} f(y)\,\mathrm{d}y$$

for some nonnegative function f, then X is said to admit the *probability density* f.

Note

Note that the probability density f satisfies

$$\int_{-\infty}^{\infty} f(y)\,\mathrm{d}y = 1.$$

1.3 Conditional Probability and Independence

Definition 1.3.1: Conditional Probability

Let B be an event with $P(B) > 0$. For any event A, we define

$$P(A \mid B) := \frac{P(A \cap B)}{P(B)}$$

and it is called the *probability of A given B*.

Definition 1.3.2: Independent Events

(1) Two events A and B are said to be *indepenent* if $P(A \cap B) = P(A)P(B)$.

(2) Let $\mathcal{A}$ be a nonempty family of events. $\mathcal{A}$ is said to be a *family of independent events* if for any finite subfamily $\langle A_1, \cdots, A_n \rangle$ of $\mathcal{A}$,

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i).$$

Note

When $P(B) > 0$, A and B are indepenent if and only if $P(A \mid B) = P(A)$.

End.