CARDIFF UNIVERSITY SCHOOL OF MATHEMATICS

MA1500

INTRODUCTION TO PROBABILITY THEORY

LECTURE NOTES

2015-16

Version: 18/5/2015

Contents

Lecture 1 Events

1.1 A brief history of probability

Games of chance have been played since antiquity, but the mathematical principles of chance and uncertainty were first established only in the 17th century.

1654	Classical principles	Blaise Pascal (1623–1662)
		Pierre de Fermat (1601–1665)
1657	*De Ratiociniis in Ludo Aleae*	Christiaan Huygens (1629–1695)
1713	*Ars Conjectandi*	Jakob Bernoulli (1654–1705)
1718	*The Doctrine of Chances*	Abraham de Moivre (1667–1754)
1812	*Theorie Analytique des Probabilites*	Pierre de Laplace (1749-1827)
1919	Relative frequency	Richard von Mises (1883–1953)
1933	Modern axiomatic theory	Andrey Kolmogorov (1903–1987)

1.2 Sample spaces

Definition 1.1

(1) Any process of observation or measurement will be called an *experiment* or *trial*.

(2) Any experiment whose outcome is uncertain is called a *random experiment*.

(3) A random experiment has a set of possible *outcomes*.

(4) Each time a random experiment is performed, *exactly one* of its outcomes will occur.

(5) The set of all possible outcomes is called the *sample space* of the experiment, denoted by Ω.

(6) Outcomes are also called *elementary events*, and denoted by $\omega \in \Omega$.

Example 1.2

The sample space of a random experiment is the set of all possible outcomes:

Experiment	Sample space
A coin is tossed once.	$\Omega = \{H, T\}$
A six-sided die is rolled once.	$\Omega = \{1, 2, 3, 4, 5, 6\}$
A coin is tossed repeatedly until a head occurs.	$\Omega = \{1, 2, 3, \ldots\}$
The height of a randomly chosen student is measured:	$\Omega = [0, \infty)$

Exercise 1.3

Think of a situation in which randomness occurs. Can you describe the set of possible outcomes? Can you write it down using mathematical notation?

1.3 Events

Definition 1.4

- An *event* A is a subset of the sample space, Ω.

- If outcome ω occurs, we say that event A *occurs* if and only if $\omega \in A$.
- Two events A and B with $A \cap B = \emptyset$ are called *disjoint* or *mutually exclusive*.
- The empty set $\emptyset$ is called the *impossible event*.
- The sample space Ω is called the *certain event*.

Remark 1.5

- If A occurs and $A \subseteq B$, then B must also occur.
- If A occurs and $A \cap B = \emptyset$, then B does not occur.

Example 1.6
A die is rolled once. The sample space can be represented by $\Omega = \{1, 2, 3, 4, 5, 6\}$. We may be interested in whether or not the following events occur:

Event	Subset
The outcome is the number 1.	$\{1\}$
The outcome is an even number.	$\{2, 4, 6\}$
The outcome is even but does not exceed 3.	$\{2, 4, 6\} \cap \{1, 2, 3\}$
The outcome is not even	$\Omega \setminus \{2, 4, 6\}$

1.4 Families of events

Definition 1.7
Let Ω be any set.

(1) The set of all subsets of Ω is called its *power set*, which we denote by $\mathcal{P}(\Omega)$.

(2) Any subset of $\mathcal{P}(\Omega)$ is called a *family of sets over* Ω.

Let Ω be the sample space of some random experiment. If we are interested in events A and B, we must also be interested in whether:

- event A occurs *or* event B occurs – this is the event $A \cup B$,
- event A occurs *and* event B occurs – this is the event $A \cap B$,
- event A does *not* occur – this is the event A^c.

We cannot therefore use arbitrary families of sets over Ω as the basis for investigating random experiments. Instead, we allow only families that are *closed* under certain set operations.

Definition 1.8
A family of sets $\mathcal{F}$ over Ω is said to be

(1) *closed under complementation* if $A^c \in \mathcal{F}$ for every $A \in \mathcal{F}$,

(2) *closed under pairwise unions* if $A \cup B \in \mathcal{F}$ for every $A, B \in \mathcal{F}$,

(3) *closed under finite unions* if $\bigcup_{i=1}^{n} A_i \in \mathcal{F}$ for every $A_1, A_2, \dots A_n \in \mathcal{F}$,

Definition 1.9
A family of sets $\mathcal{F}$ over Ω is called a *field of sets* over Ω if

(1) $\Omega \in \mathcal{F}$,

(2) $\mathcal{F}$ is closed under complementation, and

(3) $\mathcal{F}$ is closed under pairwise unions.

Table 1.1: Table of correspondence (Grimmett & Stirzaker 2001).

Notation	Set theory	Probability theory
Ω	Universal set	Sample space
$\omega \in \Omega$	Element of Ω	Elementary event, outcome
$A \subseteq \Omega$	Subset of Ω	Event A
$A \subseteq B$	Inclusion	If A occurs, then B occurs
$A \cup B$	Union	A or B occurs
$A \cap B$	Intersection	A and B occur
A^c	Complement of A	A does not occur
$A \setminus B$	Difference	A occurs, but B does not
$A \triangle B$	Symmetric difference	A or B occurs, but not both
$\emptyset$	Empty set	Impossible event
Ω	Universal set	Certain event

Example 1.10
A six-sided die is rolled once, and the score is observed. A suitable sample space for this experiment is the set $\Omega = \{1, 2, 3, 4, 5, 6\}$. The power set of Ω will always provide a field of sets to work with. However, suppose we are only interested in whether or not the outcome is an even number. In this case, we need only consider the following family of events:

$$\mathcal{F} = \big\{\emptyset, \{1, 3, 5\}, \{2, 4, 6\}, \{1, 2, 3, 4, 5, 6\}\big\}.$$

We can see that $\mathcal{F}$ is a field of sets over Ω, because

(1) it contains the sample space $\{1, 2, 3, 4, 5, 6\}$,

(2) the complement of every set in $\mathcal{F}$ is also contained in $\mathcal{F}$, and

(3) the union of any two sets in $\mathcal{F}$ is also contained in $\mathcal{F}$.

Theorem 1.11 (Properties of fields)
Let $\mathcal{F}$ be a field over Ω. Then

(1) $\emptyset \in \mathcal{F}$,

(2) $\mathcal{F}$ is closed under pairwise intersections,

(3) $\mathcal{F}$ is closed under set differences.

Proof

(1) We know that $\emptyset = \Omega^c$, and that $\Omega \in \mathcal{F}$. Because $\mathcal{F}$ is closed under complementation, it thus follows that $\emptyset \in \mathcal{F}$.

(2) Let $A, B \in \mathcal{F}$. By De Morgan's laws, we have that $A \cap B = (A^c \cup B^c)^c$. Because $\mathcal{F}$ is closed under complementation and pairwise unions, it thus follows that $A \cap B \in \mathcal{F}$.

(3) Let $A, B \in \mathcal{F}$. Set difference can be written as $A \setminus B = A \cap B^c$. Furthermore, by De Morgan's laws we see that $A \cap B^c = (A^c \cup B)^c$. Because $\mathcal{F}$ is closed under complementation and pairwise unions, it thus follows that $A \setminus B \in \mathcal{F}$.

1.5 Exercises

Exercise 1.12

1. Identify a sample space, and the subset corresponding to event A, in each of the following scenarios:
   (a) A coin is tossed three times. A is the event that at least two heads are obtained.
   (b) A game of football is played. A is the event that the match ends in a draw.
   (c) A couple have two children. A is the event that both are girls.
   (d) A shot hits a circular target of radius 10cm. A is the event that the shot hits within 3cm of the centre.

2. A family of sets $\mathcal{F}$ over Ω is said to be

   - *closed under finite unions* if $A_1 \cup A_2 \cup \ldots \cup A_n \in \mathcal{F}$ whenever $A_1, A_2, \ldots A_n \in \mathcal{F}$, and
   - *closed under finite intersections* if $A_1 \cap A_2 \cap \ldots \cap A_n \in \mathcal{F}$ whenever $A_1, A_2, \ldots A_n \in \mathcal{F}$.

   If $\mathcal{F}$ is a field of sets over Ω, show that

   (a) $\mathcal{F}$ is closed under finite unions, and

   (b) $\mathcal{F}$ is closed under finite intersections.

Lecture 2 Probability

2.1 Probability measures

Probability is defined to be a *function* that assigns numerical value to random events.

Definition 2.1
Let Ω be the sample space of some random experiment, and let $\mathcal{F}$ be a field of sets over Ω. A *probability measure* on $(\Omega, \mathcal{F})$ is a function

$$\begin{aligned} \mathbb{P}: \quad \mathcal{F} &\to [0,1] \\ A &\mapsto \mathbb{P}(A) \end{aligned}$$

such that $\mathbb{P}(\Omega) = 1$, and for any countable collection of pairwise disjoint events $\{A_1, A_2, \ldots\}$,

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *probability space.*

Remark 2.2

- The second property is called *countable additivity.*
- The number $\mathbb{P}(A)$ is called the *probability* of event $A \in \mathcal{F}$.

Example 2.3
Consider a random experiment in which a fair six-sided die is rolled once.

- A suitable sample space for the experiment is $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- A suitable field of events for the experiment is the power set, $\mathcal{F} = \mathcal{P}(\Omega)$.
- Because the die is fair, a suitable probability measure is given by the function

$$\begin{aligned} \mathbb{P}: \mathcal{F} &\to [0,1] \\ A &\mapsto \tfrac{1}{6}|A|, \qquad \text{where } |A| \text{ denotes the cardinality of } A. \end{aligned}$$

Event	Event	Probability
The outcome is the number 1.	$A = \{1\}$	$\mathbb{P}(A) = 1/6$
The outcome is an even number.	$A = \{2, 4, 6\}$	$\mathbb{P}(A) = 3/6$
The outcome is even but does not exceed 3.	$A = \{2, 4, 6\} \cap \{1, 2, 3\}$	$\mathbb{P}(A) = 1/6$
The outcome is not even	$A = \Omega \setminus \{2, 4, 6\}$	$\mathbb{P}(A) = 3/6$

Example 2.4
A fair six-sided die is rolled once. If we are only interested in whether the outcome is an odd or even number, we can take

- Sample space: $\Omega = \{1, 2, 3, 4, 5, 6\}$,
- Events: $\mathcal{F} = \big\{\emptyset, \{1, 3, 5\}, \{2, 4, 6\}, \{1, 2, 3, 4, 5, 6\}\big\}$
- Probability measure: $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\{1, 3, 5\}) = 1/2$, $\mathbb{P}(\{2, 4, 6\}) = 1/2$, $\mathbb{P}(\{1, 2, 3, 4, 5, 6\}) = 1$.

2.2 Properties of probability measures

Theorem 2.5 (Properties of probability measures)
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $A, B \in \mathcal{F}$.

(1) Complementarity: $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

(2) $\mathbb{P}(\emptyset) = 0$,

(3) Monotonicity: if $A \subseteq B$ then $\mathbb{P}(A) \leq \mathbb{P}(B)$.

(4) Addition rule: $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Proof

(1) Since $A \cup A^c = \Omega$ is a disjoint union and $\mathbb{P}(\Omega) = 1$, it follows by additivity that

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c).$$

(2) Since $\emptyset = \Omega^c$ and $\mathbb{P}(\Omega) = 1$, it follows by complemenarity that

$$\mathbb{P}(\emptyset) = \mathbb{P}(\Omega^c) = 1 - \mathbb{P}(\Omega) = 1 - 1 = 0.$$

(3) Let $A \subseteq B$ and let us write $B = A \cup (B \setminus A)$.
Since A and $B \setminus A$ are disjoint sets, it follows by additivity that

$$\mathbb{P}(B) = \mathbb{P}\big[A \cup (B \setminus A)\big] = \mathbb{P}(A) + \mathbb{P}(B \setminus A).$$

Hence, because $\mathbb{P}(B \setminus A) \geq 0$, it follows that $\mathbb{P}(B) \geq \mathbb{P}(A)$.

(4) Let us write:

- $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$
- $A = (A \setminus B) + (A \cap B)$
- $B = (B \setminus A) + (A \cap B)$

These are disjoint unions, so by additivity,

- $\mathbb{P}(A \cup B) = \mathbb{P}(A \setminus B) + \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B)$
- $\mathbb{P}(A) = \mathbb{P}(A \setminus B) + \mathbb{P}(A \cap B)$
- $\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B)$

Hence $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$, as required.

2.3 Exercises

Exercise 2.6

1. What does it mean to say that $\mathbb{P}$ is a probability measure over $(\Omega, \mathcal{F})$?

2. Show that $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$ for any two events A and B.

3. Let A and B be events such that $\mathbb{P}(A) = 0.4$, $\mathbb{P}(B) = 0.5$ and $\mathbb{P}(A \cup B) = 0.8$.
   Compute the following probabilities:
   (a) $\mathbb{P}(A \cap B)$.
   (b) $\mathbb{P}(A \cup B^c)$.

2.4 Assessment

Assessment 2.1
To be submitted by yesterday!

1. Let A and B be random events, with probabilities $\mathbb{P}(A) = 1/2$ and $\mathbb{P}(B) = 3/4$.
   (a) Show that $\frac{1}{4} \leq \mathbb{P}(A \cap B) \leq \frac{1}{2}$.
   (b) Show that $\frac{3}{4} \leq \mathbb{P}(A \cup B) \leq 1$.

Assessment 2.2 (Diagnositic)
Are you OK? YES/NO

Lecture 3 Conditional Probability

3.1 Conditional probability

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $A, B \in \mathcal{F}$ be any two events.

- If B occurs and $A \cap B = \emptyset$, then A cannot occur.
- If B occurs and $B \subseteq A$, then A is certain to occur.
- If B occurs, then A will also occur *if and only if* the event $A \cap B$ occurs.

Given that B occurs, the probability that A also occurs is $\mathbb{P}(A \cap B)$ expressed as a proportion of $\mathbb{P}(B)$.

Definition 3.1
If $\mathbb{P}(B) > 0$, the *conditional probability of A given B* is defined to be

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

Remark 3.2
- $\mathbb{P}(A|B) = 0$ whenever $A \cap B = \emptyset$, and
- $\mathbb{P}(A|B) = 1$ whenever $B \subseteq A$.

Example 3.3
Let A and B be two events, with probabilities $\mathbb{P}(A) = 0.3$, $\mathbb{P}(B) = 0.8$ and $\mathbb{P}(A \cap B) = 0.2$.
Find the probabilities $\mathbb{P}(A \cup B)$, $\mathbb{P}(A \cap B^c)$, $\mathbb{P}(A|B)$ and $\mathbb{P}(A|B^c)$.

Solution

(1) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = 0.3 + 0.8 - 0.2 = 0.9$

(2) $\mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = 0.3 - 0.2 = 0.1$

(3) $\mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B) = 0.2/0.8 = 0.25$

(4) $\mathbb{P}(A|B^c) = \mathbb{P}(A \cap B^c))/\mathbb{P}(B^c) = 0.1/0.2 = 0.5$

Example 3.4 (The Second Child Paradox)
If we know that a man has two children, and that one of them is a boy, what is the probability that he has two boys?

Solution Let $\Omega = \{BB, BG, GB, GG\}$ denote the sample space, and let $A = \{BB, BG, GB\}$ be the event that the man has at least one boy. Then

$$\mathbb{P}(\{BB\}|A) = \frac{\mathbb{P}(\{BB\} \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(\{BB\})}{\mathbb{P}(\{BB, BG, GB\})} = \frac{1/4}{3/4} = \frac{1}{3}.$$

3.2 The partition theorem

Definition 3.5
A *partition* of a set B is a collection of non-empty sets $\{A_1, A_2, \ldots\}$ such that every element of B lies in exactly one of these sets, or equivalently,

(1) $A_i \cap A_j = \emptyset$ for all $i \neq j$, and

(2) $B \subseteq \bigcup_{i=1}^{\infty} A_i$.

Theorem 3.6 (The Partition Theorem)
If $\{A_1, A_2, \ldots\}$ is a partition of B, then

$$\mathbb{P}(B) = \sum_{i=1}^{\infty} \mathbb{P}(B \cap A_i) = \sum_{i=1}^{\infty} \mathbb{P}(B|A_i)\mathbb{P}(A_i)$$

Proof First we write B as a disjoint union

$$B = (B \cap A_1) \cup (B \cap A_2) \cup \ldots = \bigcup_{i=1}^{\infty} (B \cap A_i)$$

By the countable additivity of probability measures,

$$\begin{aligned}\mathbb{P}(B) &= \mathbb{P}\left(\bigcup_{i=1}^{\infty}(B \cap A_i)\right) \\ &= \sum_{i=1}^{\infty} \mathbb{P}(B \cap A_i) \\ &= \sum_{i=1}^{\infty} \mathbb{P}(B|A_i)\mathbb{P}(A_i).\end{aligned}$$

3.3 Bayes' theorem

Lemma 3.7
For any two events A and B such that $\mathbb{P}(B) > 0$,

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}.$$

Proof Set intersection is a commutative operation, so

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}.$$

Theorem 3.8 (Bayes' Theorem)
Let $\{A_1, A_2, \ldots\}$ be a partition of an event B and suppose that $\mathbb{P}(B) > 0$. Then

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_{j=1}^{\infty} \mathbb{P}(B|A_j)\mathbb{P}(A_j)}$$

Proof By Lemma 3.7,

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_{j=1}^{\infty}\mathbb{P}(B|A_j)\mathbb{P}(A_j)}$$

where the last equality follows by the partition theorem.

Example 3.9
Bob tries to buy a newspaper every day. He tries in the morning with probability 1/3, in the evening with probability 1/2 and forgets completely with probability 1/6. The probability of successfully buying a newspaper in the morning is 9/10 (plenty of copies left), and in the evening is 2/10 (often sold out). If Bob buys a newspaper, what is the probability that he bought it in the morning?

Solution Let M be the event that Bob tries to buy a newspaper in the morning, E the event that he tries in the evening, and F the event that he forgets completely. Then

$$\mathbb{P}(M) = 1/3, \qquad \mathbb{P}(E) = 1/2, \qquad \mathbb{P}(F) = 1/6.$$

Let N denote the event that Bob buys a newspaper. Then

$$\mathbb{P}(N|M) = 9/10, \qquad \mathbb{P}(N|E) = 2/10, \qquad \mathbb{P}(N|F) = 0.$$

By Bayes' Theorem,

$$\begin{aligned}
\mathbb{P}(M|N) = \frac{\mathbb{P}(N|M)\mathbb{P}(M)}{\mathbb{P}(N)} &= \frac{\mathbb{P}(N|M)\mathbb{P}(M)}{\mathbb{P}(N|M)\mathbb{P}(M) + \mathbb{P}(N|E)\mathbb{P}(E) + \mathbb{P}(N|F)\mathbb{P}(F)} \\
&= \frac{9/10 \times 1/3}{(9/10 \times 1/3) + (2/10 \times 1/2) + (0 \times 1/6)} \\
&= 3/4
\end{aligned}$$

If Bob buys a newspaper, the probability that he bought it in the morning is 0.75.

3.4 Exercises

Exercise 3.10

1. Let A and B be events such that $\mathbb{P}(A) = 0.4$, $\mathbb{P}(B) = 0.5$ and $\mathbb{P}(A \cup B) = 0.8$. Compute the following probabilities:
   (a) $\mathbb{P}(A \cap B)$
   (b) $\mathbb{P}(A \cup B^c)$
   (c) $\mathbb{P}(A \mid B)$
   (d) $\mathbb{P}(A \mid A \cup B)$
2. A student has three opportunities to pass an exam. The probability of failing the first attempt is 0.6; the probability of failing the second attempt, given that they have failed the first is 0.75, and the probability of failing the third attempt, given that they have failed the first and second is 0.4.
   (a) What is the probability that the student eventually passes the exam.
   (b) What are the respective probabilities of passing at the first, second and third attempts.

3.5 Assessment

Exercise 3.11

1. Let A, B and C be events such that $\mathbb{P}(A) = 0.7$, $\mathbb{P}(B) = 0.6$, $\mathbb{P}(C) = 0.5$, $\mathbb{P}(A \cap B) = 0.4$, $\mathbb{P}(A \cap C) = 0.3$, $\mathbb{P}(B \cap C) = 0.2$ and $\mathbb{P}(A \cap B \cap C) = 0.1$. Compute the following probabilities:
   (a) $\mathbb{P}(A \cup B)$

(b) $\mathbb{P}(A|B)$

(c) $\mathbb{P}(A \mid A \cup B)$

(d) $\mathbb{P}(A \cup B \cup C)$

(e) $\mathbb{P}(A^c \cap B^c \cap C)$

(f) $\mathbb{P}(A^c \cap B^c \cap C|A \cup B)$.

2. An insurance company divides its customers into three categories: 60% of customers are classed as low-risk, 30% as moderate-risk and 10% as high-risk. The probabilities that low-risk customers, moderate-risk customers and high-risk customers make a claim in any given year are 0.01, 0.1 and 0.5 respectively. Given that a customer makes a claim this year, what is the probability that the customer is in the high-risk category?