Ch1. Probability

§1.1 Introduction

"Probability" "chance" "randomness";

Probability and Statistics;

Very useful: See the book!

§1.2 Sample Spaces

I. Random Experiments:

1. Experiments: An experiment is any situation in which we observe an outcome.

Two types of experiment:

① deterministic: the observed result is not subject to change.

Example: Measure the length of a straight wire by a ruler.

② random: the outcome is always subject to change.

Example: If a coin is tossed, then the outcome can either be "head" or "tail"

2. Random Experiments:

Random Experiments are the experiments for which the outcome cannot be predicted with certainty.

(So, at least two outcomes)

In our course, we shall only consider the random experiments and hereafter refer to "experiments" or "trials"

3. Examples:

① Example 1:

Experiment: A coin is tossed one time.

Outcomes: {Head; Tail}.

② Example 2:

Experiment: A coin is tossed three times

Outcomes: {HHH; HHT; HTH; HTT;
THH; THT; TTH; TTT}.

Note that in Ex. 2 "HHT", say, is a particular outcome.

③ Example 3:

Experiment: The number of jobs in a print queue of a mainframe computer

Outcomes: {0, 1, 2, 3, 4, 5, ...}

④ Example 4:

Experiment: The number of telephone calls received at a fixed time in . . .

Outcomes: $\{0, 1, 2, 3, 4, 5, 6, \cdots\}$

⑤ Example 5:

Experiment: the length of time between successive earthquakes in . . .

Outcomes: $\{t;\ t \geq 0\}$.

⑥ Example 6:

Experiment: the maximum temperature of a particular (coming) day.

Outcomes: might be $\{x;\ -10.5 \leq x \leq 30\}$

More convenient: $\{x;\ -\infty < x < +\infty\}$.

Note that: Examples 1 and 2: finite outcomes

Examples 3 and 4: sequence, Examples 5 and 6: . . .

II. Sample Space:

1. Definition: For any random experiment, we define the sample space to be the set of all possible outcomes of the experiment.

2. Notation: The sample space is usually denoted by Ω, and a generic element of Ω is denoted by ω.

3. Examples:

Ex 1. $\Omega = \{H; T\}$

Ex 2. $\Omega = \{HHH, HHT; HTH; HTT; THH; THT; TTH; TTT\}$

Ex 3 and 4: $\Omega = \{0, 1, 2, 3, 4, \cdots\}$

Ex 5: $\Omega = \{t;\ t \geq 0\}$

Ex 6: $\Omega = R = \{x;\ -\infty < x < +\infty\}$

Of course, sample space depends upon the experiment.

III. Events:

Recall that the sample space Ω is the set of all possible outcomes.

1. Definition: An event of the sample space Ω is a(any) subset of Ω. (See P9)
(So, any subset of Ω is an event)

2. Example: In the above example 2.

$$\Omega = \{HHH; HHT; HTH; HTT; THH; THT; TTH; TTT\}$$

Now, consider: at least two "heads" appear.
then it contains the following elements of Ω

$$\{HHH; HHT; HTH; THH\}.$$

This is a subset of Ω and thus an event. We may denote it by say

$$A = \{HHH; HHT; HTH; THH\}$$

then we say A is an event.

3. Notes:

① Since any subset of a set is itself a set, so <u>any event is a set</u>. We may thus use the notations and results in Set Theory, for example, operations. (See below!)

② We usually use capital letters to denote events, for example, A, B, etc.

③ "terminology": Suppose A is an event of Ω and we perform the random experiment and if the outcome is in A, we say

"A has occurred".

otherwise "A has not occurred"

Example.

④ Of course, "events" depend upon the sample space Ω.

4. Some special events:

① Impossible event: ϕ

Empty set ϕ is a subset of Ω and so ϕ is an event.

Here "empty" means "contains no element i.e. "it can not occur".

For example, in the above example 2, if we consider "four heads appear", then it is impossible and thus an impossible event.

We can see impossible event ϕ refers to a particular sample space.

② Certain event: Ω (or "Sure event")

Ω itself can be viewed as a subset of Ω and so Ω is also an event.

Since Ω contains all the possible outcomes and thus "Ω always occurs" and so "certain event".

③ Elementary event:

An elementary event of the sample space is a single element of Ω corresponding to a particular outcome of the experiment.

For example in the above example 2,

$\Omega = \{HHH; HHT; HTH; HTT; THH; THT; TTH; TTT\}$.

"THH", for example, is a particular outcome and thus is an elementary event.

Elementary event is also called a "point" in the sample space.

The sample space is actually the set of all elementary events, or the set of all "points".

④ Note: We can see, "impossible event ϕ", "Certain event Ω" and "elementary event" all correspond to the given sample space.

IV. Operations of Events: (∪)

Nothing but the operations of sets.

1. Union:

① Definition: Suppose A and B are two events of Ω, the union of A and B is the event that either A occurs or B occurs or both occur, denoted by $A \cup B$.

(It is enough just to say either A or B)

② Example: In the above example 2

$\Omega = \{HHH; HHT; HTH; HTT; THH; THT; TTH; TTT\}$

Suppose $A = \{HHH; HHT; HTT; HTH\}$

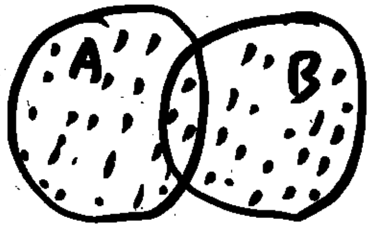
$B = \{HHH; HTH; TTH; THH\}$

then the union of A and B is the event C, where

$$C = A \cup B = \{HHH; HHT; HTT; HTH; TTH; THH\}$$

(Note that we do not write HHH and HTH two times)

③ Venn diagram:



$A \cup B$

④ Notation:

$$A \cup B = \{\omega \in \Omega;\ \ \omega \in A \text{ or } \omega \in B\}$$
$$= \{\text{either } A \text{ or } B \text{ occurs}\}$$
$$\equiv \{\text{either } A \text{ occurs or } B \text{ occurs or both}\}$$

⑤ Basic laws:

(i) Commutative law: $A \cup B = B \cup A$

(ii) Associative law: $(A \cup B) \cup C = A \cup (B \cup C)$

(iii) $A \cup \phi = A$

(iv) $A \cup \Omega = \Omega$

All these laws are easily verified.

2. Intersection:

① Definition: Suppose A and B are two events of Ω, then the intersection of A and B is the event that both A and B occur, denoted by $A \cap B$.

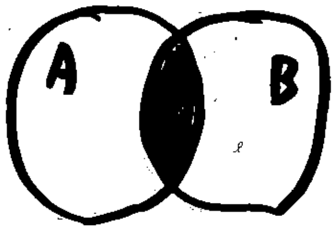
(i.e. consists of those outcomes that are common to both A and B)

(Recall and compare here that union consists of those outcomes that are in A or in B or in both).

② Example: A and B as above, then the intersection of A and B is the event C where

$$C = A \cap B = \{HHH;\ HTH\}$$

③ Venn diagram:



$A \cap B$

④ Notation:

$$A \cap B = \{ \omega \in \Omega ;\ \omega \in A \ \underline{\text{and}} \ \omega \in B \}$$
$$= \{ \text{both } A \text{ and } B \text{ occur} \}$$

⑤ Basic laws:

(i) Commutative law: $A \cap B = B \cap A$

(ii) Associative law: $(A \cap B) \cap C = A \cap (B \cap C)$

(iii) $A \cap \phi = \phi$

(iv) $A \cap \Omega = A$

⑥ Further:

(v) Distributive laws: $(A \cup B) \cap C = (A \cap C) \cup (B \cap$

(vi) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$

3. Complement:

① Definition: Suppose A is an event of Ω, then the complement of A is the event that A does not occur and thus consists of all those elements in the sample space that are not in A.

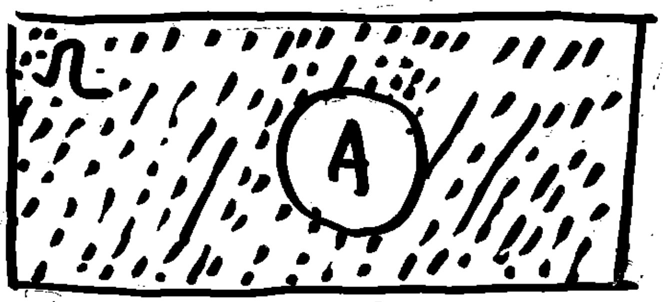
The complement of A is denoted by A^c.

② Examples: A, B as above (See P10)

then $A^c = \{THH; THT; TTH; TTT\}$

$B^c = \{HHT; HTT; THT; TTT\}$

③ Venn diagram



④ Notation:

$$A^c = \{w \in \Omega ;\ w \notin A\}$$
$$= \{A \text{ does not occur}\}$$

⑤ Basic laws:

By the above example, we notice that

$A \cup B = \{HHH; HHT; HTT; HTH; TTH; THH\}$,

$A^c = \{THH; THT; TTH; TTT\}$.

$B^c = \{HHT; HTT; THT; TTT\}$.

Thus

$(A \cup B)^c = \{THT; TTT\}$

$A^c \cap B^c = \{THT; TTT\}$

and so

$$\boxed{(A \cup B)^c = A^c \cap B^c}$$

Similarly, $A \cap B = \{HHH; HTH\}$

$\Rightarrow$: $(A \cap B)^c = \{HHT; HTT; THH; THT; TTH; TTT\}$

and $A^c \cup B^c = \{HHT; HTT; THH; THT; TTH; TTT\}$

and so

$$\boxed{(A \cap B)^c = A^c \cup B^c}$$

These are called "De-Morgan laws", i.e. for any two events A and B, we have

$(A \cup B)^c = A^c \cap B^c$
$(A \cap B)^c = A^c \cup B^c$

(Check it using the Venn diagram !!)

Also easy to see: $\phi^c = \Omega$; $\Omega^c = \phi$.

4. Remark:

Similarly, we can define the union and intersection for finitely many or even a sequence of events. The meaning should be clear.

We usually use "$\bigcup_{i=1}^{n} A_i$" "$\bigcap_{i=1}^{n} A_i$" "$\bigcap_{i=1}^{\infty} A_i$" "$\bigcup_{i=1}^{\infty} A_i$" denote these operations.

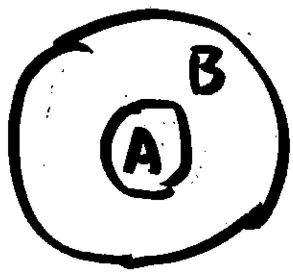
Still, De-Morgan laws apply.

V. Relations among events:

1. "Containing":

① Suppose A and B are two events, if A is a subset of B. i.e. each element of A is also an element of B, then we say B contains A. or A is included in B.

② diagram



③ Notation: $A \subset B$ or $B \supset A$.

In otherwords. $A \subset B$ means:

it A occurs then B must occur.

or " A occurs implies B occurs "

④ Example: In the above example 2.

let $E = \{HHH; HTH\}$

$G = \{HHH; HHT; HTH; TTT\}$.

Then $E \subset G$.

If "E occurs", then "G occurs".

Note that for any event A, we have

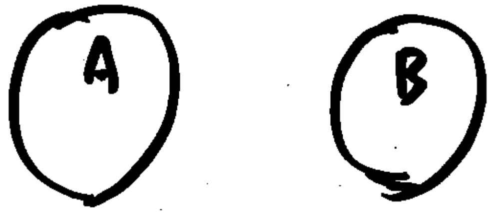
$$\phi \subset A \subset \Omega.$$

2. Disjoint:

① Two events A and B are said to be disjoint if A and B have no outcomes in common.

In otherwords, the intersection of A and B contains no element (i.e. impossible event) or "A and B are disjoint" means "$A \cap B = \phi$"

② diagram:



$$A \cap B = \phi$$

③ Example: In the above example 2,

if $A = \{HHH; HTH\}$

$B = \{TTT; THH; TTH\}$

then $A \cap B = \phi$

and so A and B are disjoint.

④ meaning: If A and B are disjoint, then if A occurs then B can not occur

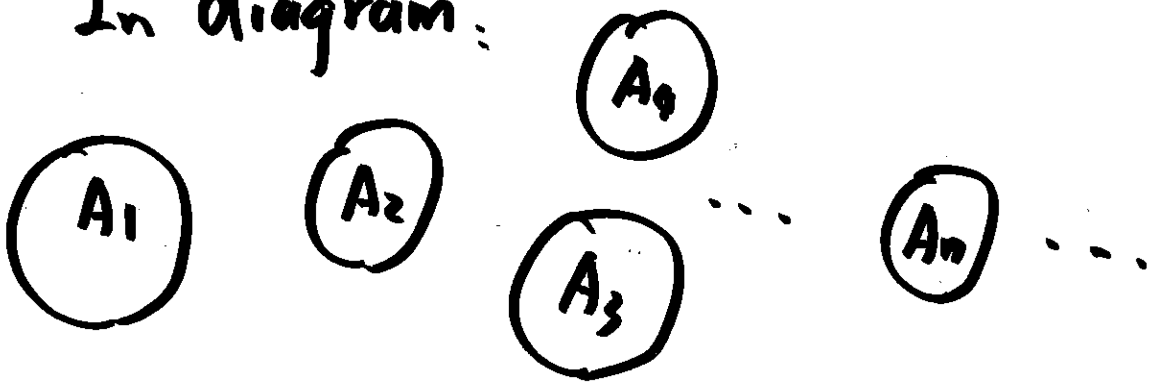
i.e. A "occurs" implies "B does not occur"

(Also, of course "B occurs" implies "A does not occur".

⑤ Remark: Similarly, we can define several events that are disjoint events. Also, the meaning of "a sequence of events are disjoint" should be clear.

Formally, a sequence of events $A_1, A_2, \cdots$ are called mutually disjoint if any two of them are disjoint.

In diagram:



In other words, a sequence of events

$$A_1, A_2, A_3, \cdots A_n, \cdots$$

are called mutually disjoint, if

for any $i \neq j$, $A_i \cap A_j = \phi$.

(Similary, for finitely many of events $A_1, A_2, \cdots, A_n$.)

§1.3 Probability Measures

I. Definition: (T)

1. Def.: For a given sample space Ω, the probability measure (or, simply, "probability") is a function $P(\cdot)$ from the event to the real numbers that satisfies the following axioms: (See P11)

(a) $P(\Omega) = 1$;

(b) for any event A, $P(A) \geq 0$;

(c) if A and B are disjoint, then

$$P(A \cup B) = P(A) + P(B). \quad (1.3.1)$$

More generally, if $A_1, A_2, \cdots A_n, \cdots$ are mutually disjoint, then

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) \quad (1.3.2)$$

Also true for finitely many of disjoint events.

2. Explanations:

① Essentially, probability is a function.
But what function?
The meaning is: for any event A, a real value, denoted by $P(A)$, is assigned. Thus, for this function,

Domain: events
Range: real values $\mathbb{R}$.
] set function

Hence Probability P: events $\to \mathbb{R}$.

② For any event A, $P(A)$ is not an event, it is a real value (and actually non-negative). $P(A)$ represents the possibility of the occurance of A. (Chance of A!!)

So, A is an event (not a number usually) but $P(A)$ is a number (not an event in general.

③ The probability, i.e. the set function

$$P(\cdot) : \text{events} \to R$$

must satisfy conditions (a)–(c).

They are axioms (!!!)

but, of course, reasonably in the meaning of "agreement with the intuition".

Condition (a): Certain event Ω consists of all possible outcomes and thus must occur. hence, the possibility is 100%, i.e. 1.

Condition (b): "Possibility" must be non-negative.

Condition (c): If two events A and B are disjoint, then the "possibility" of "either A or B occurs" equals to the sum of the possibilities of A and B.

Also, it should be true even for a sequence of disjoint events.

④ meaning of (1.3.2):

left hand side: $\bigcup_{n=1}^{\infty} A_n$ is also an event and so has a probability (real number)

right hand side: a series.

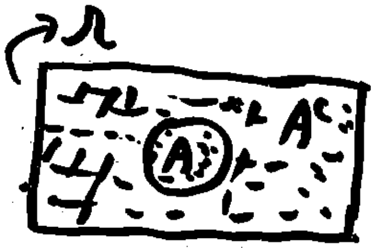
⑤ "Probability measure" or "Probability" refers to three objects: Sample space; events and the set function $P(\cdot)$.

II. Properties of Probability Measure: (Ʊ)

1. Property 1: (P15. Th.4) For any event A,

$$P(A^c) = 1 - P(A) \qquad (1.3.3)$$

Proof: Easy to see



$$A \cup A^c = \Omega, \quad A \cap A^c = \phi$$

$\Rightarrow$: $P(\Omega) = P(A \cup A^c) = P(A) + P(A^c)$

but $P(\Omega) = 1$ and so $1 = P(A) + P(A^c)$

$\Rightarrow$: $P(A^c) = 1 - P(A)$.

##

2. Property 2: $P(\phi) = 0$

(The probability of impossible event is zero.)

Proof: In Property 1, let $A = \phi$, then

$$P(\phi^c) = 1 - P(\phi).$$

But $\phi^c = \Omega$ and so $P(\phi^c) = P(\Omega) = 1$

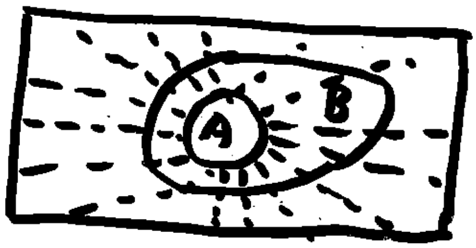
$\Rightarrow$: $1 = 1 - P(\phi)$

$\Rightarrow$: $P(\phi) = 0$. ##

3. Property 3: If $A \subset B$, then $P(A) \le P(B)$.

Proof: $\because$ $A \subset B$

we may write B as



$$B = A \cup (B \cap A^c)$$

and A and $B \cap A^c$ are disjoint. (See ↑)

Now, by axioms (c),

$$P(B) = P(A) + P(B \cap A^c)$$

But $P(B \cap A^c)$ is non-negative (axioms (b))

we get $P(B) \ge P(A)$. ##

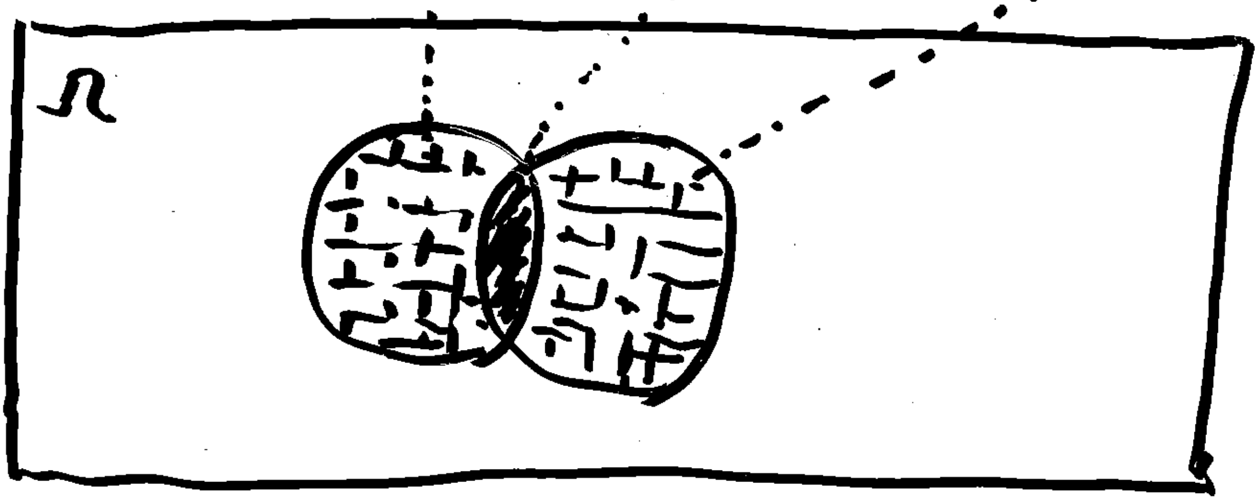
4. Property 4: (Sm P12. Th. 2)

If A and B are any two events, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (1.3.4)$$

Proof: Easy to see

$$A \cup B = (A \cap B^c) \cup (A \cap B) \cup (A^c \cap B)$$



and $A \cap B^c$, $A \cap B$ and $A^c \cap B$ are mutually disjoint

$\Rightarrow$: $$P(A \cup B) = P(A \cap B^c) + P(A \cap B) + P(A^c \cap B) \quad (1.3.5)$$

But $A = (A \cap B^c) \cup (A \cap B)$

and $A \cap B^c$ and $A \cap B$ are disjoint

So

$$P(A) = P(A \cap B^c) + P(A \cap B)$$

Similarly $B = (B \cap A) \cup (B \cap A^c)$ (disjoint!)

$\Rightarrow$ $P(B) = P(A \cap B) + P(B \cap A^c)$

Substituting the above two into (1.3.5) yields

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

i.e. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. ##

Note that, if $A \cap B = \phi$ (disjoint) then

$P(A \cap B) = P(\phi) = 0$ then we get

$P(A \cup B) = P(A) + P(B)$ if A, B disjoint.

5. Property 5

For any event A, we have $P(A) \leq 1$.

Proof: Easy. Indeed, $A \subset \Omega$ always true

thus $P(A) \leq P(\Omega)$ (Property 3)

But $P(\Omega) = 1$

$\Rightarrow$ $P(A) \leq 1$.

##

IV. Summary: (※)

Probability is a set function defined for all events that satisfies the following properties:

1. Nonnegative and bounded, i.e. for any event A,
$$0 \le P(A) \le 1$$
2. Monotone, i.e.
$$A \subset B \implies P(A) \le P(B)$$
3. Additive, i.e.
If $A_1, A_2, \cdots, A_n, \cdots$ are disjoint (finitely many or a sequence of them)
then
$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$
4. $P(\phi) = 0$, $\quad P(\Omega) = 1$
5. $P(A^c) = 1 - P(A)$
6. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$