2.1 Introduction

Terminologies of Probability

Random experiments
Outcomes

Event
Sample space

What is probability ?

概率？

2.2 Sample Space and Events

Definitions

The basic object of probability is an **experiment**: an activity or procedure that produces distinct, well-defined possibilities called **outcomes**. The **sample space** is the set of all possible outcomes of an experiment, usually denoted by S.

Example The sample space of tossing a coin:

$$S = \{\text{ head, tail }\}.$$

Example Tossing two dice:

$$\begin{aligned} S &= \{(1,1),(1,2),\ldots,(1,6),(2,1),\ldots,(6,6)\} \\ &= \{(i,j) : 1 \le i,j \le 6\}. \end{aligned}$$

2.2 Sample Space and Events

Example The lifetime of a transistor:

$$S = [0, \infty).$$

Definitions

Any subset E of the sample space is an **event**.

A sample space of a random experiment is the collection of ALL possible outcomes

An event of a random experiment is the collection of SOME possible outcomes

2.2 Sample Space and Events

Example The sample space of tossing a coin:

$$S = \{ \text{ head, tail } \}.$$

$$E = \{ \text{ head } \} \text{ is an possible event.}$$

Example Tossing two dice:

$$\begin{aligned} S &= \{(1,1),(1,2),\ldots,(1,6),(2,1),\ldots,(6,6)\} \\ &= \{(i,j) : 1 \le i,j \le 6\}. \end{aligned}$$

$$\begin{aligned} E &= \{ \text{ sum of 2 dice is 7}\} \\ &= \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\} \end{aligned}$$

is an possible event.

2.2 Sample Space and Events

Example The lifetime of a transistor:

$$S = [0, \infty).$$

$E = \{x : 0 \leq x \leq 5\}$ is an possible event.

The number of elements in the sample space and event can be finite or infinite.

If the random experiment produces an outcome in event E, we say that "event E occurs"

2.3 Operations on Events

Four basic operations

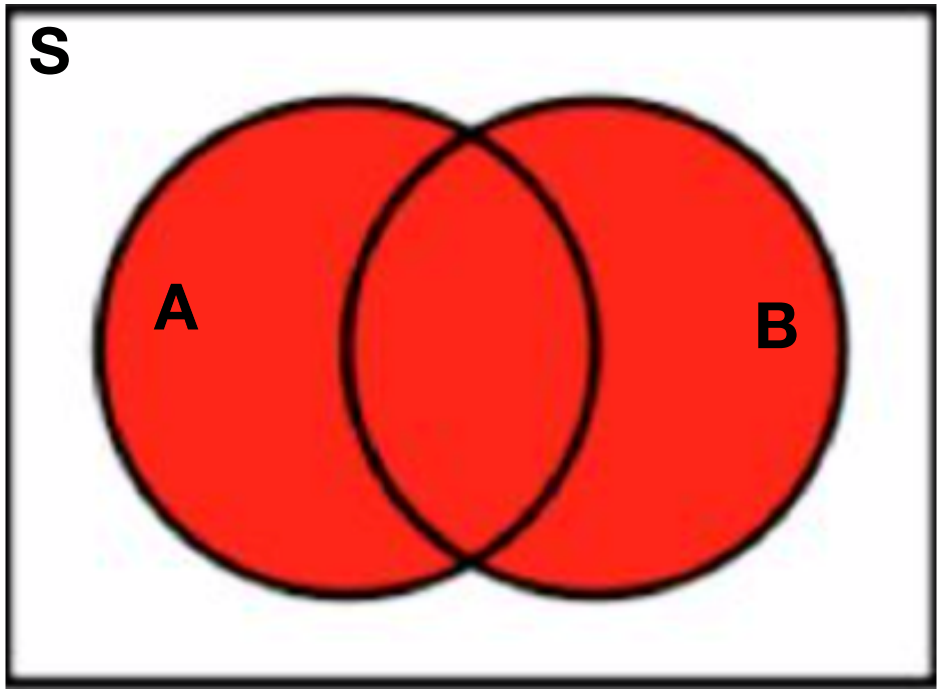
Union: A U B represents the union of event A and event B, it contains the outcomes in either A or B or both. Since AUB is also a collection of outcomes, it is also an event.

Example: A={the selected card is King}
B={the selected card is heart}
, then AUB={the selected card is either King or heart}.

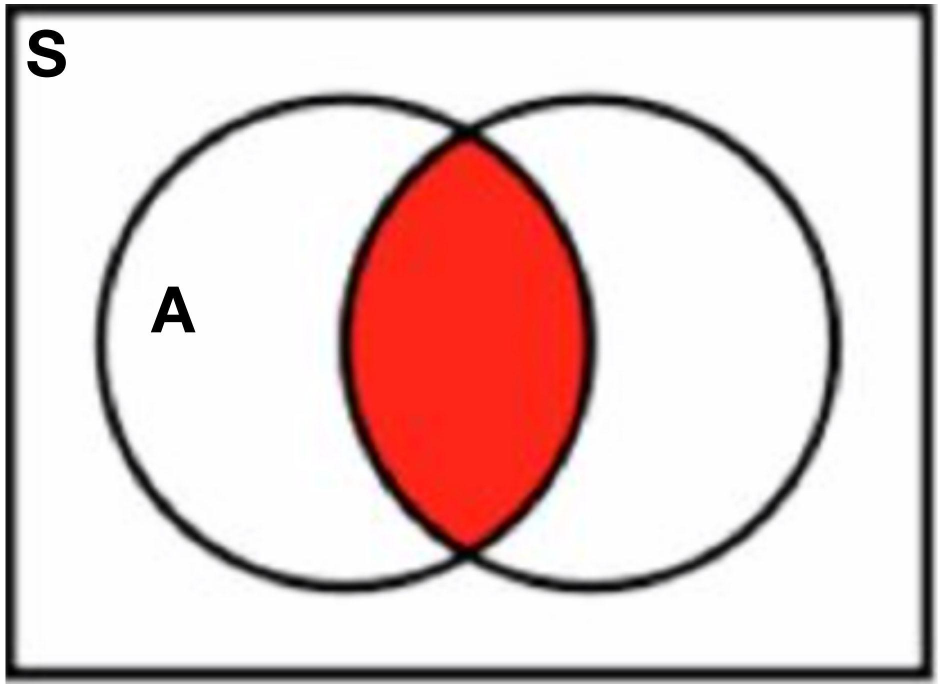
Intersection: $A \cap B$ represents the intersection of event A and event B, it contains only the outcomes in both A and B. Since $A \cap B$ is also a collection of outcomes, it is also an event.

Example: A={the selected card is King}
B={the selected card is heart}
, then $A \cap B$={the selected card is King of heart}.

Venn diagram $A \cup B$



Venn diagram $A \cap B$ *(or AB in short)*



Remember that S is the sample space

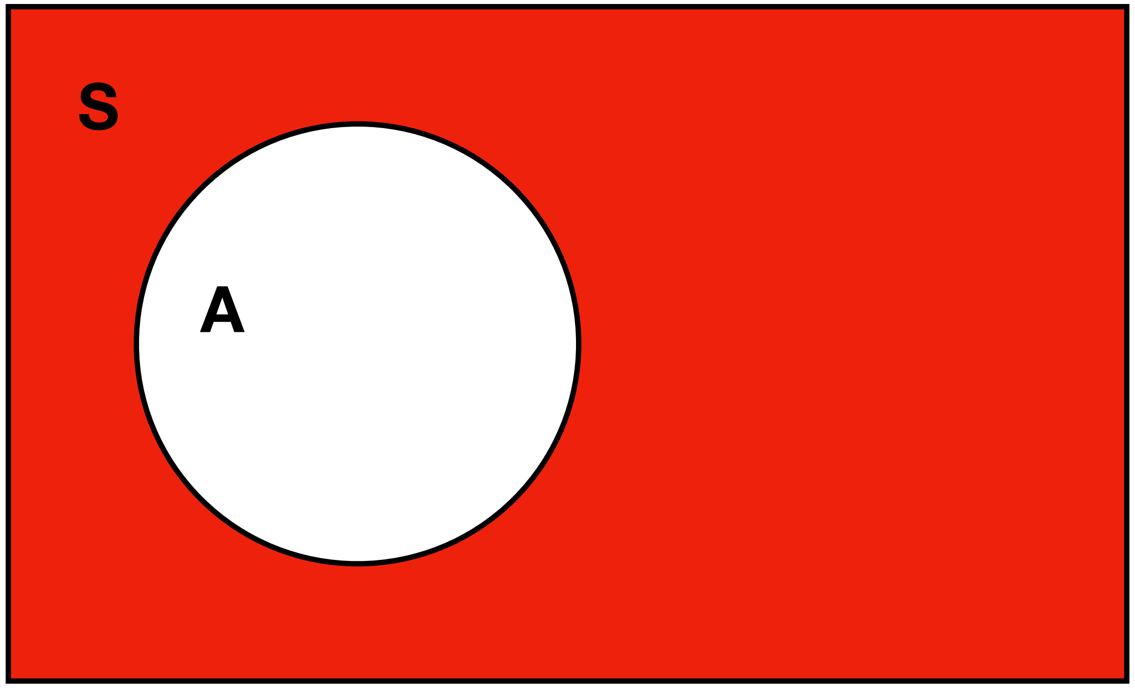
2.3 Operations on Events

Four basic operations

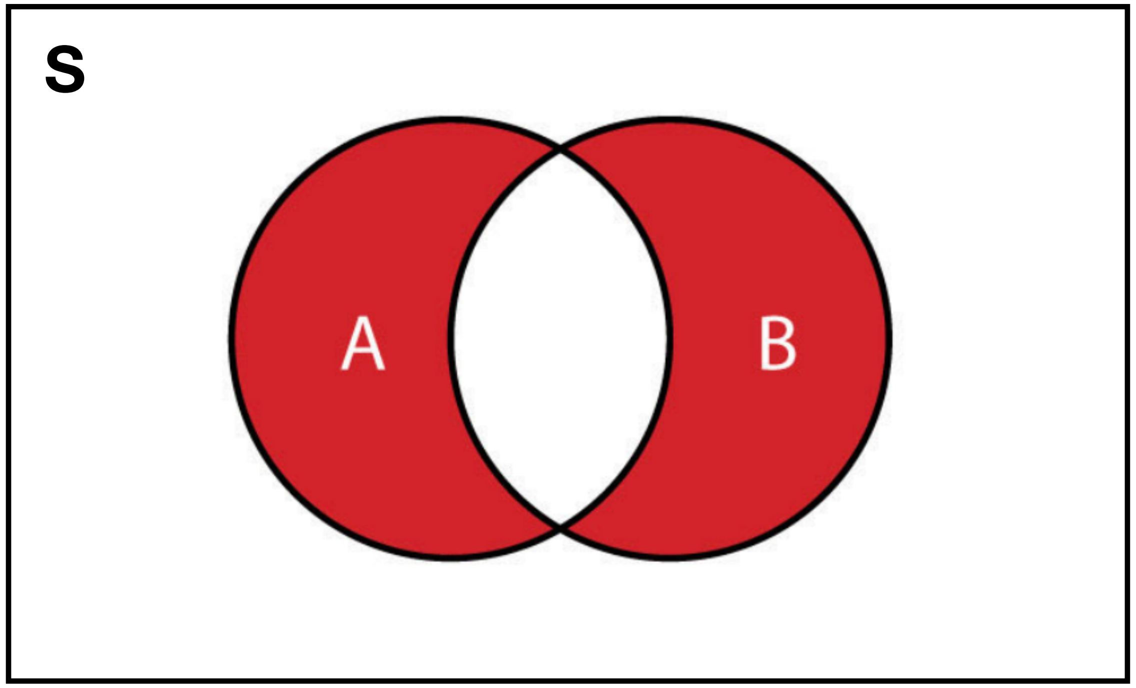
Complement: A^c represents the complement of event A , it contains the outcomes NOT in A. Since A^c is also a collection of outcomes, it is also an event.
Example: A={the selected card is King}
, then A^c={the selected card is not King}.

Symmetric difference: $A \Delta B$ represents the symmetric difference of event A and B, it contains the outcomes in either A or B, but not in both A and B.
Example: A={the selected card is King}
B={the selected card is heart}
, then $A \Delta B$ ={the selected card is either King or heart, except King of heart}.

Venn diagram A^c



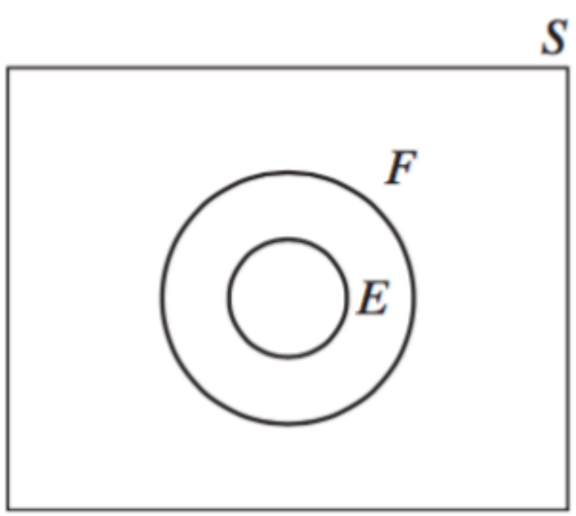
Venn diagram $A \Delta B$



Remember that S is the sample space

2.3 Operations on Events

Definitions (Inclusion of events). *For any two events E and F, if all of the outcomes in E are also in F, then we say that E is contained in F and write $E \subset F$, or $F \supset E$. Thus, if $E \subset F$, the occurrence of E necessarily implies the occurrence of F.*



Remark 2.2.7. If $E \subset F$ and $F \subset E$, we have $E = F$.

Remember that S is the sample space

2.3 Operations on Events

Additional Concepts

Disjoint: if two events share no common outcomes, that is there intersection is an empty set $\emptyset$, they are called disjoint.

Example: A={the selected card is King}, B={the selected card is Queen}.
Then, $A \cap B = \emptyset$

Mutually exclusive: if events $A_1, A_2, \cdots, A_k$ are pairwise disjoint, then they are called mutually exclusive.

Exhaustive: if the union of some events $A_1, A_2, \cdots, A_k$ is the sample space, then they are called exhaustive.

Can you represent them by Venn diagram?

2.3 Operations on Events

Fundamental Laws

Let A, B, C be any three events. Let $A_1, A_2, \cdots, A_k$ be events

Commutative law

$$1). \quad A \cap B = B \cap A, \quad A \cup B = B \cup A$$

Associative law

$$2). \quad A \cap (B \cap C) = (A \cap B) \cap C, \quad A \cup (B \cup C) = (A \cup B) \cup C$$

Distributive law

$$3). \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

DeMorgan's law

$$4). \quad \left(\cup_{i=1}^k A_i\right)^{\mathrm{c}} = \cap_{i=1}^k A_i^{\mathrm{c}}, \quad \left(\cap_{i=1}^k A_i\right)^{\mathrm{c}} = \cup_{i=1}^k A_i^{\mathrm{c}}$$

The complement of unions = intersection of the complements
The complement of intersections = the union of complements

2.3 Operations on Events

Lemma For any events $A, B,$ we have

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

Proof **by Venn diagram?**

By distributive law

$$\begin{aligned}(A \cap B) \cup (A \cap B^c) &= A \cap (B \cup B^c) \\ &= A \cap S \\ &= A\end{aligned}$$

2.4 Axioms of Probability

What is probability ?

Intuitively, a **probability** is a function that assigns numbers to events,

Example the probability of "sum of 2 dices is 7" is 1/6.

This number can characterize how likely this event occurs.

Appetizer

If I randomly roll this dice,
how likely I will get 5?
and how to describe this likelihood?

An intuitively idea is to roll this dice for 1 million times and count the number of times when 5 is observed.

Then, the relatively frequency can be regarded as a measure of how likely 5 is observed.

2.4 Axioms of Probability

Definitions (primitive definition via limiting frequency). *Suppose that an experiment, with sample space S, is repeatedly performed. For each event $E \subset S$, we define $n(E)$ to be the number of times in the first n repetitions of the experiment that the event E occurs. Then $P(E)$, the probability of the event E, is defined by*

$$P(E) = \lim_{n\to\infty} \frac{n(E)}{n}.$$

That is, $P(E)$ is defined as the (limiting) proportion of time that E occurs.

2.4 Axioms of Probability

The issues of primitive definitions

However, the above definition has some serious drawback:

1. How do we know if the limit of $\frac{n(E)}{n}$ exists or not for a sequence of repetitions of the experiment?

2. Even if the limits exist for all sequences, how do we know that the limits are the same?

For a mathematical probability model, we can certainly make an axiom to assume that $\frac{n(E)}{n}$ will converge to the same constant value. However, this statement seems too complicated to be an axiom.

Modern probability theory is built on Kolmogorov Axiom

2.4 Axioms of Probability

Kolmogorov Axioms

Probability, denoted by P, is a function on the collection of events satisfying

(i) For any event A,

$$0 \leq P(A) \leq 1.$$

(ii) Let S be the sample space, then

$$P(S) = 1.$$

(iii) For any sequence of mutually exclusive events $A_1, A_2, \ldots$ (that is, $A_i \cap A_j$ $\emptyset$ when $i \neq j$),

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

2.4 Axioms of Probability

Remarks

We call $P(A)$ the **probability** of the event A.

Axiom 1 states that the probability that the outcome is in E is some number between 0 and 1. Axiom 2 states that, with probability 1, the outcome will be a point in the sample space S. Axiom 3 states that for any sequence of mutually exclusive events the probability of at least one of these events occurring is just the sum their respective probabilities.

2.4 Axioms of Probability

Example

We roll a fair six-faced dice randomly and we consider the number of the random outcome.
Let A={1,3} and B={2,4} be two events.
By Kolmogorov's axiom of probabilities, what are P(A) and P(B) ?

There are only 6 possible outcomes: S={1,2,3,4,5,6}.
We only need to decide the probabilities P(1), P(2), P(3), P(4), P(5), P(6).

Since the dice is fair, we know that P(1)=P(2)=P(3)=P(4)=P(5)=P(6). Meanwhile, P(S)=1 we get
P(S)=P(1)+P(2)+P(3)+P(4)+P(5)+P(6)=1.
Therefore, we get P(1)=P(2)=…=P(6)= 1/6.
Then, P(A)=P(1)+P(3)= 1/3 and P(B)=P(2)+P(4)= 1/3.

2.5 Properties of Probability

Theorem

$$P(\emptyset) = 0.$$

Proof

Proof. Take $A_1 = S$, $A_2 = A_3 = \cdots = \emptyset$, then $A_1, A_2, \ldots$ are mutually exclusive(why?). Axiom (iii) says that

$$P(S) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = P(S) + \sum_{i=2}^{\infty} P(\emptyset),$$

which implies that $P(\emptyset) = 0$. $\square$

2.5 Properties of Probability

Theorem

For any finite sequence of mutually exclusive events $A_1, A_2, \ldots, A_n$,

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

Proof

Proof. Let $A_{n+1} = A_{n+2} = \cdots = \emptyset$, then the sequence $A_1, A_2 \ldots$ is still mutually exclusive, and hence

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} A_i\right) &= P\left(\bigcup_{i=1}^{\infty} A_i\right) \\ &= \sum_{i=1}^{\infty} P(A_i) \\ &= \sum_{i=1}^{n} P(A_i) \quad \text{since } P(A_i) = 0, i \geq n+1. \end{aligned}$$

2.5 Properties of Probability

Theorem

Let A be an event, then

$$P(A^c) = 1 - P(A).$$

Proof

Proof. Since $S = A \cup A^c$, it follows that

$$1 = P(S) = P(A \cup A^c) = P(A) + P(A^c).$$

Rearranging the terms, we proved the result.

2.5 Properties of Probability

Theorem

If $E \subset F$, *then*

$$P(E) \leq P(F).$$

Proof

Proof. Since $E \subset F$, it follows that we can express F as

$$F = E \cup E^c F.$$

Hence, as E and $E^c F$ are mutually exclusive, we obtain from Axiom 3 that

$$P(F) = P(E) + P(E^c F)$$

which proves the result since $P(E^c F) \geq 0$ by Axiom 1.

2.5 Properties of Probability

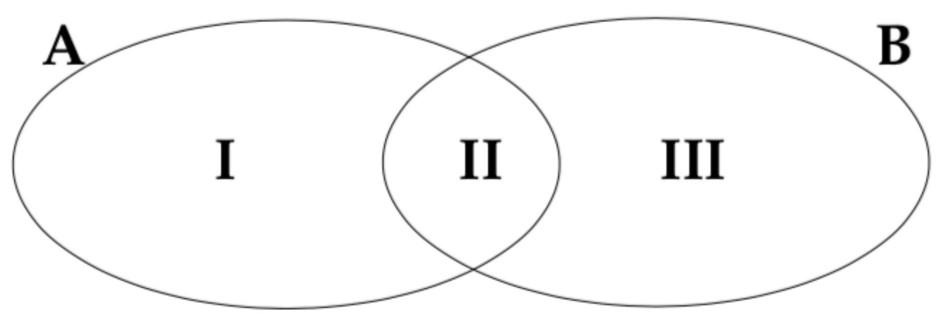
Theorem

Let A and B be any two events, then

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

Proof

Proof. Decompose $A \cup B$ into 3 mutually exclusive regions as shown:



$$\begin{aligned} P(A) &= P(I) + P(II) \\ P(B) &= P(II) + P(III) \end{aligned} \qquad \begin{aligned} P(AB) &= P(II) \\ P(A \cup B) &= P(I) + P(II) + P(III). \end{aligned}$$

proves claim

2.5 Properties of Probability

Example

Let A and B be two events such that $P(B) = \frac{5}{8}$ and $P(A \cap B) = \frac{1}{2}$. Find $P(B \cap A^c) =?$

Solution. Recall the lemma $B = (B \cap A) \cup (B \cap A^c)$

Since $(B \cap A)$ is disjoint with $(B \cap A^c)$, we get

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

By property 3) of Kolmogorov's Axiom

Therefore,

$$\begin{aligned} P(B \cap A^c) &= P(B) - P(B \cap A) \\ &= \frac{5}{8} - \frac{1}{2} = \frac{1}{8} \end{aligned}$$

2.5 Properties of Probability

Example J is taking two books along on her holiday vacation. With probability 0.5, she will like the first book; with probability 0.4, she will like the second book; and with probability 0.3, she will like both books. What is the probability that she likes neither book?

Solution.

Let B_i denote the event that J likes book i, where $i = 1, 2$. Then the probability that she likes at least one of the books is

$$P(B_1 \cup B_2) = P(B_1) + P(B_2) - P(B_1 B_2) = 0.5 + 0.4 - 0.3 = 0.6.$$

The event that she likes neither book is represented by $B_1^C B_2^C$, and its probability is given as

$$P\left(B_1^C B_2^C\right) = P\left((B_1 \cup B_2)^C\right) = 1 - P(B_1 \cup B_2) = 0.4.$$

2.5 Properties of Probability

Theorem (Inclusion-Exclusion Principle). *Let $A_1, A_2, \ldots, A_n$ be any events, then*

$$
\begin{aligned}
P(A_1 \cup A_2 \cup \cdots \cup A_n) &= \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i_1 < i_2 \le n} P(A_{i_1} A_{i_2}) + \cdots \\
&+ (-1)^{r+1} \sum_{1 \le i_1 < \cdots < i_r \le n} P(A_{i_1} \cdots A_{i_r}) \\
&+ \cdots + (-1)^{n+1} P(A_1 \cdots A_n).
\end{aligned}
$$

If n=4

$\sum_{1 \le i_1 < i_2 \le 4} P(A_{i_1} A_{i_2})$ means

$$
\begin{aligned}
&P(A_1A_2) + P(A_1A_3) + P(A_1A_4) + P(A_2A_3) \\
&+ P(A_2A_4) + P(A_3A_4).
\end{aligned}
$$

$\sum_{1 \le i_1 < i_2 < i_3 \le 4} P(A_{i_1} A_{i_2} A_{i_3})$ means

$$
P(A_1A_2A_3) + P(A_1A_2A_4) + P(A_1A_3A_4) + P(A_2A_3A_4).
$$

2.6 Sample Spaces Having Equally Likely Outcomes

For many experiments, it is natural to assume that all outcomes in the sample space (of finite number of elements) are equally likely to occur.

Example

- tossing a fair coin: $P(\{\text{ head }\}) = P(\{\text{ tail }\})$
- tossing a pair of fair dice: $P(\{(1,1)\}) = P(\{(1,2)\}) = \cdots.$

Write $S = \{s_1, s_2, \ldots s_N\}$ where N denotes the number of outcomes of S. (We will use $|S|$ to denote the number of outcomes of S.) Since outcomes are assumed to be equally likely to occur, write $P(\{s_i\}) = c$, for $i = 1, 2, \ldots, N$. As

$$1 \;=\; P(S) \;=\; P(\cup_{i=1}^{N}\{s_i\}) \;=\; \sum_{i=1}^{N} P(\{s_i\}) \;=\; Nc,$$

we get $c = 1/N$. In other words, $P(\{s_i\}) = \dfrac{1}{|S|}$.

2.6 Sample Spaces Having Equally Likely Outcomes

Then, it follows from Axiom 3 that for any event $E \subset S$,

$$P(E) = \frac{\text{number of outcomes in } E}{\text{number of outcomes in } S}.$$

In words, if we assume that all outcomes of an experiment are equally likely to occur, then the probability of any event E equals the proportion of outcomes in the sample space that are contained in E.

2.6 Sample Spaces Having Equally Likely Outcomes

Example

We select a card randomly from a standard deck of 52 playing cards.
Let A={the selected card is diamond},
B={the selected card is Jacket}.
Find P(A) and P(B)?

Solution. The sample S contains all the 52 possible cards

Then,

$$P(A) = \frac{13}{52} = \frac{1}{4}$$

$$P(B) = \frac{4}{52} = \frac{1}{13}$$

2.6 Sample Spaces Having Equally Likely Outcomes

Example A pair of fair dice is tossed. What is the probability of getting a sum of 7?

Solution.

As the dice are fair, we assume all outcomes are equally likely. So

$$\begin{aligned} A &= \{\text{ sum is } 7\} \\ &= \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\} \end{aligned}$$

therefore,

$$P(A) = \frac{|A|}{|S|} = \frac{6}{36} = \frac{1}{6}.$$

2.6 Sample Spaces Having Equally Likely Outcomes

Example If 3 balls are "randomly drawn" from an urn containing 6 white and 5 black balls, what is the probability that one of the drawn balls is white and the other two black?

Solution 1

2.6 Sample Spaces Having Equally Likely Outcomes

Example If 3 balls are "randomly drawn" from an urn containing 6 white and 5 black balls, what is the probability that one of the drawn balls is white and the other two black?

Solution 1

(I) Regard the order in which the balls are selected as being relevant.
There are a total of $11 \cdot 10 \cdot 9 = 990$ ways to draw 3 balls, this is $|S|$.
To get one white two black, there are 3 cases:

first ball is white. Number of ways $= 6 \cdot 5 \cdot 4 = 120$.

second ball is white. Number of ways $= 5 \cdot 6 \cdot 4 = 120$.

third ball is white. Number of ways $= 5 \cdot 4 \cdot 6 = 120$.

Therefore, $|A| = 120 + 120 + 120 = 360$. It follows that

$$P(A) = \frac{360}{990} = \frac{4}{11}.$$

2.6 Sample Spaces Having Equally Likely Outcomes

Example If 3 balls are "randomly drawn" from an urn containing 6 white and 5 black balls, what is the probability that one of the drawn balls is white and the other two black?

Solution 2

(II) Regard the order in which the balls are selected as being irrelevant.
There are a total of $\binom{11}{3} = 165$ ways to draw 3 balls, this is $|S|$.
To get one white two black, there are $\binom{5}{2}\binom{6}{1} = 60$ ways.
Hence

$$P(A) = \frac{60}{165} = \frac{4}{11}.$$

2.6 Sample Spaces Having Equally Likely Outcomes

Example (Birthday problem II). How large must the group be so that there is a probability of greater than 0.5 that someone will have the same birthday as you do?

Solution

2.6 Sample Spaces Having Equally Likely Outcomes

Example (Birthday problem II). How large must the group be so that there is a probability of greater than 0.5 that someone will have the same birthday as you do?

Solution

Assuming 365 equally likely birthdays. In a group of n people the probability that all will fail to have your birthday is $(364/365)^n$. Setting this equal to 0.5 and solving

$$n = \frac{\log(0.5)}{\log(364/365)} = 252.7.$$

So we need 253 people.