Section 1.3 conditional probability

郁金香球茎

1.3.1 a motivation example [Tulip Bulb combinations]

Similar in appearance

Assumption: Given 20 bulbs and all bulbs are “equally likely”

Table 1.3-1 Tulip combinations

	Early (E)	Late (L)	Totals
Red (R)	5	8	13
Yellow (Y)	3	4	7
Totals	8	12	20

Experiment 1: *Select one bulb randomly*

The probability that the selected bulb will be a red tulip (R) *is*

$$P(R) = \frac{N(R)}{N(S)} = \frac{13}{20}$$

*Experiment*2: *Select one bulb randomly from the bulbs that will boom early.*

The probability that the selected bulb will be a red tulip(R) *given that the selected bulb is known to bloom early*(E) *is*

$$P(R|E) = \frac{N(R\cap E)}{N(E)} = \frac{5}{8} = \frac{N(R\cap E)/N(S)}{N(E)/N(S)} = \frac{P(R\cap E)}{P(E)}.$$

$$Experiment2: \quad P(\underset{A}{R}|\underset{B}{E}) = \frac{N(R\cap E)}{N(E)} = \frac{5}{8} = \frac{N(R\cap E)/N(S)}{N(E)/N(S)} = \frac{P(R\cap E)}{P(E)}$$

For experiment 2, we are interested only in those outcomes which are elements of a subset B of the sample space S.

1. The essential sample space is B (reduced from S to B)
2. Study the problem of how to define a new probability function associated with this sample space B,

Under the assumption that all out comes are "equally likely", the above example give us the idea:

$$P(A|B) = \frac{N(A\cap B)}{N(B)} = \frac{N(A\cap B)/N(S)}{N(B)/N(S)} = \frac{P(A\cap B)}{P(B)}$$

leading to the next definition:

Definition 1.3-1 [conditional probability]

The **conditional probability** of an event A, given that event B has occurred, is defined by:

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

- Provided that P(B) > 0
- **Need not** to be "equally likely"!

Example 2: $P(A)=0.4,\quad P(B)=0.5,\quad P(A\bigcap B)=0.3,$

$$\Rightarrow P(A|B)=\frac{P(A\bigcap B)}{P(B)}=\frac{0.3}{0.5}=0.6$$

$$P(B|A)=\frac{P(B\bigcap A)}{P(A)}=\frac{0.3}{0.4}=0.75.$$

Can conditional probability be larger than 1 or negative ?

---

* *Conditional probability satisfy the axims for a probability function:*

1. $P(A|B)\geq 0$
2. $P(B|B)=1$
3. *If* $A_1, A_2, \ldots A_k$ *are mutually exclusive events, then*

$$P(A_1\cup A_2\cup\cdots\cup A_k|B)=P(A_1|B)+\cdots+P(A_k|B)$$

for each integer k, *and*

$$P(A_1\cup A_2\cup\cdots|B)=P(A_1|B)+P(A_2|B)+\cdots$$

for an infinitely, but countable number of events.

4. *The propability properties also holds for conditional probabilities. For example.*

$$P(A'|B)=1-P(A|B)$$

is true.

Example 3: 25 balloons on a board, of which 10 balloons are yellow, 8 are red, and 7 are green. A player randomly hits one of them.

A={The first balloon hit is yellow}

B={The second hit is yellow}

Question:

$P(A \cap B)$

What is the probability that the $\boxed{\text{first two balloons are all yellow?}}$

$$Solution: \quad P(A) = \frac{10}{25}, P(B|A) = \frac{9}{24} \quad \Rightarrow \quad P(A \cap B) = P(A)P(B|A) = \frac{10}{25} \times \frac{9}{24}.$$

Definition 1.3-2 [multiplication rule]

The probability that two events, A and B, both occur is given by the **multiplication rule**,

$$P(A \cap B) = P(A)P(B|A),$$

provided P(A) > 0 or by

$$P(A \cap B) = P(B)P(A|B),$$

provided P(B) > 0.

Example 4: A bowl contains 10 chips in total, 7 blue and 3 red. Two chips drawn successively at random and without replacement.

Our goal is to compute the probability that $\underbrace{\text{the first draw is red}}_{A}$ **and** $\underbrace{\text{the second draw is blue}}_{B}$:

$$Solution: \ P(A)=\frac{3}{10}, P(B|A)=\frac{7}{9} \quad \Rightarrow \quad P(A\cap B)=P(A)P(B|A)=\frac{3}{10}\times\frac{7}{9}=\frac{7}{30}.$$

Quiz: Roll a pair of 4-sided dice and observe the sum of the dice

A={ A sum of 3 is rolled }

B={ A sum of 3 or a sum of 5 is rolled }

C={ A sum of 3 is rolled before a sum of 5 is rolled }

Question: Compute P(A)、P(B) and P(C).

$$Consider \ \ P(A) \ \ and \ \ P(B):$$

$$the \ \ sample \ \ space \ \ S=\{(1,1),(1,2),\ldots,(4,4)\}$$

$$P(A)=\frac{\mathbb{N}(A)}{\mathbb{N}(S)}=\frac{2}{16}, \qquad P(B)=\frac{\mathbb{N}(B)}{\mathbb{N}(S)}=\frac{6}{16}.$$

Quiz (c.n.t.)

Consider P(C):

- Method 1 [by definition]:

① Figure out the random experiment

② Figure out the sample space and the event.

For ①, repeat the experiment of rolling a pair of 4-sided dice and record the sum of dice. For each repetition, we keep rolling the dice till we see either a sum of 3 or a sum of 5. Then we stop because we have an answer to the problem whether a sum of 3 is rolled before a sum of 5 is rolled.

For instance,

repetition 1:	2,4,6,3.
repetition 2:	8,6,7,4,5.
repetition 3:	6,5

The sums other than 3 and 5 don't matter and we can remove them.

Repetition 1: a sum of 3 first.

Repetition 2: a sum of 5 first.

Repetition 3: a sum of 5 first.

The problem reduces to roll the pair of dice once and compute the probability that the sum is 3.

Quiz (c.n.t.)

For ②, the reduced sample space

$$S_r = \left\{\begin{matrix} (1,2), & (2,1) \\ (2,3), & (3,2) \\ (1,4), & (4,1) \end{matrix}\right\}$$

Gives a sum of 3

Gives a sum of 5

$$\Rightarrow P(C) = P(\{\ roll\ \ the\ \ pair\ \ of\ \ dice\ \ once\ \ and\ \ the\ \ sum\ \ is\ \ 3\})$$

$$= \frac{N(\{\ roll\ \ the\ \ pair\ \ of\ \ dice\ \ once\ \ and\ \ the\ \ sum\ \ is\ \ 3\})}{N(S_r)}$$

$$= \frac{2}{6}$$

Quiz (c.n.t.)

Method 2 [by conditional probability]

$$P(C) = P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{2/16}{6/16} = \frac{2}{6}$$

Note that the event "A|B" is the same as event "C"

This is because

① Event C is concerned with the cases where the sum is either 3 or 5

②"B" happened means that the sum is either 3 or 5. If B happened, then A|B is nothing but the event { roll the pair of dice once and the sum is 3}.

Section 1.4 independent events

- Intuition and motivation examples.

Intuition: For certain pair of events, the occurrence of one of them **may** or **may not** change the probability of the occurrence of the other. In the latter case, they are said to be **independent events**.

Example 1:

Experiment: flip a coin twice and observe the sequence of heads and tails.

Sample Space: S={ HH, HT, TH, TT}

Events: A={ heads on the first flip }={ HH, HT },

B={ tails on the second flip }={ HT, TT },

C={ tails on both flip }={ TT }.

Then $P(A)=\frac{2}{4},\ P(B)=\frac{2}{4},\ P(C)=\frac{1}{4}.$

Given that C has occurred, then $P(B|C)=1\ \ because\ \ C\subset B$

Given that A has occurred, then $P(B|A)=\frac{P(B\cap A)}{P(A)}=\frac{1/4}{3/4}=\frac{1}{2}$

So the occurrence of A has not changed the probability of B.

$P(B|A)=P(B),\qquad Similarly,\quad \boxed{P(A|B)=P(A).}$ Verify by yourself

Example 1(c.n.t):

- Intuitionally, this means that the probability of B doesn't depend on the knowledge about event A.

⇒ A and B are independent events

- That is, events A and B are independent if the occurrence of one of them does not affect the probability of the occurrence of the other. In math,

$P(A)>0$ $P(B)>0$

$$P(B|A) = P(B) \qquad P(A|B) = P(A)$$

- This example motivates the following definition of independent events.

Definition 1.4-1 [independent events]

Events A and B are **independent** if and only if $P(A \cap B) = P(A)P(B)$.
Otherwise, A and B are called **dependent** events.

Example 2: A red die and a white die are rolled

- S={ (1,1), (1,2), ……, (6,6) } (Number of all outcomes is 36)
- A={ 4 on the red die }. B={ sum of dice is odd }.

Assuming the two dies are fair. Are events A and B independent ?

$$Solution: P(A)=\frac{6}{36},\ \ P(B)=\frac{18}{36}\ \ P(A\cap B)=\frac{3}{36}$$

$$P(A\cap B)=\frac{3}{36}=P(A)P(B)=\frac{6}{36}\times\frac{18}{36} \quad \Rightarrow \ A \ \ and \ \ B \ \ are \ \ independent.$$

Theorem 1.4-1

If A and B are independent events, then the following pairs of events are also independent:

(a) A and B;

(b) A' and B;

(c) A' and B'.

The proofs are in the later page.

Proofs of theorem 1.4-1

$(a) Proof:$

$$\begin{aligned}
P(A \cap B') &= P(A)P(B' | A) && (multiplication \;\; rule) \\
&= P(A)\left[1 - P(B|A)\right] && (axims \;\; for \;\; conditional \;\; probability) \\
&= P(A)\left[1 - P(B)\right] && (definiton \;\; of \;\; independent \;\; events) \\
&= P(A)P(B'). && (properties \;\; of \;\; probability \;\; function)
\end{aligned}$$

The proofs of part(b) and (c) will be written simply:

$Proof:$

$$(b) P(A' \cap B) = P(B)P(A' | B) = P(B)\left[1 - P(A|B)\right] = P(B)\left[1 - P(A)\right] = P(B)P(A')$$

$$(c) P(A' \cap B') = P\left[(A \cup B)'\right] = 1 - (A \cup B) = 1 - P(A) - P(B) + P(A \cap B)$$

$$= 1 - P(A) - P(B) + P(A)P(B) = \left[1 - P(A)\right]\left[1 - P(B)\right] = P(A')P(B').$$

$Q.E.D.$

Then let's extend the definition of independent events to more than two events.

Definition 1.4-2 [mutually independent]

Events A, B, and C are **mutually independent** if and only if the following two conditions hold:

(a) A, B, and C are **pairwise** independent; that is,

$$\begin{cases} P(A \cap B) = P(A)P(B) \\ P(A \cap C) = P(A)P(C) \\ P(B \cap C) = P(B)P(C) \end{cases}$$

(b) $P(A \cap B \cap C) = P(A)P(B)P(C)$.

Remark:

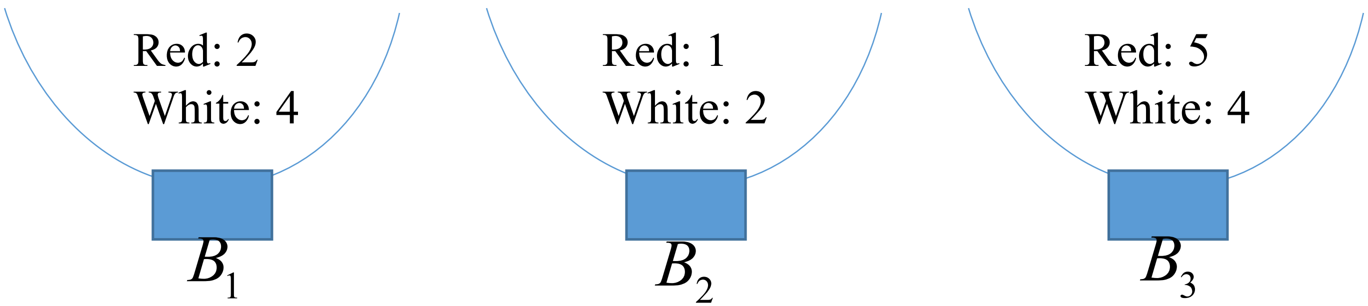
1. This definition can be extended to the mutual independence of four or more events. In such a extension, each pair, triple, quartet, and so on, must satisfy this type of multiplication rule. In the following we will discuss its definition in mathematical formal.
2. If there is no possibility of misunderstanding, independent is often used without the modifier mutually when several events are considered.

Events $A_1, A_2, \ldots, A_k$ *are independent if and only if the following condition hold:*

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_j}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_j}), \qquad j = 2, \ldots, k.$$

Section 1.5 Bayes's theorem

Motivation example:

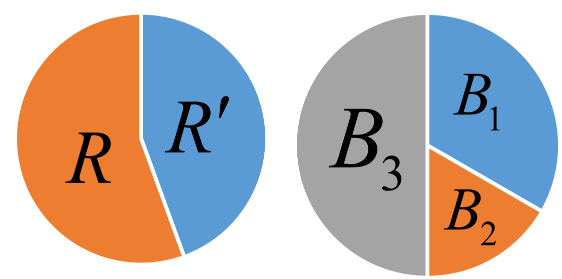


Experiment: Select a bowl first, and then draw a chip from the selected bowl.

Goal: compute the probability of event R={ draw a red chip }.

Assumption: $P(B_1)=\frac{1}{3},\ P(B_2)=\frac{1}{6},\ P(B_3)=\frac{1}{2}.$

Question1: Compute the probability of event R



$$Solution: P(R)=P(S\cap R)$$

$$=P\left[(B_1\cup B_2\cup B_3)\cap R\right]=P\left[(B_1\cap R)\cup(B_2\cap R)\cup(B_3\cap R)\right]$$

$$=P(B_1\cap R)+P(B_2\cap R)+P(B_3\cap R)$$

$$=P(B_1)P(R|B_1)+P(B_2)P(R|B_2)+P(B_3)P(R|B_3)$$

$$=\frac{1}{3}\times\frac{2}{6}+\frac{1}{6}\times\frac{1}{3}+\frac{1}{2}\times\frac{5}{9}=\frac{4}{9}.$$

Question2: Suppose now that the outcome of the experiment is a red chip, but we do not know from which bowl the chip was drawn. We are interested in the **conditional probability** that the chip was drawn from the bowl, namely, $P(B_1|R),\ P(B_2|R),\ P(B_3|R).$

From the definition of conditional probability, we consider

$$P(B_1|R)=\frac{P(B_1\cap R)}{P(R)}=\frac{P(B_1)P(R|B_1)}{P(B_1)P(R|B_1)+P(B_2)P(R|B_2)+P(B_3)P(R|B_3)}$$

$$=\frac{\frac{1}{3}\times\frac{2}{6}}{\frac{4}{9}}=\frac{1}{4}$$

$$Similarly,\ \ P(B_2|R)=\frac{1}{8},\ \ P(B_3|R)=\frac{5}{8}$$

$Recall\ \ that:$

$$P(B_1)=\frac{1}{3},\ \ P(B_2)=\frac{1}{6},\ \ P(B_3)=\frac{1}{2}\qquad \rightarrow prior\ \ prpbability$$

$$P(B_1|R)=\frac{1}{4},\ \ P(B_2|R)=\frac{1}{8},\ \ P(B_3|R)=\frac{5}{8}\qquad \rightarrow posterior\ \ probability$$

$We\ \ observe:$ ① $P(B_i)\ \ different\ \ from\ \ P(B_i|R)$

② $The\ \ changer\ \ coincide\ \ with\ \ our\ \ intuition.$

Generalization: Assume that

1. S is a sample space, and $B_1, B_2, \ldots, B_m$ are mutually exclusive and exhaustive. That is:

$$S = B_1 \bigcup B_2 \bigcup \cdots \bigcup B_m \quad and \quad B_i \bigcap B_j = \phi, i \neq j$$

2. The prior probabilities of B_i is positive. That is:

$$P(B_i) > 0, \quad i = 1, 2, \ldots, m.$$

Then we have

a) For any event A,

$$P(A) = P(A \cap S) = P\left[A \cap (B_1 \cup B_2 \cup \cdots \cup B_m)\right]$$
$$= P\left[(A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_m)\right]$$
$$= P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_m)$$
$$= \sum_{i=1}^{m} P(A \cap B_i)$$

$$\Rightarrow \boxed{P(A) = \sum_{i=1}^{m} P(B_i) P(A|B_i)} \longrightarrow \text{Total probability}$$

b) If P(A)＞0， then

$$P(B_k|A) = \frac{P(B_k \cap A)}{P(A)}, \quad k = 1, \ldots, m.$$

$$\text{Hence } \boxed{P(B_k|A) = \frac{P(B_k) P(A|B_k)}{\sum_{i=1}^{m} P(B_i) P(A|B_i)}.} \longrightarrow \text{Bayes's Theorem}$$

- $P(B_k) \rightarrow$ *propr probability*
- $P(B_k|A) \rightarrow$ *posterior probability*
- $P(A|B_i) \rightarrow$ *likelihood of* B_k, *A is called a data.*