Section 4 Conditional probability、 total probability formula and Bayes' rule

- 一、Conditional probability
- 二、Total probability and Bayes' rule
- 三、Conclusion

下页 返回

一、Conditional probability

1.e.g.

Given three families, one has only one boy and no girl, one has only one girl and no boy, and the last family has one boy and one girl:

Question? Determine the probability of the following events.

1, A: select a family randomly from the three families, what is the probability that it has a boy?

2, B: select a family randomly from the three families, what is the probability that it has a girl?

3, C: given a family selected has a girl, what is the probability that it only has a boy? $P(A|B) = ?$

Solution: **(1)** $$P(A)=\frac{2}{3}$$

(2) $$P(B)=\frac{2}{3}$$

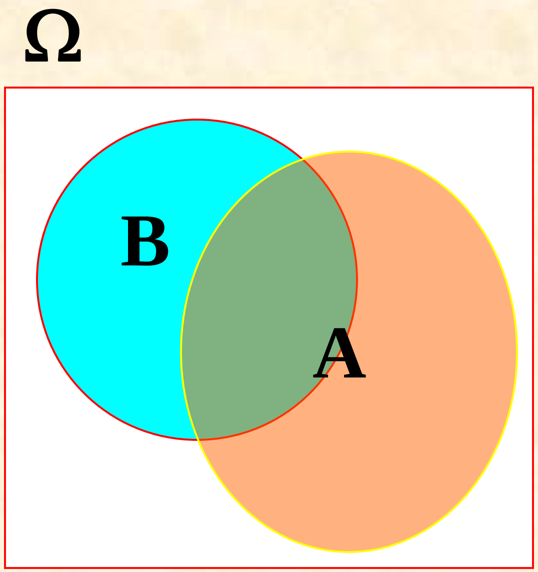
(3) $$P(AB)=\frac{1}{3}$$

$$(4)\,\mathrm{P}(\mathrm{A}|B)=\frac{P(\mathrm{AB})}{P(\mathrm{B})}=\frac{1}{2}$$

note. $1^\circ\ P(A)=\dfrac{2}{3}\neq P(A|B)$

$2^\circ\ P(AB)=\dfrac{1}{3}\neq P(A|B)$

$P(AB):\ \Omega=$



{only one boy family, only one girl family, one boy and one girl family }

$\boldsymbol{P(A|B)}:\ \Omega_B=B$

$3^\circ\ P(A|B)=\dfrac{1}{2}=\dfrac{1/3}{2/3}=\dfrac{P(AB)}{P(B)}$

Can this merely be coincidence? **No.**

2. Definition 1.8 (conditional probability)

Given two events A，*B*， **if** $P(B) > 0$**, then**

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

is the conditional probability of A**,** ***while given B*.**

Note. **The way to calculate the conditional probability:**

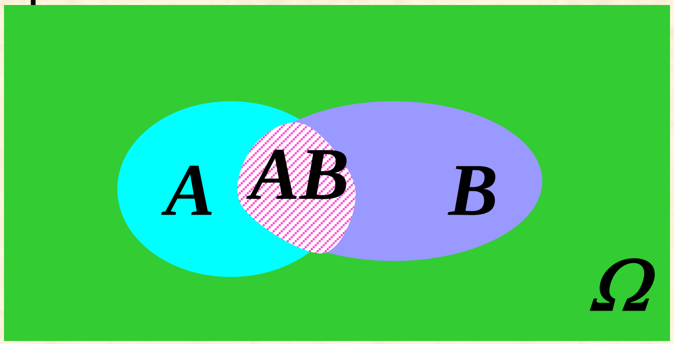
① **reduce the sample space;**

② **definition.**

Difference between $P(AB)$ and $P(A|B)$:

$P(AB):\ \Omega$

$\boldsymbol{P(A|B):}\quad \Omega_B = B$



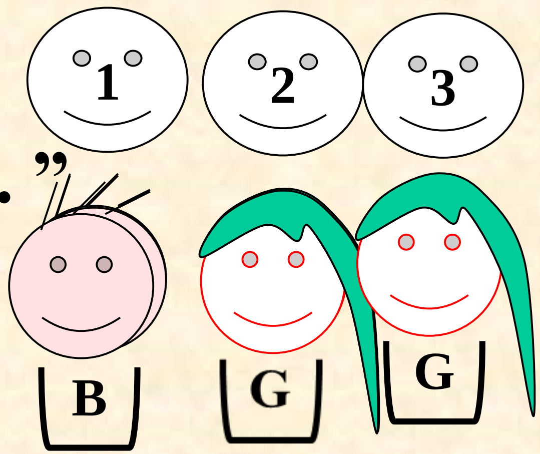
e.g.，classical probability

$$P(AB) = \frac{|AB|}{|\Omega|}$$

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

e.g.1 **(1)Given the families having 3 children， what is the probability for the event that one family selected randomly has at least one girl (suppose that the boy and girl are born equally).**

Solution: $|\Omega|\ :\ 2^3.$



A=“there is at least one girl in the family. ”

$\bar{A}$ **=“all the children are boy.”**

$$P(\bar{A}) = \frac{1}{2^3} = \frac{1}{8}$$

$$\therefore \quad P(A) = 1 - P(\bar{A}) = 1 - \frac{1}{8} = \frac{7}{8}$$

(2) **For the families having 3 children， given the family selected has at least one girl, what is the probability for the event that this family selected has at least one boy?**

Solution: denote A=“there is at least one girl in the family.”

B=“there is at least one boy in the family.”

$$\text{then } P(B) = 1 - P(\bar{B}) = 1 - \frac{1}{2^3} = \frac{7}{8},$$

Denote C=“the family has one boy and two girls.”

D=“the family has two boys and one girl.”

$$\text{then } AB = C + D \quad (\mathrm{C} \cap \mathrm{D} = \varnothing)$$

$$\therefore \quad P(AB) = P(\mathrm{C}) + P(\mathrm{D})$$

$$= 2 \times \frac{3}{2^3} = \frac{6}{8}$$

$$Thus, \quad P(B|A) = \frac{P(AB)}{P(A)}$$

$$= \frac{\mathbf{6/8}}{\mathbf{7/8}} = \frac{\mathbf{6}}{\mathbf{7}}.$$

e.g.2 **The probability for the dog that can live more than 8 years is 0.8, and that more than 10 years is 0.4. Given one dog who is already 8 years, what is the probability for it can live more than 10 years?**

Solution: **Suppose** $A =$ **" live more than 8 years" ;**

$B =$ **"live more than 10 years" ;**

then $$P(B|A) = \frac{P(AB)}{P(A)}. \qquad (\because B \subset A, \ \therefore AB = B)$$

$$\text{Because } P(A) = 0.8, \ P(B) = 0.4, \qquad P(AB) = P(B),$$

$$\text{Thus, } P(B|A) = \frac{P(AB)}{P(A)} = \frac{0.4}{0.8} = \frac{1}{2}.$$

3. Properties of conditional probability

(1)non-negativity: $0 \le P(A|B) \le 1;$

Proof: $\because AB \subset B$ $\therefore 0 \le P(AB) \le P(B)$

$\because P(B) > 0$ $\therefore 0 \le \dfrac{P(AB)}{P(B)} \le 1$

$i.e., 0 \le P(A|B) \le 1.$

(2) normalization ： $P(\Omega|B) = 1;$

Proof: $\because \Omega B = B$

$$\therefore P(\Omega|B) = \frac{P(\Omega B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

(3) additivity：

Given A_i, A_j**, i, j=1,2,…, are mutually exclusive, i.e.,** $A_i \cap A_j = \phi$

$$Then, \quad P\left(\left(\sum_{k=1}^{\infty} A_k\right) \middle| B\right) = \sum_{k=1}^{\infty} P(A_k | B)$$

Proof:

$$P\left(\left(\sum_{k=1}^{\infty} A_k\right) \middle| B\right) = \frac{P\left(\left(\sum_{k=1}^{\infty} A_k\right) B\right)}{P(B)}$$

$$= \frac{\sum_{k=1}^{\infty} P(A_k B)}{P(B)} = \sum_{k=1}^{\infty} P(A_k | B)$$

(4) :

$$P((A_1 \cup A_2)|B)$$

$$= P(A_1|B) + P(A_2|B) - P(A_1A_2|B)$$

Proof: Since the denominators are the same, it is equivalent to prove the numerators are the same

$$P((A_1 \cup A_2)B) = P(A_1B \cup A_2B)$$

$$= P(A_1B) + P(A_2B) - P(A_1A_2B)$$

(5) The conditional probability of the opposite event:

$$P(A|B) = 1 - P(\overline{A}|B).$$

4. Multiplication

$$\text{if } P(B) > 0, \text{then } P(AB) = P(B)P(A|B)$$

$$\text{if } P(A) > 0, \text{then } P(AB) = P(A)P(B|A)$$

Meaning:

The probability of the event AB equals the probability of one event(i.e.,A) times the conditional probability of the other event(i.e., B) given the event(i.e., A) happens .

Generalization: Given A, B, C , and $P(AB) > 0$, *then*

$$P(ABC) = P(A)P(B|A)P(C|AB).$$

In general, suppose $A_1, A_2, \cdots, A_n$ are events, if

$$P(A_1 A_2 \cdots A_{n-1}) > 0,$$

then

$$P(A_1 A_2 \cdots A_n) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 A_2) \cdots \cdot$$

$$\cdots \cdot P(A_n | A_1 A_2 \cdots A_{n-1}).$$

二、Theory of the total probability and the Bayes' Rule

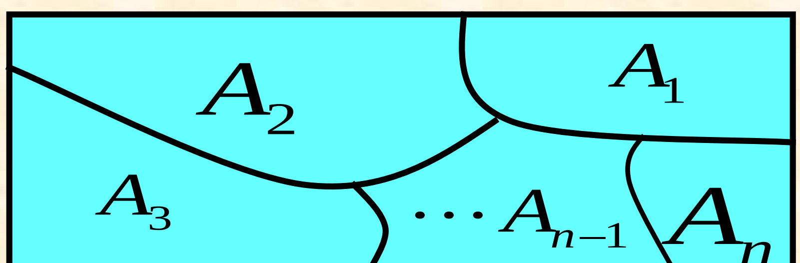
1. The partition of the sample space

Given $A_1, A_2, \ldots, A_n$ **are n events of E, and** Ω **:sample space. If**

$$(1)\quad A_i A_j = \varnothing,\ i \neq j,\ i, j = 1, 2, \cdots, n;$$

$$(2)\quad A_1 \cup A_2 \cup \cdots \cup A_n = \Omega,$$

then, $A_1, A_2, \ldots, A_n$ **constitute a partition of the sample space.**



2. Theory of the total probability

If the events $A_1, \ldots, A_n$ constitute a partition of the sample space Ω such that $P(A_i) \neq 0$, then for any event B of Ω ,

$$P(B) = P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) + \cdots + P(B \mid A_n)P(A_n)$$

$$= \sum_{i=1}^{n} P(A_i)P(B \mid A_i)$$

Total probability

Proof:

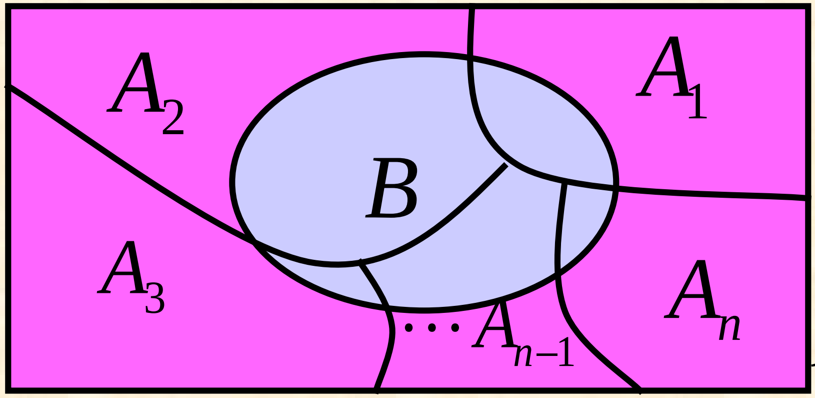
$$B = B\Omega = B\bigcap(A_1 \bigcup A_2 \bigcup \cdots A_n)$$

$$= BA_1 \bigcup BA_2 \bigcup \cdots \bigcup BA_n.$$

$$\text{Since } A_iA_j = \varnothing \Rightarrow (BA_i)(BA_j) = \varnothing$$

$$\Rightarrow P(B) = P(BA_1) + P(BA_2) + \cdots + P(BA_n)$$

$$\Rightarrow P(B) = P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + \cdots + P(A_n)P(B \mid A_n)$$



breaking up the whole into parts

Note. The condition of the total probability:

$$\sum_{i=1}^{n} A_i = \Omega$$

can be replaced by

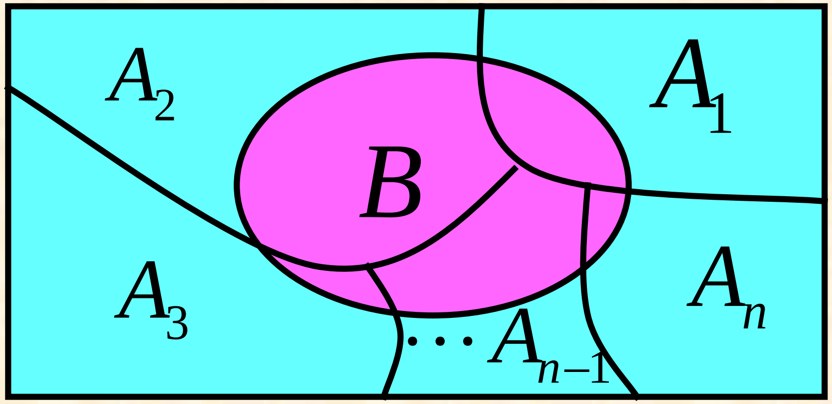
$$B \subset \sum_{i=1}^{n} A_i .$$

3.Meaning of the total probability

If the event *B happening is due to n events* A_i (i=1,2,···,n), and A_i, A_j ($i \neq j$) are mutually exclusive, then

P(*B*) *has relationship with P*(*B*A_i)(i=1,2,···,n).

$$P(B) = \sum_{i=1}^{n} P(BA_i).$$



目录 上页 下页 返回 结束

e.g.4

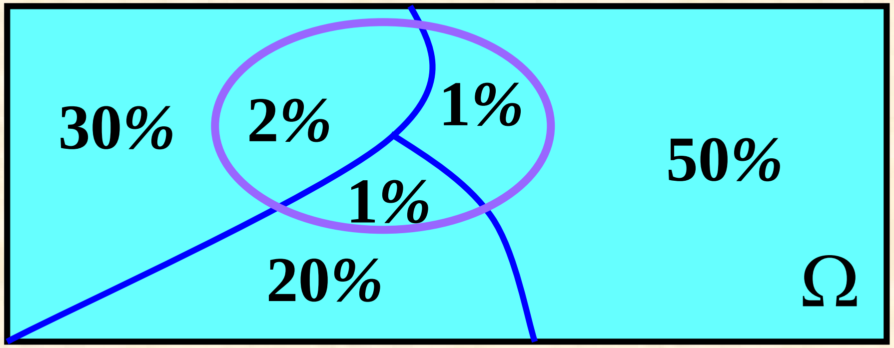
Given a batch of products coming from **three firms**, 30%, *50% and 20%of the products are from firm 1, firm 2 and firm 3*, respectively. The defect rates of the three firms are 2% , 1%, 1% *respectively*. **What is the probability that one product selected randomly is defected?**

Solution: $A =$ "one product selected randomly is defected".

$B_i =$ "the product selected randomly is from firm i, i=1,2,3".

then, $B_1 \cup B_2 \cup B_3 = \Omega,$

$$B_i B_j = \varnothing, \quad i, j = 1,2,3.$$



By the theory of total probability .

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3).$$

$$P(B_1) = 0.3,\quad P(B_2) = 0.5,\quad P(B_3) = 0.2,$$

$$P(A|B_1) = 0.02,\ P(A|B_2) = 0.01,\ P(A|B_3) = 0.01,$$

Thus, $P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)$

$$= 0.02 \times 0.3 + 0.01 \times 0.5 + 0.01 \times 0.2 = 0.013.$$

4. Bayes' Rule

Theorem **If the events** $A_1, A_2, \ldots, A_n$ **constitute a partition of** Ω**, such that** $P(A_j) \neq 0, j = 1, \ldots, n$**, then for any event B of** Ω **, with** $P(B) \neq 0$**,**

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_{j=1}^{n} P(B \mid A_j)P(A_j)}, \quad i = 1, 2, \cdots, n.$$

Such a formula is called Bayes' Rule.

Proof: $$P(A_i|B)=\frac{P(A_iB)}{P(B)}\qquad i=1,2,\cdots,n.$$

$$=\frac{P(B|A_i)P(A_i)}{P(B)}$$

$$=\frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)}$$

e.g. 5 One box contains 12 balls， 9 are new, 3 old.

In the first set 3 balls are used . After the first set, the balls are put back. Then the second set begins, still 3 balls are chosen from the 12 balls.

Question? Determine the probability of the following events.

(1) B=“in the second set, all the 3 balls chosen are new.”;

(2) Given the 3 balls chosen in the second set are new, what is the probability of the event that all the three balls chosen in the first set are new ;

Solution: suppose $A_i =$ “in the first set, there are i new balls are chosen ” $(i = 0, 1, 2, 3)$

(1) in the second set, all the 3 balls chosen are new;

A_i =“first set, i new balls”

B=“second set, all 3 balls are new”

$$\because \quad P(A_i) = \frac{C_9^i \cdot C_3^{3-i}}{C_{12}^3} \quad (i = 0, 1, 2, 3)$$

$$P(B|A_i) = \frac{C_{9-i}^3}{C_{12}^3}$$

$$\therefore \quad P(B) = \sum_{i=0}^{3} P(A_i)P(B|A_i)$$

$$= \sum_{i=0}^{3} \frac{C_9^i \cdot C_3^{3-i}}{C_{12}^3} \cdot \frac{C_{9-i}^3}{C_{12}^3} = 0.146.$$

First set

new: 9

old: 3

Second set

new: 9-*i*

old: 3+*i*

(the new balls used become old)

(2) Given the 3 balls chosen in the second set are new, what is the probability of the event that all the three balls chosen in the first set are new ;

$$P(A_3|B) = \frac{P(A_3B)}{P(B)}$$

$$\because \quad P(A_3B) = P(A_3)P(B|A_3)$$

$$= \frac{C_9^3 \cdot C_3^0}{C_{12}^3} \cdot \frac{C_6^3}{C_{12}^3}$$

$$\therefore \quad P(A_3|B) = \frac{P(A_3B)}{P(B)} = \frac{5}{21} = 0.24$$

If in the first set, all the balls chosen are new, then in the second set

new: 9-3

old: 3+3

(balls are put back)

三、Conclusion

1.Conditional probability

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

multiplication principle

$$P(AB)=P(A)P(B|A)$$

Theory of total probability

$$P(A)=P(B_1)P(A|B_1)+P(B_2)P(A|B_2)+\cdots+P(B_n)P(A|B_n)$$

Bayes' Rule

$$P(B_i|A)=\frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{n}P(B_j)P(A|B_j)},\quad i=1,2,\cdots,n.$$

目录 上页 下页 返回 结束

2. The difference between P(A| B) and P(AB).

P(AB) is calculated by considering Ω, while P(B|A) is calculated by considering Ω_A.

For classical probability,

$$P(B|A)=\frac{|AB|}{|\Omega_A|},$$

$$P(AB)=\frac{|AB|}{|\Omega|}.$$

Normally, P(AB) is smaller than P(B|A).

Information of Bayes

Thomas Bayes

Born: **1702 in London, England**

Died: **17 April 1761 in Tunbridge Wells, Kent, England**