

자료구조및실습 1장. 자료구조와 알고리즘

# 第3章 概率

- 3.1 样本空间与事件
- 3.2 概率的定义
- 3.3 条件概率
- 3.4 贝叶斯(Bayes)定理

2/31

## 3.1 样本空间与事件



## [定义 3-1] 样本空间(sample space)

- 进行随机试验(或者观察),可能出现的所有结果的集合(S)

■ 元素(element):构成样本空间的要素 在随机试验中出现的各种结果

## [定义 3-2] 事件(event)

- 在构成样本空间的元素中,成为关心对象的元素的集合
- 样本空间的子集合(subset)



### [例 3-1] 请求下列情况下的样本空间。

① 掷一颗骰子的随机试验

$$S = \{1, 2, 3, 4, 5, 6\}$$

② 依次投掷两个骰子的随机试验

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

#### ③ 同时投掷两颗骰子的随机试验

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
<b>(4,1)</b>	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

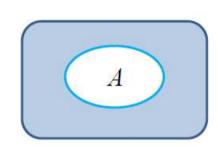
$$|S| = 1 + 2 + \dots + 6$$
$$= \frac{6 \times 7}{2} = 21$$

 $|S| = 6 \times 6 = 36$ 



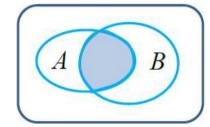
- 事件的基本运算
  - ✓ 互补事件(complement)

 $A^{c}$ 



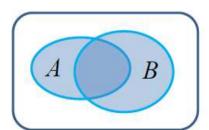
✓ 交事件(intersection)

 $A \cap B$ 

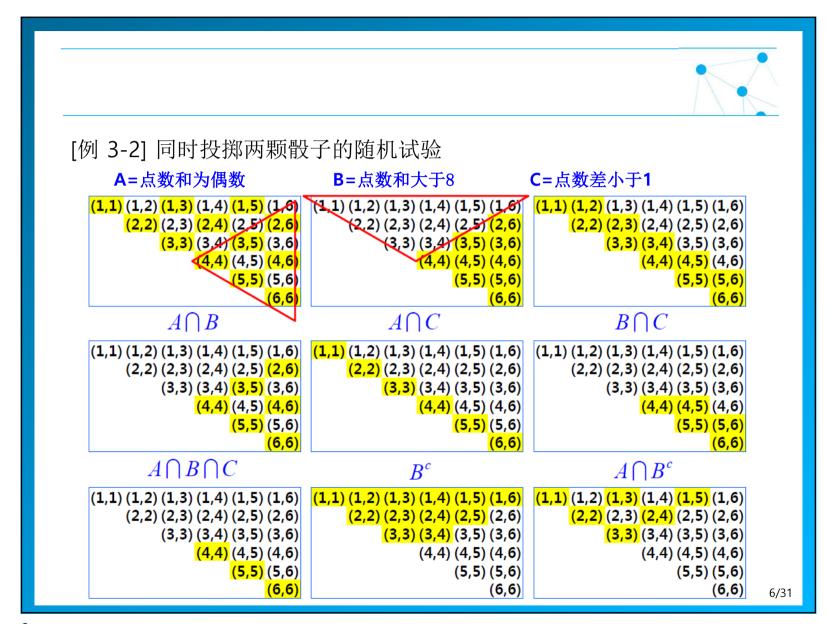


✓ 和事件(union)

 $A \bigcup B$ 



5/31



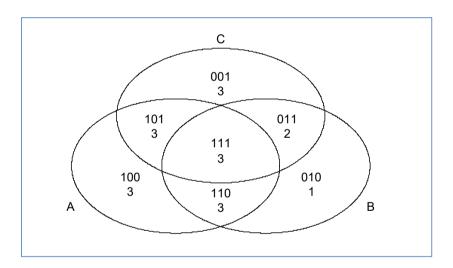
6

자료구조및실습 1장. 자료구조와 알고리즘



### [例 3-3] 同时投掷两颗骰子的随机试验的韦恩图

A=点数和为偶数 B=点数和大于8 C=点数差小于1



 $A = 100 \ \bigcup \ 101 \ \bigcup \ 111 \ B = 010 \ \bigcup \ 011 \ \bigcup \ 111 \ C = 001 \ \bigcup \ 011 \ \bigcup \ 101 \ \bigcup \ 111$ 

 $A \cap B = 110 \bigcup 111$   $A \cap C = 101 \bigcup 111$   $B \cap C = 011 \bigcup 111$ 

 $A \cap B \cap C = 111$   $A \cap B^c = 100 \cup 101$   $A^c \cap B = 010 \cup 011$ 



(6,6)

### [例 3-4] 同时投掷两颗骰子的随机试验(相互排斥)

#### A=点数差大于3

B=点数积大于20

# 事件 A (点数差大于3)

$$A = \{(1,4), (1,5), (1,6), (2,5), (2,6), (3,6)\}$$

(1,1) (1,2) (1,3) (1,4) (1,5) (1,6) (2,2) (2,3) (2,4) (2,5) (2,6) (3,3) (3,4) (3,5) (3,6) (4,4) (4,5) (4,6) (5,5) (5,6)

# 事件 B (点数积大于20)

$$B = \{(4,5), (4,6), (5,5), (5,6), (6,6)\}$$

(1,1) (1,2) (1,3) (1,4) (1,5) (1,6) (2,2) (2,3) (2,4) (2,5) (2,6) (3,3) (3,4) (3,5) (3,6) (4,4) (4,5) (4,6) (5,5) (5,6)

# 确认是否相互排斥 ⇒ 由于没有交事件, 因此是相互排斥

$$A \cap B = \phi$$

## 3.2 概率的定义



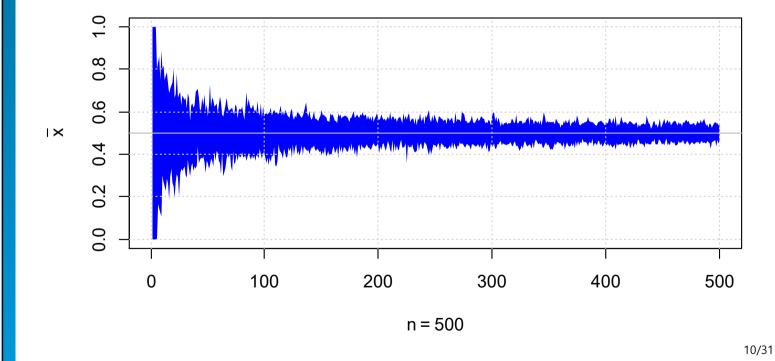
- 古典概率(classical probability)
  - 预测掷硬币时出现正面的概率为1/2
- 相对频数(relative frequency)概率
  - 实际掷硬币时出现正面的次数的比率(相对频数)
- 大数定律(law of large numbers)
  - 若无数次地掷硬币的话,相对频数概率将越来越接近古典概率

# 大数定律(law of large numbers)



[例 3-5] 投掷掷硬币1到500次,估计出现正面的概率





10

자료구조및실습 1장. 자료구조와 알고리즘

## 3.2 概率的定义



## [定义 3-3] 概率(probability)

- 用数字来表示某一事件会发生的可能性
- 如果样本空间的所有元素的发生概率都是一样的话,事件A的概率就是属于事件A的元素的个数占所有元素的个数的比率。
- 概率的特征
  - $P(S)=1, 0 \le P(A) \le 1, P(\emptyset)=0$



### [例 3-6] 在投掷硬币四次的试验中,正面出现两次以上的概率

$$S = \{HHHH, HHHT, HHTH, \dots, TTTH, TTTT\}$$

$$\Rightarrow |S| = 2^{4} = 16$$

$$|A| = {}_{4}C_{4} + {}_{4}C_{3} + {}_{4}C_{2} = 1 + 4 + 6 = 11$$

$$\Rightarrow P(A) = 11/16 = 0.6875$$

[例 3-7] 同时投掷四颗骰子的试验

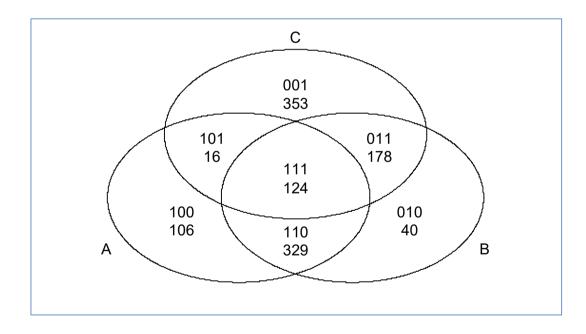
① A= 数字之和为15 ② B= 6的个数大于1 ③ C= 1的个数大于1 ④ P(A∩B), P(A∩C), P(B∩C), P(A∩B∩C)

$$6^4 - 5^4 = 1296 - 625 = 671$$



### [例 3-7] 同时投掷四颗骰子的试验

- ① A= 数字之和为15 ② B= 6的个数大于1
- ③ C= 1的个数大于1



13/31

13

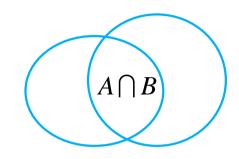
자료구조및실습 1장. 자료구조와 알고리즘

## 3.2.2 概率的运算



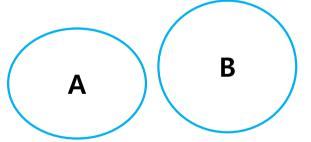
• 和事件的概率计算

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



- 和事件的概率计算 2
  - 若事件A与B相互排斥,由于A∩B=Ø, P(Ø)=0,因此,

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





#### [例 3-8] 投掷四颗骰子的试验

A =数字之和大于15, B = 6的个数大于1, C = 1的个数大于1

(1)  $P(A \cup B)$  (2)  $P(A \cup C)$  (3)  $P(B \cup C)$  (4)  $P(A \cup B \cup C)$ 

A=575, B=671, C=671, AB=453, AC=140, BC=302, ABC=124

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{575 + 671 - 453}{1296} = \frac{793}{1296} \doteq 0.6119$$

$$P(A \cup C) = P(A) + P(C) - P(A \cap C) = \frac{575 + 671 - 140}{1296} = \frac{1106}{1296} = 0.8534$$

$$P(B \cup C) = P(B) + P(C) - P(B \cap C) = \frac{671 + 671 - 302}{1296} = \frac{1040}{1296} = 0.8025$$

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

$$=\frac{575+671+671-453-140-302+124}{1296}=\frac{1146}{1296}\doteq 0.8843$$



[例 3-9] 在同时投掷四颗骰子的试验中, 出现四个连续数字的概率

$$\{1,2,3,4\}, \{2,3,4,5\}, \{3,4,5,6\} \Rightarrow \frac{3\times(4!)}{6^4} = \frac{3\times24}{1296} = \frac{72}{1296}$$

16/31



#### [例 3-10] 投掷四颗骰子的试验

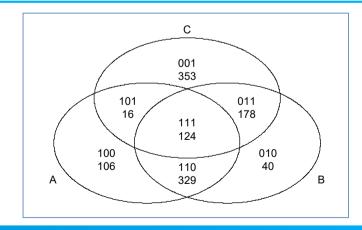
A =数字之和大于15, B = 6的个数大于1, C = 1的个数大于1

(1)  $P(A \cap B^c)$  (2)  $P(A^c \cap B^c)$  (3)  $P(A \cap B \cap C^c)$  (4)  $P(A \cap B^c \cap C)$ 

A=575, B=671, C=671, AB=453, AC=140, BC=302, ABC=124, AuB=793, AuC=1106, BuC=1040, AuBuC=1146

$$P(A \cap B^c) = P(A) - P(A \cap B) = \frac{575 - 453}{1296} = \frac{122}{1296} \square 0.0941$$
$$P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - \frac{793}{1296} = \frac{503}{1296} \square 0.3881$$

$$P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - \frac{793}{1296} = \frac{503}{1296} \square 0.3881$$





#### [例 3-10] 投掷四颗骰子的试验

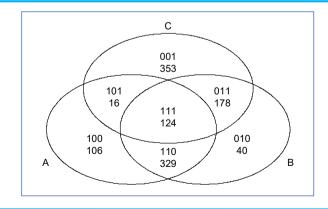
A =数字之和大于15, B = 6的个数大于1, C = 1的个数大于1

①  $P(A \cap B^c)$  ②  $P(A^c \cap B^c)$  ③  $P(A \cap B \cap C^c)$  ④  $P(A \cap B^c \cap C)$ 

A=575, B=671, C=671, AB=453, AC=140, BC=302, ABC=124, AuB=793, AuC=1106, BuC=1040, AuBuC=1146

$$P(A \cap B \cap C^c) = P(A \cap B) - P(A \cap B \cap C) = \frac{453 - 124}{1296} = \frac{329}{1296} \doteq 0.2539$$

$$P(A \cap B^c \cap C) = P(A \cap C) - P(A \cap B \cap C) = \frac{140 - 124}{1296} = \frac{16}{1296} \doteq 0.0123$$



## 3.3 条件概率



## [定义 3-4] 条件概率

- 在给定的某一条件(B)下,某事件(A)发生的概率

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

- 若所有元素的发生概率都相同的话...



### [表 3-1] 关于学科新生的列联表

	男生(M)	女生(F)	合计
文科出身(A)	15	25	40
理科出身(B)	40	20	60
合计	55	45	100

• 边缘概率(marginal probability)与联合概率(joint probability)

	男生(M)	女生(F)	合计
文科出身(A)	15/100	25/100	40/100
理科出身(B)	40/100	20/100	60/100
合计	55/100	45/100	100/100

• 条件概率(conditional probability) 行条件 / 列条件

	男生(M)	女生(F)
文科出身(A)	15/40	25/40
理科出身(B)	40/60	20/60

	男生(M)	女生(F)
文科出身(A)	15/55	25/45
理科出身(B)	40/55	20/45

20/31

자료구조및실습



### [例 3-11] 关于学科新生的列联表

	男生(M)	女生(F)	合计
文科出身(A)	15	25	40
理科出身(B)	40	20	60
合计	55	45	100

• 某学生是文科出身(A)时,该学生是女生(F)的条件概率

	男生(M)	女生(F)
文科出身(A)	15/40	25/40
理科出身(B)	40/60	20/60

$$P(F \mid A) = \frac{P(A \cap F)}{P(A)} = \frac{25/100}{40/100} = \frac{25}{40}$$

• 某学生是女生(F)时, 该学生是文科出身(A)的条件概率

	男生(M)	女生(F)
文科出身(A)	15/55	25/45
理科出身(B)	40/55	20/45

$$P(A \mid F) = \frac{P(A \cap F)}{P(F)} = \frac{25/100}{45/100} = \frac{25}{45}$$



#### [例 3-12] 在投掷四颗骰子的试验中的条件概率

A= 数字之和大于15, B= 6的个数大于1, C= 1的个数大于1 P(A|B)P(A|C) P(A|B∩C) P(A|B∪C)

A=575, B=671, C=671, AB=453, AC=140, BC=302, ABC=124, AuB=793, AuC=1106, BuC=1040, AuBuC=1146

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{453/1296}{671/1296} = \frac{453}{671} \doteq 0.6751$$

$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{140/1296}{671/1296} = \frac{140}{671} \doteq 0.2086$$

$$P(A|B \cap C) = \frac{P(A \cap B \cap C)}{P(B \cap C)} = \frac{124/1296}{302/1296} = \frac{124}{302} \doteq 0.4160$$

$$P(A \cap (B \cup C)) = P(A \cap B) + P(A \cap C) - P(A \cap B \cap C)$$
$$= \frac{453 + 140 - 124}{1296} = \frac{469}{1296}$$

$$\Rightarrow P(A|B \cup C) = \frac{P(A \cap (B \cup C))}{P(B \cup C)} = \frac{469/1296}{1040/1296} = \frac{469}{1040} \doteq 0.4510$$

# 3.3.2 乘法公式(multiplicative law)



• 事件A和事件B同时发生的概率

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

• 推广至n个事件的情况

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2 \cap A_3 \cap \dots \cap A_n \mid A_1)$$

- $= P(A_1)P(A_2 \mid A_1)P(A_3 \cap \cdots \cap A_n \mid A_1 \cap A_2)$
- $= P(A_1)P(A_2 | A_1)P(A_3 | A_1 \cap A_2)P(A_4 \cap \cdots \cap A_n | A_1 \cap A_2 \cap A_3)$

= ...

$$= P(A_1)P(A_2 | A_1)P(A_3 | A_1 \cap A_2) \times \dots \times P(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

23/31

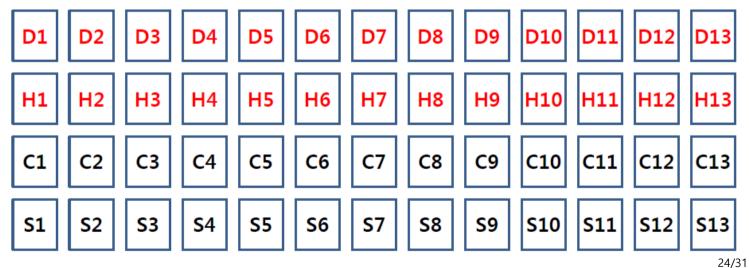


[例 3-13] 在一个罐子内放入了52张标有1到13的四种条纹的卡片,从中随意抽取 四张,所有卡片均为同一条纹的概率

$$P(A_1 \cap A_2 \cap A_3 \cap A_4)$$

$$= P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2)P(A_4 \mid A_1 \cap A_2 \cap A_3)$$

$$=1 \times \frac{12}{51} \times \frac{11}{50} \times \frac{10}{49} = \frac{1,320}{124,950} \doteq 0.0106$$



## 3.3.3 独立事件(independent events)。

• 对于构成样本空间的事件A和事件B, 若两事件相互独立的话

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

• 即,在P(B|A)中,由于作为条件的事件A对事件B不造成任何影响,因此事件A和事件B相互独立的关系成立。

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



## [定义 3-5] 独立事件(independent events)

- $P(A \cap B) = P(A)P(B)$
- 若相互独立  $\rightarrow P(A) = P(A \mid B) = \frac{P(A \cap B)}{P(B)}$
- 若P(A∩B)=P(A)P(B)成立,

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

■ 扩展(n个事件)

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2)P(A_3)\cdots P(A_n)$$

26/31



### [例 3-14] 关于学科新生的列联表

	男生(M)	女生(F)	合计
文科出身(A)	15	25	40
理科出身(B)	40	20	60
合计	55	45	100

■ 某学生是文科出身的事件(A)和该学生是女生的事件(F)间的独立性判定

$$P(A) = 0.40, P(F) = 0.45, P(A \cap F) = 0.25$$
  
 $P(A)P(F) = 0.40 \times 0.45 = 0.18 \neq P(A \cap F) = 0.25$ 



[例 3-15]\* 从52张标有1到13的四种条纹的卡片中,随机抽取4张时,判定以下两事件是否相互独立

事件 A = 四张卡片的条纹都不一样

事件 B = 四张卡片的数字都不一样

- ① 与之前抽出的卡片条纹不同,事件 Ai
- ② 与之前抽出的卡片数字不同,事件 Bi
- ③ 与之前抽出的卡片条纹不同,数字也不同,事件 Ci

$$P(A) = P(A_1 \cap A_2 \cap A_3 \cap A_4) = 1 \times \frac{39}{51} \times \frac{26}{50} \times \frac{13}{49} = \frac{13,182}{124,950} \doteq 0.1055$$

$$P(B) = P(B_1 \cap B_2 \cap B_3 \cap B_4) = 1 \times \frac{48}{51} \times \frac{44}{50} \times \frac{40}{49} = \frac{84,480}{124,950} \doteq 0.6761$$

$$P(A \cap B) = P(C_1 \cap C_2 \cap C_3 \cap C_4) = 1 \times \frac{36}{51} \times \frac{22}{50} \times \frac{10}{49} = \frac{7,920}{124,950} \doteq 0.0634$$

$$\Rightarrow P(A)P(B) = \frac{13,182}{124,950} \times \frac{84,480}{124,950} \doteq 0.0713 \neq P(A \cap B)$$

$$P(A) = P(A_1 \cap A_2 \cap A_3 \cap A_4) = 1 \times \frac{39}{51} \times \frac{26}{50} \times \frac{13}{49}$$

$$P(B) = P(B_1 \cap B_2 \cap B_3 \cap B_4) = 1 \times \frac{48}{51} \times \frac{44}{50} \times \frac{40}{49}$$

$$P(A \cap B) = P(C_1 \cap C_2 \cap C_3 \cap C_4) = 1 \times \frac{36}{51} \times \frac{22}{50} \times \frac{10}{49}$$

$$D1 \quad D2 \quad D3 \quad D4 \quad D5 \quad D6 \quad D7 \quad D8 \quad D9 \quad D10 \quad D11 \quad D12 \quad D13$$

$$H1 \quad H2 \quad H3 \quad H4 \quad H5 \quad H6 \quad H7 \quad H8 \quad H9 \quad H10 \quad H11 \quad H12 \quad H13$$

$$C1 \quad C2 \quad C3 \quad C4 \quad C5 \quad C6 \quad C7 \quad C8 \quad C9 \quad C10 \quad C11 \quad C12 \quad C13$$

$$S1 \quad S2 \quad S3 \quad S4 \quad S5 \quad S6 \quad S7 \quad S8 \quad S9 \quad S10 \quad S11 \quad S12 \quad S13$$



[例 3-16] 在前面的[例 3-15]中出现了条纹和数字之间相互影响的结果,让人多少有些意外。为了支撑这一结果,请判定以下两事件是否相互独立。

事件 A= 4张卡片的条纹相同 事件 B= 4张卡片的数字相同

- P(A)>0, P(B)>0
- 由于相同条纹和相同数字不可能出现两次以上,因此属于事件A∩B 的元素的个数为0
- P(A)P(B)>0=P(A∩B),所以不是独立事件

[例 3-17] 在不合格率为0.1的工序中生产的产品中,随机抽取10个

- ① 均为合格品的概率
- ② 1个不合格品, 9个合格品的概率

$$P(A) = 0.9^{10} \doteq 0.3487$$

$$P(B) = 10 \times (0.1 \times 0.9^9) \doteq 0.3847$$

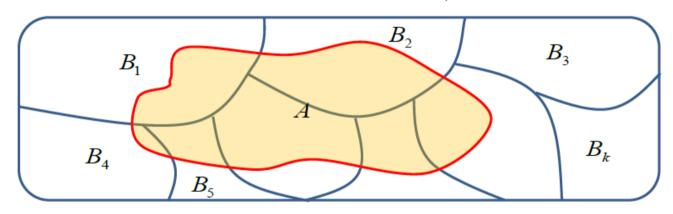


## 3.4 贝叶斯定理



## [定理 3-1] 全概率公式(theorem of total probability)

■ 利用相互排斥事件对样本空间S进行分割(partition)



$$A = (B_1 \cap A) \cup (B_2 \cap A) \cup \cdots \cup (B_k \cap A)$$

$$P(A) = P(B_1 \cap A) + P(B_2 \cap A) + \dots + P(B_k \cap A)$$

$$P(B_i \cap A) = P(B_i)P(A \mid B_i)$$

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i) P(A \mid B_i)$$



# [例 3-18] 四条生产线上的生产比例和不合格率 随机抽取一个产品时,该产品为不合格品(F)的概率

生产线	Α	В	C	D	
生产比例	20%	40%	30%	10%	⇒100%
不合格率	0.04	0.02	0.01	0.05	$\Leftrightarrow P(F \mid \Box)$

$$P(F) = P(F \cap A) + P(F \cap B) + P(F \cap C) + P(F \cap D)$$
  
=  $P(A)P(F \mid A) + P(B)P(F \mid B) + P(C)P(F \mid C) + P(D)P(F \mid D)$   
=  $0.2 \times 0.04 + 0.4 \times 0.02 + 0.3 \times 0.01 + 0.1 \times 0.05 = 0.024$   
# 生产比例, 各生产线的不合格率

prior <- c(0.2, 0.4, 0.3, 0.1) cond <- c(4, 2, 1, 5)/100

# 不合格品在各生产线上发生的概率

tot <- prior\*cond; tot [1] 0.008 0.008 0.003 0.005

# 合计 sum(tot)

[1] 0.024

32/31

## 3.4 贝叶斯定理



## [定理 3-2] 贝叶斯法则(Bayes theorem)

■ 利用相互排斥事件对样本空间S进行分割(partition)

$$P(B_r | A) = \frac{P(B_r)P(A | B_r)}{\sum_{i=1}^k P(B_i)P(A | B_i)}$$

$$P(A) = \sum_{i=1}^k P(B_i \cap A) = \sum_{i=1}^k P(B_i)P(A | B_i)$$

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

$$= \frac{P(B_r)P(A | B_r)}{\sum_{i=1}^k P(B_i \cap A)} = \frac{P(B_r)P(A | B_r)}{\sum_{i=1}^k P(B_i)P(A | B_i)}$$



#### [例 3-19] 出现一个不合格品时, 其出自生产线A, B, C, D的概率

生产线	Α	В	C	D
生产比例	20%	40%	30%	10%
不合格率	0.04	0.02	0.01	0.05

$$P(F) = P(F \cap A) + P(F \cap B) + P(F \cap C) + P(F \cap D)$$

$$= P(A)P(F \mid A) + P(B)P(F \mid B) + P(C)P(F \mid C) + P(D)P(F \mid D)$$

$$= 0.2 \times 0.04 + 0.4 \times 0.02 + 0.3 \times 0.01 + 0.1 \times 0.05 = 0.024$$

$$\Rightarrow P(A \mid F) = \frac{P(A \cap F)}{P(F)} = \frac{0.2 \times 0.04}{0.024} = \frac{1}{3}$$

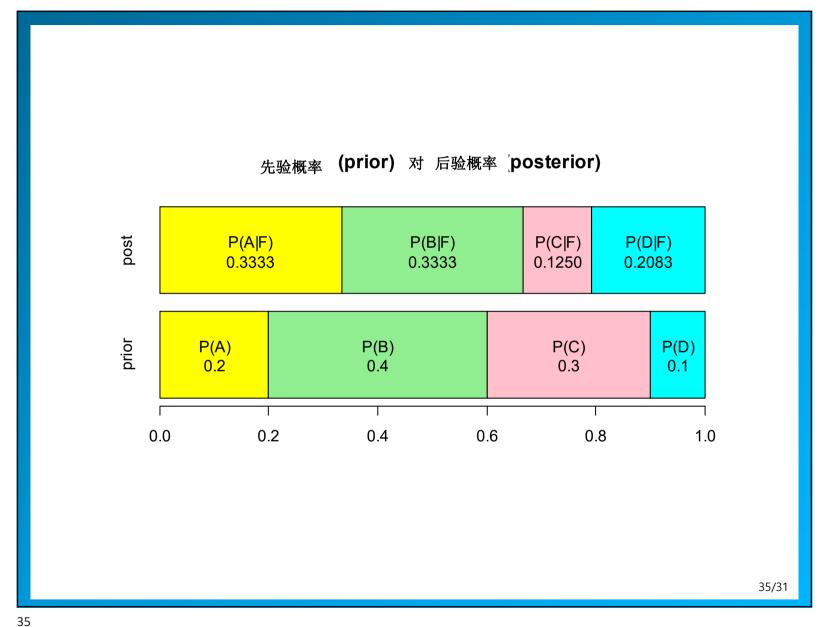
$$\Rightarrow P(B \mid F) = \frac{P(B \cap F)}{P(F)} = \frac{0.4 \times 0.02}{0.024} = \frac{1}{3}$$

$$\Rightarrow P(C \mid F) = \frac{P(C \cap F)}{P(F)} = \frac{0.3 \times 0.01}{0.024} = \frac{1}{8}$$

$$\Rightarrow P(D \mid F) = \frac{P(D \cap F)}{P(F)} = \frac{0.1 \times 0.05}{0.024} = \frac{5}{24}$$

34/31

자료구조및실습



자료구조및실습 1장. 자료구조와 알고리즘