浙江大学
ZHEJIANG UNIVERSITY

第7讲　全概率公式与贝叶斯公式

在第5讲有一个抽签问题例子：

一袋中有a个白球，b个黄球，记a+b=n.

设每次摸到各球的概率相等，每次从袋中摸一球，

不放回地摸n次. 则第k次摸到白球的概率均为a/n.

现在用另一种方法计算第2次取到白球的概率.

解：设A_i表示第i次取到白球，$i=1,2.$

$$P(A_2)=P[(A_1\cup\overline{A}_1)A_2]=P(A_1A_2\cup\overline{A}_1A_2)$$

$$=P(A_1A_2)+P(\overline{A}_1A_2)$$

$$P(A_1A_2)=P(A_1)P(A_2\left|A_1\right.)=\frac{a}{n}\times\frac{a-1}{n-1}$$

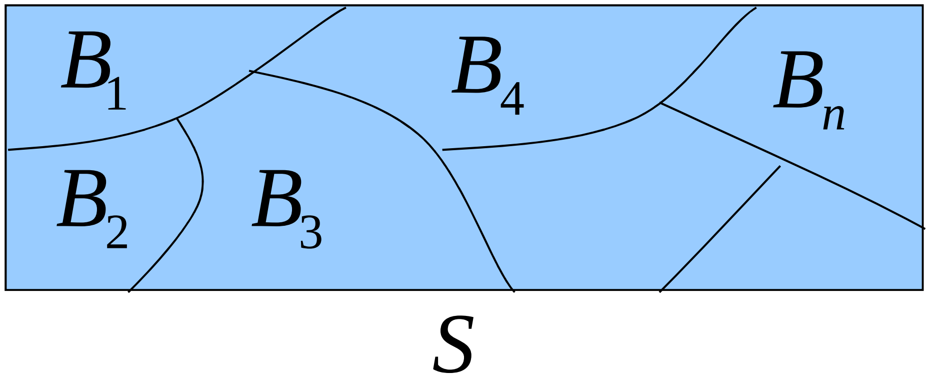
$$P(\overline{A}_1A_2)=P(\overline{A}_1)P(A_2\left|\overline{A}_1\right.)=\frac{b}{n}\times\frac{a}{n-1}$$

$$P(A_2)=P(A_1)P(A_2\left|A_1\right.)+P(\overline{A}_1)P(A_2\left|\overline{A}_1\right.)=\frac{a}{n}.$$

定义: 称$B_1, B_2, \cdots, B_n$为S的一个**划分**, 若

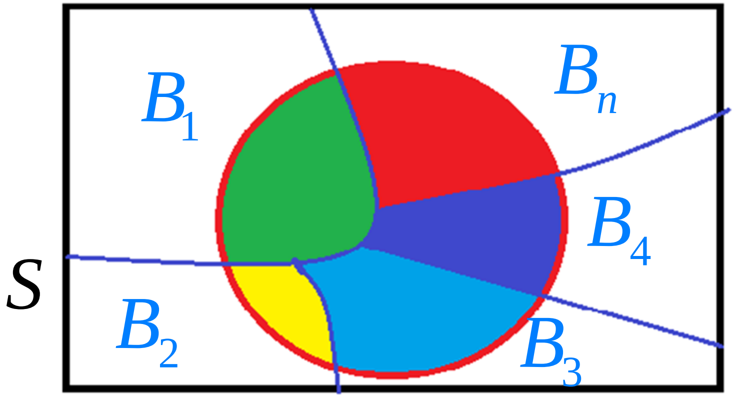
(i) **不漏** $B_1 \cup B_2 \cup \cdots \cup B_n = S$,

(ii) **不重** $B_i B_j = \varnothing, \ \ i \neq j.$



定理：设$B_1, B_2, \cdots, B_n$为S的一个划分且$P(B_i) > 0$. 则有**全概率公式**：

$$P(A) = \sum_{j=1}^{n} P(B_j) \cdot P(A \mid B_j)$$



证明：$A = AS$

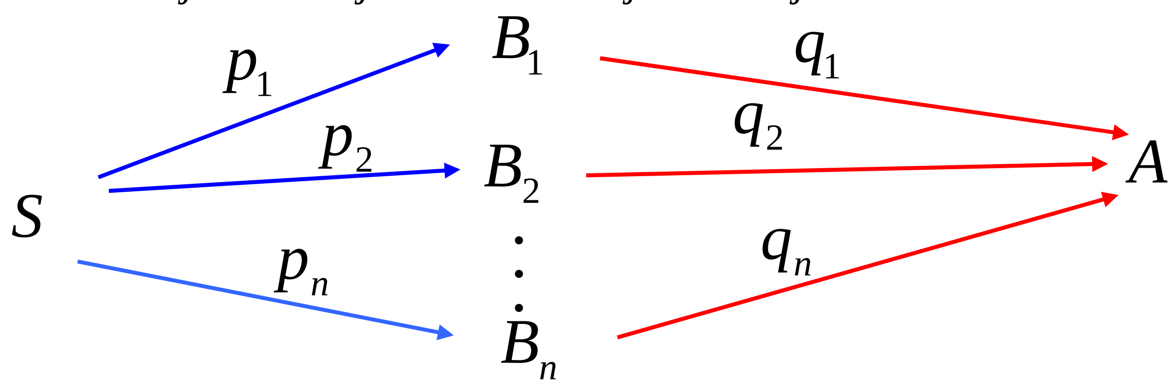
$= AB_1 \cup AB_2 \cup \cdots \cup AB_n$

AB_i与AB_j不相容$(i \neq j)$

$$\therefore P(A) = \sum_{j=1}^{n} P(AB_j)$$

$$= \sum_{j=1}^{n} P(B_j) \cdot P(A \mid B_j)$$

设$P(B_j)=p_j, P(A|B_j)=q_j, j=1,2,\ldots,n.$



则 $P(A)=\sum_{j=1}^{n} p_j q_j.$

注意：在运用全概率公式时，关键是构造合适的划分.

定理: 设$B_1, B_2, \cdots, B_n$为S的一个划分且$P(B_i)>0$.

对$P(A)>0$有**Bayes公式**:

$$P(B_i \mid A) = \frac{P(B_i)P(A \mid B_i)}{\sum_{j=1}^{n} P(B_j)P(A \mid B_j)} = \frac{p_i q_i}{\sum_{j=1}^{n} p_j q_j}$$

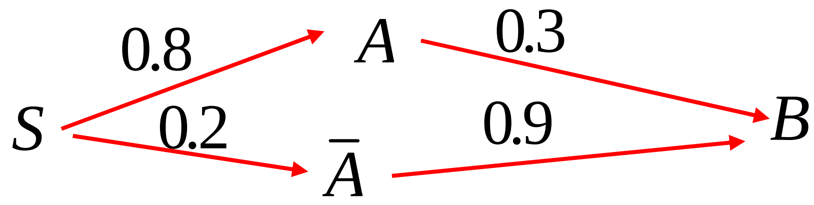
例1：一小学举办家长开放日,欢迎家长参加活动.小明的母亲参加的概率为80%.若母亲参加，则父亲参加的概率为30%；若母亲不参加,则父亲参加的概率为90%.

(1) 求父母都参加的概率；

(2) 求父亲参加的概率；

(3) 在已知父亲参加的条件下，求母亲参加的概率.

解：设$A=\{$母亲参加$\}$, $B=\{$父亲参加$\}$.



由题意, $P(A)=0.80$,

$P(B|A)=0.30, P(B|\bar{A})=0.90.$

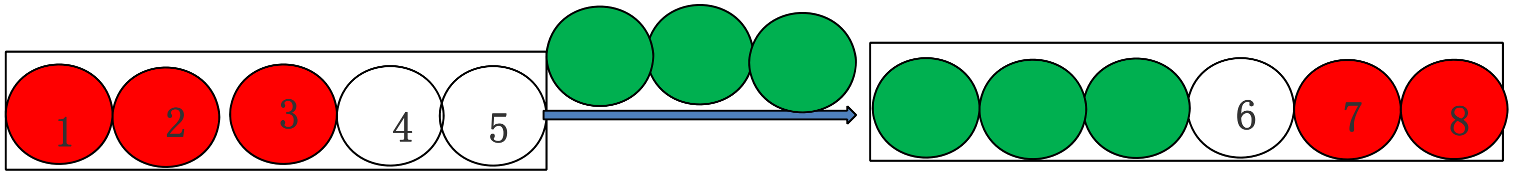
(1) $P(AB)=P(A)\,P(B|A)=0.24$

(2)由全概率公式：

$$P(B)=P(A)\,P(B|A)+P(\bar{A})P(B|\bar{A})$$
$$=0.8\times 0.3+0.2\times 0.9=42\%$$

(3)由**Bayes**公式： $P(A|B)=\dfrac{P(AB)}{P(B)}=\dfrac{P(A)P(B|A)}{P(B)}=\dfrac{4}{7}$

例2：有甲乙两盒，甲盒有3个红球2个白球，乙盒有2个红球，1个白球。先从甲盒中采用不放回抽样取3球放入乙盒，再从乙盒中取一个球，求取到的是红球的概率。



解：设$A_i=\{$从甲盒中取到i个红球$\}, i=1,2,3.$

$$P(A_1)=\frac{C_3^1C_2^2}{C_5^3}=0.3,\ \ P(A_2)=\frac{C_3^2C_2^1}{C_5^3}=0.6,\ P(A_3)=\frac{C_3^3}{C_5^3}=0.1,$$

$$P(B\mid A_1)=1/2,\ \ P(B\mid A_2)=2/3, P(B\mid A_3)=5/6.$$

$$\begin{tikzcd}
 & A_1 \arrow[dr, "1/2"] & \\
S \arrow[ur, "0.3"] \arrow[r, "0.6"] \arrow[dr, "0.1"'] & A_2 \arrow[r, "2/3"] & B \\
 & A_3 \arrow[ur, "5/6"'] &
\end{tikzcd}$$

$B=\{$从乙盒中取到红球$\}$

$$P(B)=P(A_1)P(B\mid A_1)+P(A_2)P(B\mid A_2)+P(A_3)P(B\mid A_3)$$

$$=\frac{3}{10}\times\frac{1}{2}+\frac{6}{10}\times\frac{2}{3}+\frac{1}{10}\times\frac{5}{6}=\frac{19}{30}.$$

例3：根据以往的临床记录某种诊断癌症的试验具有5%的假阳性及3%的假阴性：

若设A={试验反应是阳性}，C={被诊断患有癌症}，

则有：$P(A|\overline{C})=5\%, P(\overline{A}|C)=3\%,$

已知某一群体$P(C)=0.005$，问这种方法能否用于普查？

解：

$$\begin{tikzcd}[column sep=large] & C \arrow[dr, "0.97"] & \\ S \arrow[ur, "0.005"] \arrow[dr, "0.995"'] & & A \\ & \overline{C} \arrow[ur, "0.05"'] & \end{tikzcd}$$

由Bayes公式：

$$P(C\mid A)=\frac{P(CA)}{P(A)}=\frac{P(C)P(A|C)}{P(C)P(A|C)+P(\overline{C})P(A|\overline{C})}=0.089$$

若用于普查，100个阳性病人中被诊断患有癌症的大约有8.9个，所以不宜用于普查.如果发现检验结果为阳性，则需要作进一步的检查.

若$P(C)$很大, 比如$P(C)=0.8$,则

$$
\begin{tikzcd}
 & C \arrow[dr, "0.97"] & \\
S \arrow[ur, "0.8"] \arrow[dr, "0.2"'] & & A \\
 & \overline{C} \arrow[ur, "0.05"'] &
\end{tikzcd}
$$

$$P(C \mid A)=\frac{0.8\times 0.97}{0.8\times 0.97+0.2\times 0.05}=0.987.$$

说明此方法在肿瘤医院等专门医院适用.