§1·4 条件概率与统计独立性

一 条件概率

我们知道，对任何概率问题的讨论，都必须建立起一个概率空间$(\Omega,\mathbb{F},\mathrm{P})$，并且是在$P(\Omega)=1$的前提下进行讨论. 但有时除了这个总前提之外，还会出现附加前提.

例如，抛掷一枚均匀的骰子，已知掷出的点数为奇数（记为事件B），要求出点数大于1 (记为事件A)的概率.

显然，这个附加前提"掷出的点数为奇数"会影响事件A发生的概率.

下面求"在事件 B 发生的前提下，事件 A 发生的概率"（以 $P(A|B)$ 表示，称之为**条件概率**）

$$\Omega=\{1,2,3,4,5,6\},\quad B=\{1,3,5\},\quad A=\{2,3,4,5,6\}.$$

由于 B 已经发生，因此掷出的所有可能结果只有3种情况：点数为1，点数为3，点数为5．若事件 A 再发生，则事件 $AB=\{3,5\}$ 发生．这相当于在样本空间 $\Omega'=B=\{1,3,5\}$ 中，求事件 AB 发生的概率, 这也是古典概型问题．因此

$$P(A|B)=\frac{|AB|}{|B|}=\frac{2}{3}.$$

那么在原概率空间Ω中，如何求$P(A|B)$呢？我们有

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

这个例子启发我们：可以以$P(AB)$与$P(B)$的比作为条件概率$P(A|B)$的定义.

定义 1 设$(\Omega,\mathbb{F},\mathrm{P})$是一个概率空间, $B\in\mathbb{F}$. 而且$P(B)>0$，则对任意$A\in\mathbb{F}$，记

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

并称$P(A|B)$为在事件B发生的前提下, 事件A发生的**条件概率**.

可验证上述的$P(A|B)$满足非负性、规范性、可列可加性.

（1）非负性　　对任何事件A，　有　$P(A|B)\geq 0$;

（2）规范性　　$P(\Omega|B)=1$;

（3）可列可加性　若$A_1,A_2,\cdots$两两互不相容，则

$$P\left(\sum_{n=1}^{+\infty}A_n\,\middle|\,B\right)=\sum_{n=1}^{+\infty}P(A_n|B).$$

因此$\left(\Omega,\mathbb{F},P(\cdot|B)\right)$也是概率空间. 对条件概率$P(A|B)$也可由三条基本性质导出概率的其他一些性质.

在具体计算 $P(A|B)$ 时，可以用公式右端来求，也可以像刚才的例子那样，直接从缩小了的样本空间来求，后一种求法有时更方便、实用.

【例1】某班级共有100名同学，其中男生60名，女生40名.在期末的高等数学考试中有8人考试不及格，其中有6名男生,2名女生. 若A表示同学是男生这一事件,B表示考试不及格这一事件. 试求

(1) 条件概率 $P(B|A)$，$P(A|B)$，$P(\bar{A}|B)$.

(2)解释条件概率$P(B|A)$，$P(B)$的含义,并比较其大小.

解：(1)
$$P(B|A)=\frac{P(AB)}{P(A)}=\frac{6/100}{60/100}=10\%.$$

$$P(A|B)=\frac{P(AB)}{P(B)}=\frac{6/100}{8/100}=75\%$$

$$P(\bar{A}|B)=1-P(A|B)=25\%.$$

(2) 条件概率$P(B|A)$表示男生的不及格率，$P(B)$表示班级的不及格率. 由于

$$P(B|A)=10\%,\ P(B)=8\%;\ P(B|A)>P(B).$$

这表明男生的不及格率高于班级的不及格率.

下面给出条件概率特有的三个非常实用的公式：乘法公式、全概率公式、贝叶斯公式. 这些公式可以帮助我们计算一些复杂事件的概率.

二 乘法公式

根据条件概率的定义，不难得下列乘法公式：

定理 1 设A, B与$A_1, A_2, \cdots, A_n$都是概率空间$(\Omega, \mathbb{F}, \mathrm{P})$中的事件.

(1) $P(AB) = P(B)P(A|B)$（这里要求$P(B) > 0$）.

$P(AB) = P(A)P(B|A)$（这里要求$P(A) > 0$）.

(2) $$P(A_1A_2\cdots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1})$$

（这里要求$P(A_1A_2\cdots A_{n-1}) > 0$）

证明　根据条件概率的定义，即得(1).

由于：$P(A_1) \geq P(A_1A_2) \geq \cdots \geq P(A_1A_2\cdots A_{n-1}) > 0$

因此(2)中的各条件概率均有意义，根据条件概率的定义，(2)式的右边等于

$$P(A_1)\cdot\frac{P(A_1A_2)}{P(A_1)}\cdot\frac{P(A_1A_2A_3)}{P(A_1A_2)}\cdots\frac{P(A_1A_2\cdots A_n)}{P(A_1A_2\cdots A_{n-1})}$$

$$=P(A_1A_2\cdots A_n)$$

【例2】 (坛子模型)设坛子中有a个白球及b个黑球，每次随机取出一个球，取出后将原球放回，并加入c个同色球和d个异色球. 再摸第二次，这样下去共摸了n次. 记A_i 表示“第i 次取出的是白球”，B_j表示“第j 次取出的黑球”，试求.

(1) $P(A_1B_2B_3)$, $P(B_1A_2B_3)$, $P(B_1B_2A_3)$,

(2) $P(A_1A_2\cdots A_{n_1}B_{n_1+1}B_{n_1+2}\cdots B_n)$.

解：(1) $P(A_1B_2B_3)=P(A_1)P(B_2\,|\,A_1)P(B_3\,|\,A_1B_2)$

$$=\frac{a}{a+b}\cdot\frac{b+d}{a+b+c+d}\cdot\frac{b+d+c}{a+b+2c+2d};$$

$P(B_1A_2B_3)=P(B_1)P(A_2\,|\,B_1)P(B_3\,|\,B_1A_2)$

$$=\frac{b}{a+b}\cdot\frac{a+d}{a+b+c+d}\cdot\frac{b+d+c}{a+b+2c+2d};$$

$P(B_1B_2A_3)=P(B_1)P(B_2\,|\,B_1)P(A_3\,|\,B_1B_2)$

$$=\frac{b}{a+b}\cdot\frac{b+c}{a+b+c+d}\cdot\frac{a+2d}{a+b+2c+2d}.$$

以上概率与白球在第几次被取出有关.

$$(2)\qquad P(A_1)P(A_2 \mid A_1)\cdots P(A_{n_1} \mid A_1A_2\cdots A_{n_1-1})$$

$$\cdot P(B_{n_1+1} \mid A_1A_2\cdots A_{n_1})\cdots\cdot P(B_n \mid A_1A_2\cdots A_{n_1}B_{n_1+1}\cdots B_{n-1})$$

$$=\frac{a}{a+b}\cdot\frac{a+c}{a+b+(c+d)}\cdots\frac{a+(n_1-1)c}{a+b+(n_1-1)(c+d)}\cdot\frac{b+n_1d}{a+b+n_1(c+d)}$$

$$\cdot\frac{b+n_1d+c}{a+b+(n_1+1)(c+d)}\cdots\frac{b+n_1d+(n-n_1-1)c}{a+b+(n-1)(c+d)}.$$

坛子模型也称为波利亚模型，这个模型有如下各种变化：

（1） 当$c=0,d=0$ 时，即为返回抽样. 此时前次抽取结果不会影响后次的抽样结果. 故上述三个概率都是一样的，本例中有

$$P(A_1B_2B_3)=P(B_1A_2B_3)=P(B_1B_2A_3)=\frac{ab^2}{(a+b)^3}$$

（2）当c = -1,d=0 时，即为不返回抽样．此时前次抽取结果会影响后次的抽样结果．但只要抽取的白球与黑球个数一样，则概率与抽出球的次序无关，它们都是一样的，本例中有

$$P(A_1B_2B_3)=P(B_1A_2B_3)=P(B_1B_2A_3)$$
$$=\frac{ab(b-1)}{(a+b)(a+b-1)(a+b-2)}$$

（3）当c >0, d=0 时，即为传染病模型．此时每次取出球后就会增加下一次取到同色球的概率，或换句话说，每次发现一个传染病患者，以后都会增加再传染的概率．与（1）、（2）一样，以上三个概率相等，且都等于

$$P(A_1B_2B_3)=P(B_1A_2B_3)=P(B_1B_2A_3)$$
$$=\frac{ab(b+c)}{(a+b)(a+b+c)(a+2b+2c)}$$

（4）当$c=0$, $d>0$时，称为安全模型．此模型可解释为：每当事故发生了（黑球被取出），安全工作就抓紧一些，下次再发生的概率就会减少；而当事故没有发生时（白球被取出），安全工作就放松一些，下次再发生事故的概率就会增大．在这种场合，上述三个概率分别为

$$P(A_1B_2B_3)=P(A_1)P(B_2|A_1)P(B_3|A_1B_2)$$
$$=\frac{a}{a+b}\cdot\frac{b+d}{a+b+d}\cdot\frac{b+d}{a+b+2d}$$

$$P(B_1A_2B_3)=\frac{b}{a+b}\cdot\frac{a+d}{a+b+d}\cdot\frac{b+d}{a+b+2d}$$

$$P(B_1B_2A_3)=\frac{b}{a+b}\cdot\frac{b}{a+b+d}\cdot\frac{a+2d}{a+b+2d}$$

三　事件的独立性

独立性是概率论又一个重要概念，利用独立性可以简化概率的计算. 下面先讨论两个事件的独立性，然后讨论多个事件的相互独立性，最后讨论试验之间的独立性.

两个事件的独立性

设A和B是概率空间$(\Omega, \mathbb{F}, \mathrm{P})$中的两个事件.一般来说，$P(A)$与$P(A|B)$，$P(B)$与$P(B|A)$是不同的. 见上面例1.

条件概率$P(B|A)$表示男生的不及格率，$P(B)$表示班级的不及格率，通常二者是不同的；同样$P(A|B)$表示不及格的学生中男生所占的比例，而$P(A)$表示整个班级的学生中男生所占的比例，二者一般也不相等。

$$P(B|A) \neq P(B), \quad P(A|B) \neq P(A)$$

这反映了在一般情况下，无条件概率不等于条件概率，但也有例外.

例如：如果例1中的不及格人数修改为5人，其中男生3人，女生2人. 不难推得

$$P(B|A)=P(B)=5\%，P(A|B)=P(A)=60\%.$$

这种情形意味着：事件A发生对事件B发生的概率没有影响，事件B发生对事件A发生的概率也没有影响. 我们把这种现象称之为相互独立. 根据乘法公式可知

$$P(B|A)=P(B) \quad \text{等价于} \quad P(AB)=P(A)P(B)$$

同样 $P(A|B)=P(A)$ 等价于 $P(AB)=P(A)P(B)$

因此，我们有如下的定义.

定义 2 设A和B是概率空间$(\Omega,\mathbb{F},\mathrm{P})$中的两个事件，如果有

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

则称事件A与事件B相互独立，简称独立. 否则称A与B不独立或相依. 按照这个定义，必然事件和不可能事件与任何事件相互独立.

事件A与事件B相互独立意味着：任何一个事件的发生与否不影响另一个事件的发生的概率.

由独立性的定义容易推出下述结论.

推论1　若事件A与B相互独立，则下列各对事件也相互独立：

$$\{\bar{A},\ B\},\qquad \{A,\ \bar{B}\},\qquad \{\bar{A},\ \bar{B}\}$$

证明　事件A与事件B相互独立，有

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

由于　$\bar{A}B=B-AB,\ AB\subset B$　　所以

$$P(\bar{A}B)=P(B)-P(AB)=P(B)-P(A)P(B)$$
$$=\left[1-P(A)\right]P(B)=P(\bar{A})P(B)$$

因而事件$\bar{A}$与事件B相互独立，其他情况可类似证明.

在许多实际问题中，我们可以通过问题的本身性质来判断事件的独立性，例如济南下雨（事件A）与纽约下雨（事件B）可以认为相互独立。有时甚至可以人为地假定独立性的存在, 这种假定的合理性往往可以通过计算结果与实际情况的吻合程度来检验.

【例3】 若$P(A)>0$, $P(B)>0$. 则当它们互不相容时，它们必不互相独立；反之，如果它们相互独立时，则它们一定相容.

证明 A与B相互独立 $\Rightarrow P(AB)=P(A)P(B)>0$

$\Rightarrow A$与B相容.

A与B不相容 $\Rightarrow 0=P(AB)\neq P(A)P(B)\Rightarrow$ 不互相独立.

多个事件的独立性

多个事件的独立性是在建立在两个事件独立性的基础上的，但要复杂许多.首先研究三个事件的相互独立性的定义.

定义3　设A、B、C是同一概率空间中的三个事件.

如果有

$$\begin{cases} P(AB) = P(A)P(B) \\ P(AC) = P(A)P(C) \\ P(BC) = P(B)P(C) \end{cases}$$

则称A、B、C两两独立.　若还有

$$P(ABC) = P(A)P(B)P(C)$$

则称A、B、C相互独立.

由此我们可以定义三个以上事件的相互独立性.

定义4 设有n个事件$A_1, A_2, \cdots, A_n$，如果对任意的$1 \le i < j < k < \cdots \le n$. 以下$2^n - n - 1$个等式均成立.

$$\begin{cases} P(A_i A_j) = P(A_i)P(A_j) \\ P(A_i A_j A_k) = P(A_i)P(A_j)P(A_k) \\ \qquad \cdots \quad \cdots \quad \cdots \quad \cdots \\ P(A_1 A_2 \cdots A_n) = P(A_1)P(A_2)\cdots P(A_n) \end{cases}$$

则称n个事件$A_1, A_2, \cdots, A_n$**相互独立.**

从上述定义可以看出，n个相互独立的事件中的任意一部分仍是相互独立的，与推论2.2.1类似，可以证明：将互相独立的n个事件中的任意一部分换成对立事件，所得到的n个事件仍是相互独立的.

定义5 设事件序列$A_1, A_2, \cdots, A_n, \cdots$，如果其中任意有限个事件都相互独立，则称事件序列 $A_1, A_2, \cdots, A_n, \cdots$，是**相互独立**的. 并称事件序列 $A_1, A_2, \cdots, A_n, \cdots$ 是**独立事件序列.**

【例4】 设 $\Omega = (0,1)$，事件域 $\mathbb{F}$ 由Ω的所有博雷尔子集组成，P为勒贝格测度。令

$$A = \left(0, \frac{1}{2}\right),\ B = \left(\frac{1}{4}, \frac{3}{4}\right),\ C = \left(0, \frac{1}{4}\right) \cup \left(\frac{1}{2}, \frac{3}{4}\right)$$

试讨论事件A, B, C是否相互独立.

解： 由题设不难得

$$P(A) = P(B) = P(C) = 1/2$$

$$P(AB) = P(AC) = P(BC) = 1/4$$

$$P(ABC)=P(\phi)=0\neq P(A)P(B)P(C)$$

因而事件A,B,C两两独立，但不相互独立.

思考题： 上例中A,B不变，而

$$C_1=\left(\frac{3}{8},\frac{7}{8}\right),\qquad C_2=\left(\frac{1}{8},\frac{3}{8}\right)\cup\left(\frac{5}{8},\frac{7}{8}\right)$$

试讨论

（1）事件A,B,C_1是否相互独立？事件A,B,C_2是否相互独立？

（2）根据讨论的结果及例3的结论，给出三个事件相互独立的4个条件之间的相互关系.

思考题答案:

$$P(A)=P(B)=P(C_1)=1/2$$

$$P(AB)=1/4,\quad P(AC_1)=1/8,\quad P(BC_1)=3/8$$

$$P(ABC_1)=P(A)P(B)P(C_1)=1/8$$

因而事件A,B,C_1不相互独立.

$$P(A)=P(B)=P(C_2)=1/2$$

$$P(AB)=P(AC)=P(BC_2)=1/4$$

$$P(ABC_2)=P(A)P(B)P(C_2)=1/8$$

因而事件A,B,C_2相互独立.

练习题:

设事件A,B,C相互独立，试证明$A\cup B$与C相互独立.

三　独立性场合下的概率计算

当$A_1,A_2,\cdots,A_n$是同一概率空间$(\Omega,\mathbb{F},\mathrm{P})$中的$n$个相互独立的事件时，有

（1）$A_1,A_2,\cdots,A_n$至少有一个发生的概率为

$$P(\bigcup_{i=1}^{n}A_i)=1-\prod_{i=1}^{n}(1-P(A_i))$$

（2）$A_1,A_2,\cdots,A_n$都发生的概率为

$$P(\bigcap_{i=1}^{n}A_i)=\prod_{i=1}^{n}P(A_i)$$

证明（1）$A_1,A_2,\cdots,A_n$相互独立时，$\bar{A}_1,\bar{A}_2,\cdots,\bar{A}_n$也相互独立时，再由De Morgan定理即得到.

(2) 由独立性定义即得.

注： n个事件并的概率:

$$P(A_1\cup A_2\cup\cdots\cup A_n)=$$

$$\begin{cases}\sum\limits_{i=1}^{n}P(A_i) & A_1,A_2,\cdots A_n\text{互不相容}\\[2em] 1-\prod\limits_{i=1}^{n}P(\overline{A_i}) & A_1,A_2,\cdots A_n\text{相互独立}\\[2em] \sum\limits_{i=1}^{n}P(A_i)-\sum\limits_{1\le i<j\le n}P(A_iA_j)+\cdots+(-1)^nP(A_1\cdots A_n) & \\ & A_1,A_2,\cdots A_n\text{不相容,不独立}\end{cases}$$

【例5】设某种型号的高射炮命中率为0.6,若干门炮同时发射(每炮射一发). 问:欲以99%以上的把握击中敌机,至少配备几门高炮?

解:设至少需要n 门炮

A_i表示“第i 门炮击中敌机” $i=1,2\cdots n.$

B表示“敌机被击中”.

则 $B=A_1\cup A_2\cup\cdots\cup A_n$

$$P(B)=P(\bigcup_{i=1}^{n}A_i)=1-P(\overline{\bigcup_{i=1}^{n}A_i})=1-P(\bigcap_{i=1}^{n}\overline{A_i})$$

$$=1-P(\overline{A}_1)P(\overline{A}_2)\cdots P(\overline{A}_n)=1-(0.4)^n$$

由题意： $1-(0.4)^n \geq 0.99 \therefore$

$$(0.4)^n \leq 0.01$$

$$n \geq \frac{\lg 0.01}{\lg 0.4} \approx 5.027$$

至少需配置六门炮才能以99%以上的把握击中敌机.

【例6】有两名选手比赛射击，轮流对同一目标进行射击，甲命中目标的概率为α，乙命中目标的概率为β. 甲先射，谁先命中谁得胜. 问甲、乙两人获胜的概率各为多少？

解： 记事件A_i为“第 i 次射击命中目标”，因为甲先射，所以事件“甲获胜”可以表示为

$$A_1+\bar{A}_1\bar{A}_2A_3+\bar{A}_1\bar{A}_2\bar{A}_3\bar{A}_4A_5+\cdots$$

又因为各次射击是独立的，所以得

$$P\{\text{甲获胜}\}=P(A_1+\bar{A}_1\bar{A}_2A_3+\bar{A}_1\bar{A}_2\bar{A}_3\bar{A}_4A_5+\cdots)$$

$$=\alpha+(1-\alpha)(1-\beta)\alpha+(1-\alpha)^2(1-\beta)^2\alpha+\cdots$$

$$=\frac{\alpha}{1-(1-\alpha)(1-\beta)}$$

同理可得

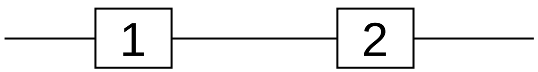
$$P\{乙获胜\} = P(\bar{A}_1 A_2 + \bar{A}_1 \bar{A}_2 \bar{A}_3 A_4 + \bar{A}_1 \bar{A}_2 \bar{A}_3 \bar{A}_4 \bar{A}_5 A_6 + \cdots)$$

$$= (1-\alpha)\beta + (1-\alpha)(1-\beta)(1-\alpha)\beta + \cdots$$

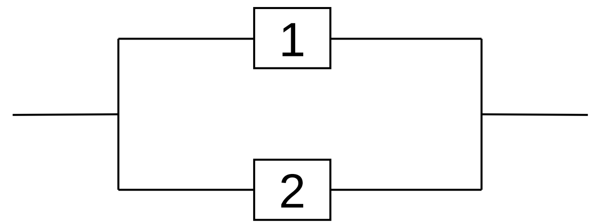
$$= \frac{\beta(1-\alpha)}{1-(1-\alpha)(1-\beta)}$$

可靠性理论是研究系统或组成系统的元件正常工作的概率问题的, 是应用概率论的一个重要分支. 对于一个元件, 它能正常工作的概率称为它的可靠性. 对于一个系统, 它能正常工作的概率称为该系统的可靠性, 并且在一般情况下, 通常假定组成系统的"各个元件能否正常工作"是相互独立的.

【例7】 系统由多个元件组成，且所有元件都独立地工作.设每个元件正常工作的均为$p=0.9$，试求以下系统正常工作的概率.



串联系统S_1



并联系统S_2

解： 设S_i表示“第i 个系统正常工作”，A_i表示“第i 个元件正常工作”.

（1）对串联系统而言，系统正常工作相当于所有元件正常工作，即$S_1 = A_1A_2$，所以

$$P(S_1) = P(A_1A_2) = P(A_1)P(A_2) = p^2 = 0.81.$$

（2）对并联系统而言, 系统正常工作相当于至少有一个元件正常工作，即$S_2 = A_1 \cup A_2$，所以

$$P(S_2) = P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1A_2)$$

$$= p + p - p^2 = 0.99$$

或者 $P(S_2) = P(A_1 \cup A_2) = 1 - P(\bar{A}_1\bar{A}_2)$

$$= 1 - (1 - p)(1 - p) = 0.99$$

试验的独立性

利用事件的独立性可以定义两个或多个试验的独立性.

定义6 设有两个试验E_1和E_2，假如试验的任一结果（事件）与试验的任一结果（事件）都相互独立，则称这两个试验相互独立.

例如：抛掷一枚硬币(试验E_1)与抛掷一颗骰子(试验E_2)是相互独立的试验.

类似地可以定义n个试验$E_1, E_2, \cdots, E_n$的相互独立性：如果试验E_1的任一结果，试验E_2的任一结果,……，试验E_n的任一结果都是相互独立的n个事件，则称试验$E_1, E_2, \cdots, E_n$相互独立.

例如:抛掷n枚硬币、检查n件产品、调查n个人是否观看过某电视节目等，都是n重伯努利试验.

【例8】一口袋中装有a个白球和b个黑球，采用有放回摸球，试证明第一次摸球(试验E_1)与第二次摸球(试验E_2)相互独立.

证明：第一次摸球可能结果有两个:

摸出白球(记为事件A)，摸出黑球(记为事件$\bar{A}$).

第二次摸球可能结果也有两个:

摸出白球(记为事件B), 摸出黑球(记为事件$\bar{B}$). 即

$$F_1=\left\{\Omega,\phi,A,\bar{A}\right\},\qquad F_2=\left\{\Omega,\phi,B,\bar{B}\right\}.$$

要证明F_1中任何一个事件与F_2中的任何一个事件相互独立，只需证明事件A与B相互独立即可.

$$P(A)=\frac{a}{a+b};\quad P(AB)=\frac{a^2}{(a+b)^2};\quad P(\bar{A}B)=\frac{ba}{(a+b)^2}.$$

所以

$$P(B)=P(AB)+P(\bar{A}B)=\frac{a^2}{(a+b)^2}+\frac{ba}{(a+b)^2}=\frac{a}{a+b}$$

因而 $$P(AB)=P(A)P(B)=\frac{a^2}{(a+b)^2}$$

事件A与事件B相互独立，证毕.