目　录

第一章　事件与概率

§1.5　条件概率

1. 条件概率的定义

一般讲, 条件概率就是在知道了一定的信息下所得到的随机事件的概率. 如两个工厂A和B生产同一品牌的电视机, 商场中该品牌有个统一的次品率, 比如0.5%, 如果你从某个途径知道该商场的这批电视机是A厂生产的, 则你买到的电视机的次品率不再是0.5%, 而应该比0.5%要小, 这个概率就是条件概率, 即你在知道了这批电视机是A厂生产的附加条件下的概率就是条件概率.

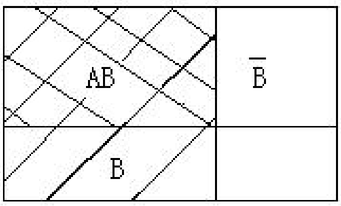
保险中应用的存活人数死亡率也是条件概率.

定义 1.5.1. 设事件A和B是随机试验Ω 中的两个事件, $P(B) > 0$, 称

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

为事件B发生条件下事件A发生的条件概率.

注 1.5.1. $P(A)$和$P(A|B)$ 是不同的两个概率. 如图, 设矩形A的面积为1, 则$P(A)$表示A的面积, 而$P(A|B)$表示在B中, A所占的比例, 即AB这块面积在B中所占的比例.



也可以从概率的统计定义, 即用频率来近似概率这一角度来理解条件概率. 设在n次独立试验中, 事件A发生了n_A次, 事件B发生了n_B次, 事件AB发生了n_{AB}次, 事件B发生下事件A发生的频率为

$$\frac{n_{AB}}{n_B} \approx \frac{P(AB)}{P(B)}$$

注 1.5.2. 事实上, 我们所考虑的概率都是在一定条件下计算的, 因为随机试验就是在一定的条件下进行的, 所以样本空间是相对而言的. 如果把在一定条件下的随机试验看成无条件的, 则在补充条件下进行的随机试验的结果一般而言相对于原有结果要少, 即样本空间改变了. 所以所得随机事件的概率一般是不相同的.

例 1.5.1. *有10个产品, 内有3个次品, 从中一个个地抽取(不放回) 检验, 问第一次取到次品后第二次再取到次品的概率.*

解: 样本空间Ω是从10个产品中有序取出2个产品的不同方法, 这是一个排列问题, 易知$\#\Omega = 10\times 9 = 90$, 记$A =$\{第一次取出的是次品\}, $B =$\{第二次取出的是次品\}, $\#(AB) = 6$, $\#A = 3$, 故

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{6/90}{3/10} = 2/9$$

注意, $P(B|A) = 2/9 \neq P(A) = 3/10$.

例 1.5.2. *有三张相同的卡片和一顶帽子, 第一张卡片两面都画有圈, 第二张卡片一面画圈, 一面画星, 第三张卡片两面都画星. 现在庄家把卡片放在帽中摇晃, 然后让你任取一张, 把它放在桌上, 设你看到卡片上面的图案为圈, 然后庄家与你打赌下面的图案与上面一样时算庄家赢, 不一样是为你赢. 请问这样的赌博是否是公平的?*

这是著名数学家, 信息论的创建者之一A. Weaver 设计的, 他曾在50年的《科学美国人》上介绍过这个例子. 请大家想一想, 很有意思.

例 1.5.3. *掷两个骰子, 观测出现的点数, 分别以x和y表示第一和第二颗骰子掷出的点数, 记$A = \{(x, y) : x + y \geq 9\}$, $B = \{(x, y) : x > y\}$, 求$P(A|B)$ 和$P(B|A)$.*

容易算出$P(A|B) = 2/15$, $P(B|A) = 1/3$, 这说明这两个条件概率不是一回事.

2. 乘法定理

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

由归纳法容易推广为n个事件同时发生的概率有如下公式:

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

上面公式的右边看似麻烦, 其实在实际中很容易算出. 在没有给出n个事件之间相互关系时, 这是计算n个事件同时发生的一个重要公式.

例 1.5.4. *某人忘了某饭店电话号码的最后一个数字, 因而随意拨号, 问他三次之内拨通电话的概率.*

解: 令A_i=\{第i次打通电话\}, $i = 1, 2, 3$, 则

$$P(3\text{次内拨通电话}) = P(A_1 \cup A_2 \cup A_3)$$

$$
\begin{aligned}
&= 1 - P(\bar{A}_1\bar{A}_2\bar{A}_3) \\
&= 1 - \frac{9}{10}\frac{8}{9}\frac{7}{8} = 0.3
\end{aligned}
$$

例 1.5.5. 将n根短绳的$2n$个端头任意两两连接,试求恰好连成n个圈的概率.

解:以Ω表示所有不同连结结果的集合,设想把$2n$个端头排成一行,然后规定将第$2k-1$个端头与第$2k$个端头相连接,$k=1,2,\cdots,n$.于是每一种排法对应一种连结结果,从而$|\Omega|=(2n)!$.以A表示恰好连成n个圈的事件.设想已将n根短绳作了编号,以A_k表示第k号短绳被连成1个圈的事件,于是有$A=\bigcap\limits_{k=1}^{n}A_k$.

当A_1发生时,1号短绳被连成1个圈,这相当于有一个$k\in\{1,2,\cdots,n\}$, 使得在$2n$个端头的排列中,1号短绳的两个端头排在第$2k-1$和第$2k$个位置上,所以$|A_1|=2n(2n-2)!$.因此

$$P(A_1)=\frac{|A_1|}{|\Omega|}=\frac{1}{2n-1}\ .$$

我们来求$P(A_2|A_1)$,即要在已知1号短绳被连成1个圈的情况下,求2号短绳也被连成1个圈的概率.既然1号短绳已经自成1个圈,我们就可以不考虑它,只要对剩下的$n-1$根短绳讨论其中的头一号短绳被连成1个圈的问题就行了.就是说,我们只要在变化了的概率空间上按计算无条件概率的公式来计算条件概率$P(A_2|A_1)$就行了.由于现在的情况与原来的情况完全类似,只不过总的绳数变为$n-1$根,故通过类比,即知

$$P(A_2|A_1)=\frac{1}{2(n-1)-1}=\frac{1}{2n-3}\ .$$

同理可得

$$P(A_k|A_1A_2\cdots A_{k-1})=\frac{1}{2[n-(k-1)]-1}=\frac{1}{2n-2k+1},\quad k=3,4,\cdots,n.$$

于是由概率乘法定理中的(2.3.6)式得到

$$P(A)=P(\bigcap_{k=1}^{n}A_k)=\prod_{k=1}^{n}\frac{1}{2n-2k+1}=\frac{1}{(2n-1)!!}\ .$$

在这个解法中,充分体现了利用变化了的概率空间计算条件概率的好处.

§1.5.1 全概率公式和Bayes公式

1. 全概率公式

定义 1.5.2. *设$B_1, B_2, \cdots B_n$ 是样本空间Ω 中的两两不相容的一组事件, 即$B_iB_j = \phi$, $i \neq j$, 且满足$\bigcup_{i=1}^n B_i = \Omega$, 则称$B_1, B_2, \cdots B_n$ 是样本空间Ω 的一个分割(又称为完备事件群,英文为partition).*

全概率公式:设$\{B_1, B_2, \cdots B_n\}$是样本空间Ω的一个分割, 且$P(B_i) > 0 (i = 1, \cdots, n)$, A为Ω中的一个事件, 则

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

目的: 有时不容易直接计算事件A的概率, 但是在每个B_i上A的条件概率容易求出.

注意: 应用中最重要的是验证$\{B_1, B_2, \cdots B_n\}$构成样本空间的一个分割.

例 1.5.6. *设某厂产品的一个零部件是由三家上游厂商供货的. 已知有一半是B_1 厂提供的, B_2厂商和B_3 分别提供25% . 已知厂商B_1 和B_2 的次品率都是2%, B_3 的次品率为4%, 从该厂产品中任取一个产品, 问该产品的这个零部件是次品的概率.*

解: 记$A =$\{取出的产品为次品\}, $B_i =$\{取到的产品是B_i 厂生产的\}, $i = 1, 2, 3$, 易见B_1, B_2, B_3 构成样本空间的一个分割, 且$P(B_1) = 0.5$, $P(B_2) = P(B_3) = 0.25$, $P(A|B_1) = P(A|B_2) = 0.02$, $P(A|B_3) = 0.04$, 由全概率公式马上得到

$$P(A) = 0.02 \times 0.5 + 0.02 \times 0.25 + 0.04 \times 0.25 = 0.025$$

例 1.5.7. *将n根短绳的$2n$个端头任意两两连接,求恰好连成n个圈的概率.*

解:现在再来利用全概率公式给出一个解答. 以A_n表示n根短绳恰好连成n个圈的事件,记$p_n = P(A_n)$.再以B表示第1根短绳连成1个圈的事件,用B和B^c作为对Ω的一个分划.于是由全概率公式得

$$p_n = P(A_n) = P(B)P(A_n|B) + P(B^c)P(A_n|B^c).$$

在前面例子中已经求得$P(B) = \frac{1}{2n-1}$;易见$P(A_n|B^c) = 0$;而$P(A_n|B)$ 则是在已知第1根短绳连成1个圈的条件下,其余$n-1$根短绳连成$n-1$个圈的概率,此时第1根短绳已经与其余$n-1$根短绳无关,所以$P(A_n|B) = P(A_{n-1}) = p_{n-1}$,代入上式即可得到

$$p_n = P(A_n) = \frac{1}{2n-1} p_{n-1}, \quad n = 2, 3, \cdots.$$

反复利用该式,并注意$p_1 = 1$,即得

$$p_n = \frac{1}{(2n-1)!!}\ ,\quad n = 1, 2, \cdots\ .$$

例 1.5.8. (Polya 罐子模型)罐中放有a个白球和b个黑球,每次从罐中随机抽取一个球,并连同c个同色球一起放回罐中,如此反复进行.试证明:在第n次取球时取出白球的概率为$\frac{a}{a+b}$.

证: 以A_k表示在第k次取球时取出白球的事件,于是A_k^c就是在第k次取球时取出黑球的事件.我们来对n作归纳.显然有$P(A_1) = \frac{a}{a+b}$.假设$n = k-1,\ k \geq 2$时结论成立,要证$n = k$时结论也成立.我们以A_1和A_1^c作为对Ω的一个分划.注意此时可将$P(A_k|A_1)$看成是从原来放有$a+c$个白球和b 个黑球的罐中按规则取球,并且在第$k-1$次取球时取出白球的概率,因此由归纳假设知$P(A_k|A_1) = \frac{a+c}{a+b+c}$,同理亦有$P(A_k|A_1^c) = \frac{a}{a+b+c}$,于是由全概率公式得

$$\begin{aligned} P(A_k) &= P(A_1)P(A_k|A_1) + P(A_1^c)P(A_k|A_1^c) \\ &= \frac{a}{a+b}\ \frac{a+b}{a+b+c} + \frac{b}{a+b+c}\ \frac{a}{a+b+c} = \frac{a}{a+b}\ . \end{aligned}$$

因此结论对一切n成立.

在上面解答中的对Ω的分划的选取方式值得我们注意.这里易走的一条歧路是把A_{k-1}和A_{k-1}^c作为对Ω的分划.在这种选取之下,我们难以利用归纳假设算出条件概率$P(A_k|A_{k-1})$和$P(A_k|A_{k-1}^c)$.因为此时我们只知道罐中有$a+b+(k-1)c$个球,而难于知道其中的白球和黑球数目.相反地,在A_1和A_1^c发生的情况下,罐中的白球和黑球数目则十分清楚.这个事实再次表明正确选取分划方式的重要性.

例 1.5.9. 一罐内有a个黑球和b个白球, 从中任意取一球, 如果是白球则将它放回去, 如果是黑球, 则从另一罐内取一白球替换它放回去。在重复n次这样的做法后, 求第n+1次取出的是白球的概率。

解: 记A={第n次取出的是白球}, $p_n = P(A)$, B为所求事件。则

$$\begin{aligned} p_{n+1} &= P(B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A}) \\ &= p_n * p_n + (p_n + \frac{1}{a+b})(1 - p_n) \end{aligned}$$

$$= \left(1-\frac{1}{a+b}\right)p_n+\frac{1}{a+b}$$

结合初值$p_1=\frac{a}{a+b}$，得到

$$p_{n+1}=1-\left(1-\frac{1}{a+b}\right)^n\frac{b}{a+b}.$$

2. Bayes公式

设$\{B_1,B_2,\cdots B_n\}$是样本空间的一个分割, A为Ω中的一个事件, $P(B_i)>0$, $i=1,2,\cdots,n$, $P(A)>0$, 则

$$P(B_i|A)=\frac{P(A|B_i)P(B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)}$$

什么情况下用Bayes公式? 由公式知, 分母就是事件A的概率, 而分子和等式左边的条件概率中的条件正好反过来. 所以我们知道在因果关系互换时必须用Bayes公式.

例 1.5.10. *一种诊断某癌症的试剂, 经临床试验有如下记录: 有癌症病人阳性的概率为95%, 无癌症病人阴性的概率为95%. 现用这种试剂在某社区进行癌症普查, 设该社区癌症发病率为0.5%, 问某人反应为阳性时该人患癌症的概率.*

解: 设$A=$\{反应为阳性\}, $C=$\{被诊断者患癌症\}, 由题意,

$$P(A|C)=0.95, P(\bar{A}|\bar{C})=0.95, P(C)=0.005,$$

现在要算的是$P(C|A)$. 这是典型的因果关系互换, 只能用Bayes公式.

$$\begin{aligned}P(C|A) &= \frac{P(A|C)P(C)}{P(A|C)P(C)+P(A|\bar{C})P(\bar{C})}\\ &= \frac{0.95\times0.005}{0.95\times0.005+0.05\times0.995}\\ &= 0.087=8.7\%\end{aligned}$$

这说明用该试剂进行普查, 准确性只有8.7%. 计算表明, 如果两次反应为阳性时患癌症的概率达到了64%.

§1.5.2 事件的独立性

为了计算两个事件同时发生的概率, 可以运用乘法定理, $P(AB)=P(A|B)P(B)$. 什么情况下$P(AB)=P(A)P(B)$? 即AB同时发生的概率等于两个事件单独发生概率的乘积? 为此我们有如下的定义:

定义 1.5.3. 设A, B是随机试验中的两个事件, 若满足$P(AB) = P(A)P(B)$, 则称事件A和B相互独立.

关于独立的概念, 应该是从实际出发, 如果能够判断事件B的发生与否对事件A的发生与否不产生影响, 则事件A, B即为独立. 如把一个硬币掷两次, 观测正反面出现的情况, $A =$\{第一次出现正面\}, $B =$\{第二次出现正面\}, $AB =$\{两次都出现正面\}, 样本空间Ω有4个基本事件, $\#(AB) = 1, \#(A) = 2, \#(B) = 2$, 故

$$P(AB) = 1/4,\ P(A)P(B) = 1/2 \cdot 1/2 = 1/4$$

即事件A, B相互独立. 事实上, 我们容易判断第一次是否出现正面与第二次是否出现正面没有任何影响, 即独立的. 设$\tilde{A}$ 表示事件A发生和不发生之一, $\tilde{B}$ 表示事件B发生和不发生之一. 由独立性的定义可以推知$P(\tilde{A}\tilde{B}) = P(\tilde{A})P(\tilde{B})$, (这儿一共4个等式) . 独立性的定义可以推广到n个事件.

定义 1.5.4. 设$A_1, A_2, \cdots A_n$是随机试验中的n个事件, 以$\tilde{A}_i$ 表示A_i或$\bar{A}_i$之一. 若满足

$$P(\tilde{A}_1\tilde{A}_2 \cdots \tilde{A}_n) = P(\tilde{A}_1)P(\tilde{A}_2) \cdots P(\tilde{A}_n),$$

则称事件列$A_1, A_2, \cdots A_n$ 相互独立.

(上面有2^n个等式)

注意: 上面等式等价于对$A_1, A_2, \cdots A_n$中的任意k个事件$A_{i_1}, A_{i_2}, \cdots, A_{i_k}$, $k = 2, \cdots n$, 有

$$P(A_{i_1}A_{i_2} \cdots A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k})$$

若$A_1, \cdots, A_n$中任意两个事件相互独立,则称为**两两独立**.

注意: 独立和不相容是不同的两个概念.

例 1.5.11. (两两独立而不相互独立的反例) 有四个同样的小球, 分别在其上写上"*1*","*2*","*3*"和"*1,2,3*"。引进三个事件: A_i=\{随机取一球, 球上有数字i\}, $i = 1, 2, 3$. 试讨论事件A_1, A_2, A_3是否相互独立.

解:易知$P(A_1) = P(A_2) = P(A_3) = \frac{1}{2}$,$P(A_1A_2) = P(A_2A_3) = P(A_3A_1) = \frac{1}{4}$,但是却有$P(A_1A_2A_3) = \frac{1}{4} \neq P(A_1)P(A_2)P(A_3)$, 所以事件$A_1, A_2, A_3$两两独立,但不相互独立.这个例子说明**两两独立不一定独立**。

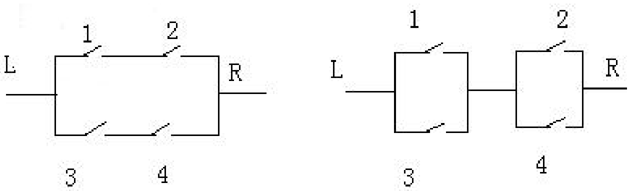
例 1.5.12. A, B, C 三人独立地破译密码, 每人能破译密码的概率分别为$1/3, 1/4, 1/5$. 问密码能被破译的概率有多大?

解: 设D={密码被破译}, A, B和C 分别表示A, B和C三人能破译密码这三个事件, 由独立性,

$$\begin{aligned} P(D) &= P(A \cup B \cup C) = 1 - P(\bar{A}\bar{B}\bar{C}) \\ &= 1 - P(\bar{A})P(\bar{B})P(\bar{C}) = 1 - \frac{2}{3}\frac{3}{4}\frac{4}{5} = 0.6 \end{aligned}$$

例 1.5.13. 在元件可靠性研究中, 我们考虑如下两种电路:

其中1-4表示4个继电器, 它们是否开通是相互独立的, 设继电器导通的概率为$p, (0 < p < 1)$, 求两种电路从L到R为通路的概率.



解: 左图为串联后并联, 右边为并联后串联, 记A_i ={第i个继电器导通}, 则左图LR为通路的表达为$A_1A_2 \cup A_3A_4$,右图LR为通路的表达为$(A_1 \cup A_3) \cap (A_2 \cup A_4)$, 由于$P(A_1A_2) = P(A_1)P(A_2) = p^2 = P(A_3A_4)$, 故

$$P(A_1A_2 \cup A_3A_4) = p^2 + p^2 - p^4 = p^2(2 - p^2)$$

同理,

$$P((A_1 \cup A_3) \cap (A_2 \cup A_4)) = (2p - p^2)^2 = p^2(2 - p)^2,$$

由于$2 - p^2 < (2 - p)^2$, 故并联后串联的电路比串联后并联的电路的可靠性高一点.

例 1.5.14. n个人独立向同一目标射击, 第i个人命中目标的概率为$p_i, i = 1, 2, \cdots, n$, 求至少有一人命中目标的概率.

解: 令A_i ={第i个人命中目标}, D ={至少有一人命中目标}, 则

$$
\begin{aligned}
D &= \bigcup_{i=1}^{n} A_i, \\
P(D) &= 1 - P(\bar{D}) = 1 - P(\bar{A}_1\bar{A}_2\cdots\bar{A}_n) \\
&= 1 - (1-p_1)(1-p_2)\cdots(1-p_n) \\
&\approx 1 - \exp\{-\Sigma p_i\}
\end{aligned}
$$

上面约等号在p_i较小时成立. 例如$p_i = 0.04$, $n = 100$ 时, $P(D) \approx 1 - \exp\{-4\} = 0.98168$