大数据的统计学基础

第3周

法律声明

【声明】本视频和幻灯片为炼数成金网络课程的教学资料，所有资料只能在课程内使用，不得在课程以外范围散播，违者将可能被追究法律和经济责任。

课程详情访问炼数成金培训网站

http://edu.dataguru.cn

关注炼数成金企业微信

■提供全面的数据价值资讯，涵盖商业智能与数据分析、大数据、企业信息化、数字化技术等，各种高性价比课程信息，赶紧掏出您的手机关注吧！

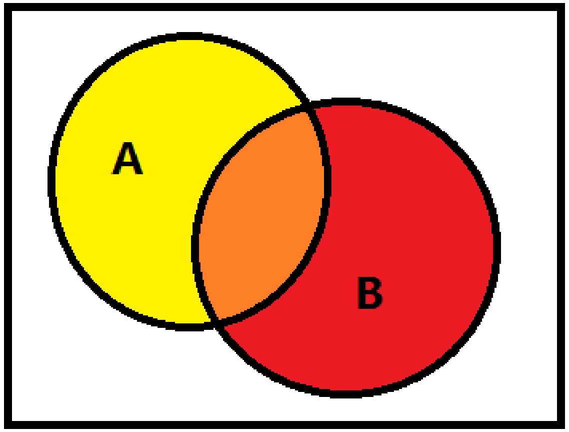
条件概率

- 回顾：赌大小的赌博游戏
- 通过上一次课的计算，我们知道开大或是开小的概率都是0.486111
- 问题：如果你知道三颗骰子中其中一颗的点数是3，那么你会选择押大还是押小呢？这时候开大或开小的概率分别是多少呢？

3	1	3	7	1	0
3	1	4	8	1	0
3	1	5	9	1	0
3	1	6	10	1	0
3	2	1	6	1	0
3	2	2	7	1	0
3	2	3	8	1	0
3	2	4	9	1	0
3	2	5	10	1	0
3	2	6	11	0	1
3	3	1	7	1	0
3	3	2	8	1	0
3	3	3			
3	3	4	10	1	0
3	3	5	11	0	1
3	3	6	12	0	1
3	4	1	8	1	0
3	4	2	9	1	0
3	4	3	10	1	0

条件概率的计算

- 已知某个事件A发生的条件下，另一个事件B发生的概率称为条件概率，记为P(B|A)
- 如何计算条件概率P(B|A)？？
- 甲乙两人各抛一颗骰子，点数大的赢。如果甲先抛骰子，得到点数4，那么乙获胜的概率是多少？
- 记A={甲得到点数为4}，B={乙获胜}
- P(A)=1/6；P(AB)=2/36=1/18；P(B|A)2/6=1/3
- 看一下P(B|A)与P(A)、P(B)的关系：P(B|A)=P(AB)/P(A)



条件概率——概率

- 条件概率说到底也是概率的一种，所以也符合概率定义的三个条件：
- 1. 非负性：P(B|A) ≥0;
- 2. 规范性：对于必然事件S，有P(S|A)=1
- 3. 可列可加性：对于两两互不相容的事件B_1， B_2， B_3……，即$B_i \cdot B_j = \varnothing$，i ≠j，i,j=1,2,……，有$P(B_1 \cup B_2 \cup \ldots\ldots|A)=P(B_1|A)+P(B_2|A)+\ldots\ldots$
- 对于概率的一些公式，条件概率也同样适用
- 如$P(A \cup B|C)=P(A|C)+P(B|C)-P(AB|C)$

例子

- 某公司年终决定举行抽奖活动，从全部员工中选取一名特等奖。公司人事架构如下：

部门	男	女	合计
行政部	10	10	20
销售部	20	10	30
技术部	10	4	14
客户部	20	16	36
合计	60	40	100

- （1）若被抽中的人是销售部的，问该员工是女性的概率？
- （2）若被抽中的人是女生的，问该员工是销售部的概率是？
- A={被抽中的是销售部的}，B={被抽中的是女生}
- (1)$P(B|A)=P(AB)/P(A)=(10/100)/(30/100)=1/3$
- (2)$P(A|B)=P(AB)/P(B)=(10/100)/(40/100)=1/4$

汽车与山羊

- 美国的一个电视游戏节目Let ‘s Make a Deal上有一个游戏，规则如下：参赛者会看见三扇关闭了的门，其中一扇的后面有一辆汽车，选中后面有车的那扇门就可以赢得该汽车，而另外两扇门后面则各藏有一只山羊。当参赛者选定了一扇门，但未去开启它的时候，节目主持人会开启剩下两扇门的其中一扇，露出其中一只山羊。主持人其后会问参赛者要不要换另一扇仍然关上的门。

汽车与山羊

- 面对这个问题，有两种观点：
- 1. 换与不换都一样，因为当一道藏有山羊的门被打开时，剩下的两道门中，汽车在任一道门的概率都是1/2，所以换与不换获得汽车的概率都一样
- 2. 换比不换好。有三种可能的情况，全部都有相等的可能性(1/3)：

  参赛者挑山羊一号，主持人挑山羊二号。转换将赢得汽车。

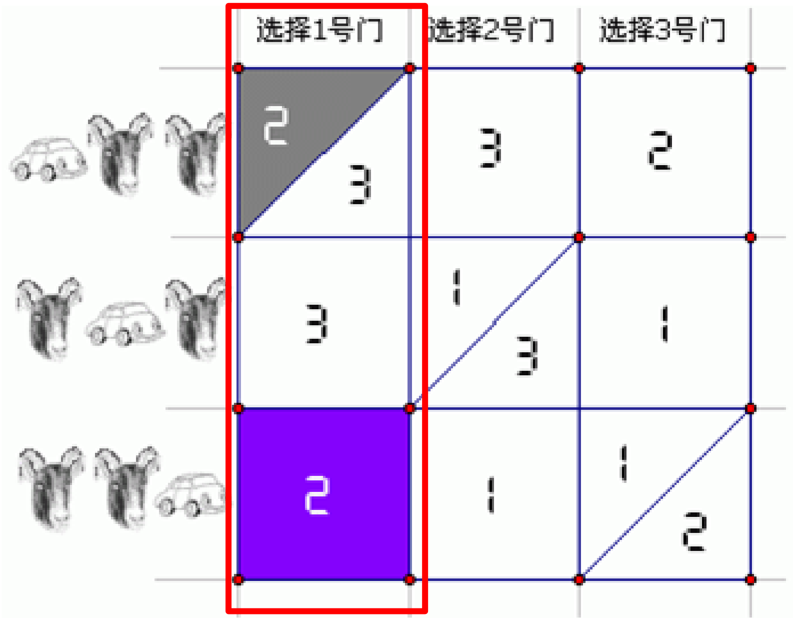
  参赛者挑山羊二号，主持人挑山羊一号。转换将赢得汽车。

  参赛者挑汽车，主持人挑两头山羊的任何一头。转换将失败。

  在头两种情况，参赛者可以通过转换选择而赢得汽车。第三种情况是唯一一种参赛者通过保持原来选择而赢的情况。因为三种情况中有两种是通过转换选择而赢的，所以通过转换选择而赢的概率是2/3。
- 你支持哪种看法？？

汽车与山羊——用数字说话

- 将3个门记为1,2,3号，假设参赛者先选择的是1号门。记A = { 1号门是汽车};B={2号门是汽车}；C={3号门是汽车}，则P(A)=P(B)=P(C)=1/3。原来的选择有1/3的机会获得汽车。
- 假设主持人开启了2号门，这个事件记为D。那么参赛者坚持选择或是改变选择而赢得汽车的概率又是多少？
- 从图中的第一列看出，当参赛者选择了1号门，2号门被打开的概率P(D)=1.5/3；汽车在1号门并且主持人打开了1号门的概率P(AD)=0.5/3



- 1. 坚持选择：P(A|D)=P(AD)/P(D)=1/3
- 2. 改变选择：P(CD)=1/3

  P(C|D)=P(CD)/P(D)=2/3
- 所以，改变选择将有更大的几率获得汽车。

汽车与山羊

- 历史上这个问题刚被提出的时候却引起了相当大的争议。这个问题源自美国电视娱乐节目Let’s Make a Deal，内容如前所述。作为吉尼斯世界纪录中智商最高的人，Savant在Parade Magazine对这一问题的解答是应该换，因为换了之后有2/3的概率赢得车，不换的话概率只有1/3。她的这一解答引来了大量读者信件，认为这个答案太荒唐了。因为直觉告诉人们：如果被打开的门后什么都没有，这个信息会改变剩余的两种选择的概率，哪一种都只能是1/2。持有这种观点的大约有十分之一是来自数学或科学研究机构，有的人甚至有博士学位。还有大批报纸专栏作家也加入了声讨Savant的行列。在这种情况下，Savant向全国的读者求救，有数万名学生进行了模拟试验。一个星期后，实验结果从全国各地飞来，是2/3和1/3。随后，MIT的数学家和阿拉莫斯国家实验室的程序员都宣布，他们用计算机进行模拟实验的结果，支持了Savant的答案。

乘法定理

- 由条件概率的定义，很容易得到P(AB)=P(B|A)P(A)，其中P(A)>0
- 这条公式很容易推广到P(ABC)=P(C|AB)P(B|A)P(A)=P(A|BC)P(B|C)P(C)

例 4 设某光学仪器厂制造的透镜，第一次落下时打破的概率为 1/2，若第一次落下未打破，第二次落下打破的概率为 7/10，若前两次落下未打破，第三次落下打破的概率为 9/10. 试求透镜落下三次而未打破的概率.

解 以 $A_i(i=1,2,3)$ 表示事件"透镜第 i 次落下打破"，以 B 表示事件"透镜落下三次而未打破". 因为 $B=\overline{A}_1\overline{A}_2\overline{A}_3$，故有

$$P(B)=P(\overline{A}_1\overline{A}_2\overline{A}_3)=P(\overline{A}_3\mid\overline{A}_1\overline{A}_2)P(\overline{A}_2\mid\overline{A}_1)P(\overline{A}_1)$$

$$=\left(1-\frac{9}{10}\right)\left(1-\frac{7}{10}\right)\left(1-\frac{1}{2}\right)=\frac{3}{200}.$$

乘法定理

另解，按题意

$$\overline{B}=A_1\cup\overline{A}_1A_2\cup\overline{A}_1\overline{A}_2A_3.$$

而 $A_1,\overline{A}_1A_2,\overline{A}_1\overline{A}_2A_3$ 是两两互不相容的事件，故有

$$P(\overline{B})=P(A_1)+P(\overline{A}_1A_2)+P(\overline{A}_1\overline{A}_2A_3).$$

已知 $P(A_1)=\frac{1}{2},P(A_2|\overline{A}_1)=\frac{7}{10},P(A_3|\overline{A}_1\overline{A}_2)=\frac{9}{10}$，即有

$$P(\overline{A}_1A_2)=P(A_2|\overline{A}_1)P(\overline{A}_1)=\frac{7}{10}\left(1-\frac{1}{2}\right)=\frac{7}{20},$$

$$P(\overline{A}_1\overline{A}_2A_3)=P(A_3|\overline{A}_1\overline{A}_2)P(\overline{A}_2|\overline{A}_1)P(\overline{A}_1)$$

$$=\frac{9}{10}\left(1-\frac{7}{10}\right)\left(1-\frac{1}{2}\right)=\frac{27}{200}.$$

故得

$$P(\overline{B})=\frac{1}{2}+\frac{7}{20}+\frac{27}{200}=\frac{197}{200}.$$

$$P(B)=1-\frac{197}{200}=\frac{3}{200}.$$

例子

- 某行业进行专业劳动技能考核，一个月安排一次，每人最多参加3次；某人第一次参加能通过的概率为60%；如果第一次未通过就去参加第二次，这时能通过的概率为80%；如果第二次再未通过，则去参加第三次，此时能通过的概率为90%。求这人能通过考核的概率。
- 解：
- 设 Ai={ 这人第i次通过考核 }，i=1,2,3。 A={ 这人通过考核 }， $A = A_1 \cup \bar{A}_1 A_2 \cup \bar{A}_1 \bar{A}_2 A_3$

$$P(A) = P(A_1) + P(\bar{A}_1 A_2) + P(\bar{A}_1 \bar{A}_2 A_3)$$

$$= P(A_1) + P(\bar{A}_1) \cdot P(A_2 \mid \bar{A}_1) + P(\bar{A}_1) \cdot P(\bar{A}_2 \mid \bar{A}_1) P(A_3 \mid \bar{A}_1 \bar{A}_2)$$

$$= 0.60 + 0.4 \times 0.8 + 0.4 \times 0.2 \times 0.9 = 0.992$$

亦可：

$$P(A) = 1 - P(\bar{A}) = 1 - P(\bar{A}_1 \bar{A}_2 \bar{A}_3) = 1 - P(\bar{A}_1) P(\bar{A}_2 \mid \bar{A}_1) P(\bar{A}_3 \mid \bar{A}_1 \bar{A}_2)$$

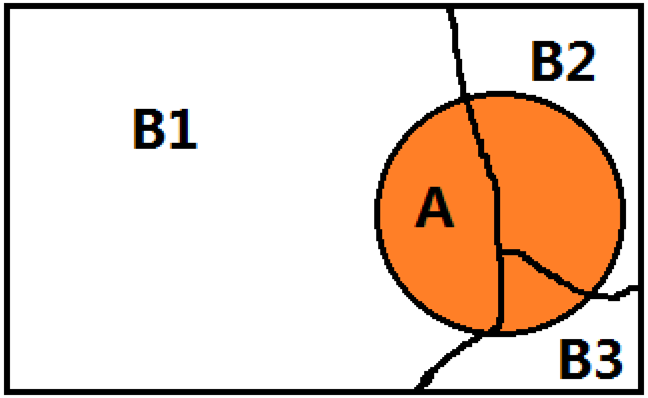
$$= 1 - 0.4 \times 0.2 \times 0.1 = 0.992$$

全概率公式

- 小明是今年的应届毕业生，他现受到了3家公司的面试通知，但不巧的是，面试时间基本一样，并且不能更改面试时间。小明只能也必须选择其中一家公司进行面试。如果小明有0.7的概率选择A公司，有0.5的概率面试成功；0.2的概率选择B公司，0.7的概率面试成功；0.1的概率选择C公司，0.3的概率面试成功。那么请算一算，小明面试成功的概率是多少？
- A=｛面试成功｝
- B1={到A公司面试}
- B2={到B公司面试}
- B3={到C公司面试}
- 根据题意，$P(B1)=0.7$，$P(A|B1)=0.5$；
- $P(B2)=0.2$，$P(A|B2)=0.7$
- $P(B3)=0.1$，$P(A|B3)=0.3$

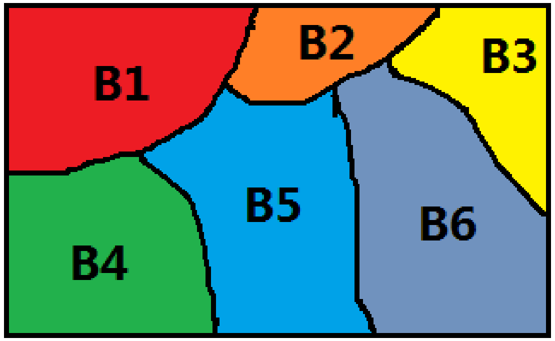
全概率公式

- {面试成功}={到A公司面试并成功}∪ {到B公司面试并成功} ∪ {到C公司面试并成功}
- 即A=AB1∪AB2 ∪ AB3，又B1，B2，B3为互斥事件，故
- P(A)=P(AB1)+P(AB2)+P(AB3)=P(A|B1)P(B1)+P(A|B2)P(B2)+P(A|B3)P(B3)=0.5*0.7 +0.7*0.2+0.3*0.1=0.52
- 上面其实使用了全概率公式



划分

- 设S为试验E的样本空间，B1，B2……Bn为E的一组事件。若
- （1）BiBj=∅，i≠j，i，j=1,2，……，n
- （2）B1 ∪B2 ∪…… ∪Bn=S
- 则称B1，B2……Bn是样本空间S的一个划分
- 上述例子中，B1，B2，B3就是样本空间的一个划分
- 例如，设试验为“抛两次硬币，观察正面朝上的次数”，那么样本空间就是S={0,1,2}，那么事件B1={正面朝上的次数不多于1}和B2={正面朝上的次数超过1}就是样本空间的一个划分

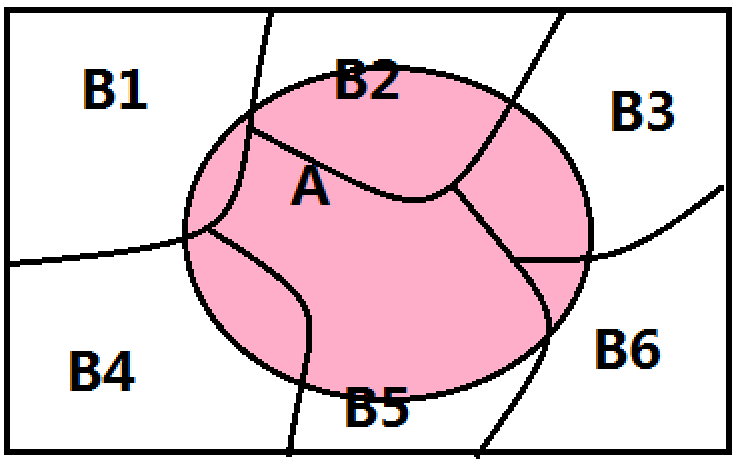


全概率公式

- 设试验E的样本空间为S，A为 E 的一个事件，B1,B2……Bn是S的一个划分，且P(Bi)>0(i=1,2……n)，则

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n)$$ ——全概率公式

在某些时候，事件A的概率不好求，但是通过全概率公式却可以很容易求得。



例子

- 例：假设在某时期内影响股票价格变化的因素只有银行存折利率的变化。经分析，该时期内利率下调的概率为60%，利率不变的概率为40% 。根据经验，在利率下调时某支股票上涨的概率为80%，在利率不变时，这支股票上涨的概率为40%。求这支股票上涨的概率。
- 设B1={利率下调}，B2={利率不变}，A={股票上涨}。依题意，$B1 \cup B2=S$，$B1B2=\varnothing$。
- 故根据全概率公式，有
- $P(A)=P(A|B1)P(B1)+P(A|B2)P(B2)=0.8*0.6+0.4*0.4=0.64$

例子

26. 病树的主人外出，委托邻居浇水，设已知如果不浇水，树死去的概率为 0.8. 若浇水则树死去的概率为 0.15. 有 0.9 的把握确定邻居会记得浇水.

（1）求主人回来树还活着的概率.

- 解：记A ={树活着}，B1={邻居记得浇水}，B2={邻居不记得浇水}
- B1∪B2=S ,B1B2=∅
- P(A)=P(A|B1)P(B1)+P(A|B2)P(B2)=0.85*0.9+0.2*0.1=0.785

（2）若主人回来树已死去，求邻居忘记浇水的概率.

- 求P(B2|A)，如何求？

乘法公式

- 根据条件概率的定义，有$\mathrm{P}(B_2|\mathrm{A}) = \frac{\mathrm{P}(AB_2)}{\mathrm{P(A)}} = \frac{P(A|B_2)P(B_2)}{P(A|B_1)P(B_1)+P(A|B_2)P(B_2)}$

全概率公式

贝叶斯公式

贝叶斯公式

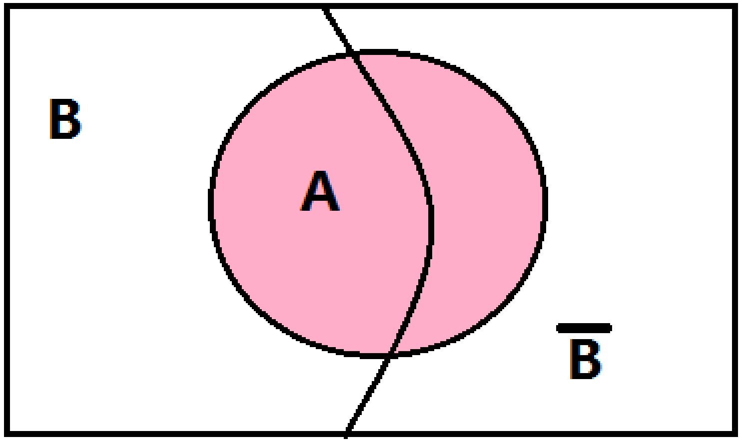
- 设试验E的样本空间为S。A为E的一个事件，B1，B2……Bn是S的一个划分，且P(A)>0，P(Bi)>0（i=1,2，……，n），则

- $P(B_i|A)=\dfrac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n}P(A|B_j)P(B_j)}$ ——————贝叶斯公式

- 当对样本空间的划分由一对对立事件B与$\bar{B}$组成时，全概率公式和贝叶斯公式可以简化为

$$P(A)=P(A\mid B)P(B)+P(A\mid \bar{B})P(\bar{B}),$$

$$P(B\mid A)=\frac{P(AB)}{P(A)}=\frac{P(A\mid B)P(B)}{P(A\mid B)P(B)+P(A\mid \bar{B})P(\bar{B})}.$$



例子

例 7 对以往数据分析结果表明，当机器调整得良好时，产品的合格率为98%，而当机器发生某种故障时，其合格率为55%。每天早上机器开动时，机器调整良好的概率为95%。试求已知某日早上第一件产品是合格品时，机器调整良好的概率是多少？

解 设 A 为事件"产品合格"，B 为事件"机器调整良好"。已知 $P(A|B)=0.98$，$P(A|\bar{B})=0.55$，$P(B)=0.95$，$P(\bar{B})=0.05$，所需求的概率为 $P(B|A)$。由贝叶斯公式

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

$$=\frac{0.98\times0.95}{0.98\times0.95+0.55\times0.05}=0.97.$$

这就是说，当生产出第一件产品是合格品时，此时机器调整良好的概率为0.97。这里，概率0.95是由以往的数据分析得到的，叫做**先验概率**。而在得到信息（即生产出的第一件产品是合格品）之后再重新加以修正的概率（即0.97）叫做**后验概率**。有了后验概率我们就能对机器的情况有进一步的了解。 □

贝叶斯公式的应用——诉讼

- 1981 年3 月30 日, 一个大学退学学生欣克利(John Hinckley Jr.) 企图对里根总统行刺. 他打伤了里根、里根的新闻秘书以及两个保安. 在1982 年宣判他时, 欣克利的辩护律师以精神病为理由作为其无罪的辩护。 作证的医师告诉法院当给被诊断为精神分裂症的人以CAT 扫描时, 扫描显示30% 的案例为脑萎缩, 而给正常人以CAT 扫描时, 只有2%的扫描显示脑萎缩. 欣克利的辩护律师试图拿欣克利的CA T 扫描结果为证据, 争辩说因为欣克利的扫描显示了脑萎缩, 他极有可能患有精神病, 从而应免受到法院的起诉。

贝叶斯公式的应用——诉讼

- 我们尝试用贝叶斯方法对欣克利是否患有精神病做出判断. 一般地, 在美国精神分裂症的发病率大约为1. 5% ：设A = {CAT 扫描显示脑萎缩}；B = { 做扫描的人患有精神病} . 根据上文的叙述可知,
- P(B)=0.015，P(A|B)=0.3，P($\bar{B}$)=1-0.015=0.985, P(A|$\bar{B}$)=0.02
- 由贝叶斯公式得：P(B|A)=$\frac{P(A|B)P(B)}{P(A|B)P(B)+P(A|\bar{B})P(\bar{B})} = 0.3*0.015/(0.3*0.015+0.02*0.985 = 0.18595041 \approx 18.6\%$
- 这意味着即使欣克利的扫描显示了脑萎缩, 他也只有18. 6%的可能患有精神病, 因此CAT 扫描无法作为其无罪的证据.

贝叶斯公式的应用——疾病诊断

例 8 根据以往的临床记录，某种诊断癌症的试验具有如下的效果：若以 A 表示事件"试验反应为阳性"，以 C 表示事件"被诊断者患有癌症"，则有 $P(A|C)=0.95, P(\overline{A}|\overline{C})=0.95$. 现在对自然人群进行普查，设被试验的人患有癌症的概率为 0.005，即 $P(C)=0.005$，试求 $P(C|A)$.

解 已知 $P(A|C)=0.95, P(A|\overline{C})=1-P(\overline{A}|\overline{C})=0.05, P(C)=0.005, P(\overline{C})=0.995$，由贝叶斯公式

$$P(C|A)=\frac{P(A|C)P(C)}{P(A|C)P(C)+P(A|\overline{C})P(\overline{C})}=0.087.$$

本题的结果表明，虽然 $P(A|C)=0.95, P(\overline{A}|\overline{C})=0.95$，这两个概率都比较高. 但若将此试验用于普查，则有 $P(C|A)=0.087$，亦即其正确性只有 8.7%（平均 1000 个具有阳性反应的人中大约只有 87 人确患有癌症）. 如果不注意到这一点，将会得出错误的诊断，这也说明，若将 $P(A|C)$ 和 $P(C|A)$ 混淆了会造成不良的后果. □

贝叶斯公式的应用——垃圾邮件判别

- 原理：若已知某些字词经常出现在垃圾邮件中，却很少出现在合法邮件中，当一封邮件含有这些字词时，那么他是垃圾邮件的可能性就很大。
- (1)创建基于字词符号的贝叶斯数据库——垃圾邮件与非垃圾邮件
- (2)创建贝叶斯概率库——垃圾概率
- (3)创建个性化的贝叶斯库——根据个人需求更改先验概率

公式比较

- 乘法公式、全概率公式与贝叶斯公式
- 1 乘法公式是求“几个事件同时发生”的概率；
- 2 全概率公式是求“最后结果”的概率；
- 3 贝叶斯公式是已知“最后结果”，求“某个事件”的概率.

- 先验概率与后验概率
- 1 P(Bj|A)是在事件A发生的条件下, 某个事件Bj发生的概率, 称为 “后验概率”；
- 2 Bayes公式又称为“后验概率公式”或“逆概公式”；
- 3 称P(Bj) 为“先验概率”.

独立性

◆ 抛一颗骰子两次，记A={第一次得到6点}，B={第二次得到6点}。（1）已知第一次得到6点，问第二次得到6点的概率。（2）已知第一次不是得到6点，问第二次得到6点的概率。（这里事件A的发生与否，对事件B发生的概率有影响吗？）

◆ 这个实验的样本空间S={(1,1),(1,2)……(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)}

共36种可能情况

◆ 根据前面所学知识，P(A)=1/6；P(B)=1/6；P(AB)=1/36

◆ $P(B|A)=P(AB)/P(A)=(1/36)/(1/6)=1/6=P(B)$

◆ $P(B|\bar{A})=P(\bar{A}B)/P(\bar{A})=(5/36)/(5/6)=1/6=P(B)$

$P(B|A)=P(B)$, $P(B|\bar{A})=P(B)$表示事件A的发生与否对事件 B 发生的概率都没有影响，这时我们可以说A、B相互独立

独立性

- 设A、B是两个事件，如果满足：P(AB)=P(A)P(B)，则称事件A、B相互独立。简称A、B独立。
- 由事件独立的定义可以推出：
- 1. A、B独立，且P(A)>0 $\Leftrightarrow$ P(B|A)=P(B)
- P(B|A)=P(AB)/P(A)=P(A)P(B)/P(A)=P(B)
- 2. 若A、B独立，则A与$\bar{B}$、$\bar{A}$与B、 $\bar{A}$与$\bar{B}$ 也相互独立

  $P(A)=P(A|B)P(B)+P(A|\bar{B})P(\bar{B})=P(A)P(B)+P(A\bar{B})$

  故$P(A\bar{B})=P(A)-P(A)P(B)=P(A)[1-P(B)]=P(A)P(\bar{B})$

多个事件相互独立

- 设A、B、C是三个事件，若满足

$$\left.\begin{aligned}&\boxed{\begin{aligned}&P(AB)=P(A)P(B),\\&P(BC)=P(B)P(C),\\&P(AC)=P(A)P(C),\end{aligned}}\\&P(ABC)=P(A)P(B)P(C),\end{aligned}\right\}$$

$\Rightarrow$ A、B、C两两独立

- 称A、B、C相互独立。

一般，设 $A_1, A_2, \cdots, A_n$ 是 n $(n\geq 2)$ 个事件，如果对于其中任意 2 个，任意 3 个，…，任意 n 个事件的积事件的概率，都等于各事件概率之积，则称事件 $A_1, A_2, \cdots, A_n$ 相互独立.

例子

- 多个事件相互独立≠多个事件两两独立
- 盒中编号为1,2,3,4的4只球，随机地从盒中抽取一只球，事件A为“取得的是1号球或2号球”，事件B为”取得的是1号球或3号球“，事件C为”取得1号球或4号球“。
- 则样本空间 S ＝｛ “取得 1 号球”，“取得 2 号球”，“取得 3 号球”，“取得 4 号球” ｝
- P(A)=2/4=1/2，P(B)=2/4=1/2，P(C)=2/4=1/2
- P(AB) =P(取得的是1号球)=1/4=P(A)P(B)
- P(AC)=P(取得的是1号球)=1/4=P(A)P(C)
- P(BC)=P(取得的是1号球)=1/4=P(B)P(C)
- 所以A、B、C两两独立。
- 但P(ABC)=P(取得的是1号球)=1/4≠P(A)P(B)P(C)=1/8，A、B、C没有相互独立

相互独立事件与互斥事件、对立事件

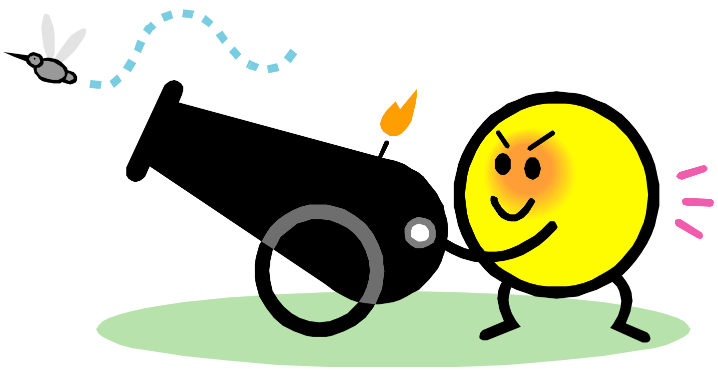
- 不要混淆相互独立事件、互斥事件、对立事件
- 相互独立事件：风马牛不相及。两个事件没有一点关系。例如，A、 B分别表示甲、乙两人患感冒，且甲乙两人的活动范围相距甚远，那么甲是否患感冒跟乙没什么关系，所以可以认为A、B独立。
- 互斥事件：要么只有其中一个事件发生，要么两个事件都不发生。在某次抽奖活动中，一等奖只有一个名额，A={甲中一等奖}，B={乙中一等奖}。那么A、B互为互斥事件，实际情况可能是甲中一等奖，可能是乙中一等奖，当然，更有可能甲乙都不中奖。
- 对立事件：两个只能活一个，不是你死就是我亡。跟互斥事件相比，对立事件必然会有一个事件发生。例如在上述的抽奖活动中，C={甲不中一等奖}，那么A与C是对立事件。
- 互斥事件与对立事件都不是相互独立事件！

相互独立事件与互斥事件、对立事件

- 设A和B为两事件，且P(A)=a,P(B)=b,问：
- (1) 当A和B独立时,$P(A\cup B)$为何值？
- (2) 当A和B互不相容时, $P(A\cup B)$为何值？
- (3) 当A和B互逆事件， $P(A\cup B)$为何值？
- （1） $P(A\cup B)=P(A)+P(B)-P(AB)=P(A)+P(B)-P(A)P(B)=a+b-ab$
- （2） $P(A\cup B)=P(A)+P(B)-P(AB)=a+b-0=a+b$
- （3） $P(A\cup B)=P(A)+P(B)-P(AB)=a+b-0=1$

例子

- 甲、乙两人同时向一目标射击（甲、乙同时射击，其结果互不影响），甲击中率为0.8，乙击中率为0.7，求目标被击中的概率。
- 设 A={甲击中}，B={乙击中}， C={目标被击中}。则
- $P(C) = P(A \cup B)=P(A)+P(B)-P(AB)=P(A)+P(B)-P(A)P(B)=0.8+0.7-0.8*0.7=0.94$



例子

例 2 一个元件(或系统)能正常工作的概率称为元件(或系统)的可靠性．如图 1－8,设有 4 个独立工作的元件 1,2,3,4 按先串联再并联的方式连接(称为串并联系统)．设第 i 个元件的可靠性为 $p_i(i=1,2,3,4)$,试求系统的可靠性．

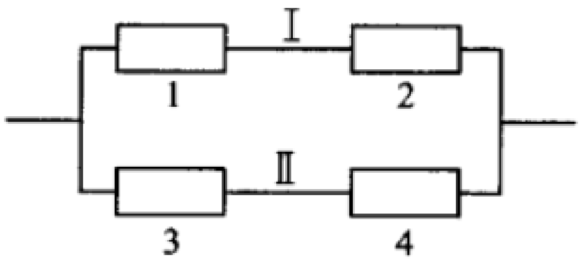


图 1－8

解 以 $A_i(i=1,2,3,4)$表示事件"第 i 个元件正常工作",以 A 表示事件"系统正常工作".

系统由两条线路 I 和 II 组成(如图 1－8)．当且仅当至少有一条线路中的两个元件均正常工作时这一系统正常工作,故有

$$A = A_1A_2 \cup A_3A_4.$$

由事件的独立性,得系统的可靠性

$$\begin{aligned} P(A) &= P(A_1A_2)+P(A_3A_4)-P(A_1A_2A_3A_4) \\ &= P(A_1)P(A_2)+P(A_3)P(A_4)-P(A_1)P(A_2)P(A_3)P(A_4) \\ &= p_1p_2+p_3p_4-p_1p_2p_3p_4. \end{aligned}$$

□

每人脑袋里有个贝叶斯

- 论人类思考时的贝叶斯过程：http://fmajor.lamost.org/blog/?p=1177

炼数成金逆向收费式网络课程

- Dataguru（炼数成金）是专业数据分析网站，提供教育，媒体，内容，社区，出版，数据分析业务等服务。我们的课程采用新兴的互联网教育形式，独创地发展了逆向收费式网络培训课程模式。既继承传统教育重学习氛围，重竞争压力的特点，同时又发挥互联网的威力打破时空限制，把天南地北志同道合的朋友组织在一起交流学习，使到原先孤立的学习个体组合成有组织的探索力量。并且把原先动辄成千上万的学习成本，直线下降至百元范围，造福大众。我们的目标是：低成本传播高价值知识，构架中国第一的网上知识流转阵地。
- 关于逆向收费式网络的详情，请看我们的培训网站 http://edu.dataguru.cn

Thanks

FAQ时间