Ch.1 Basic Concepts of Probability

第一章 概率论的基本概念

上节回顾

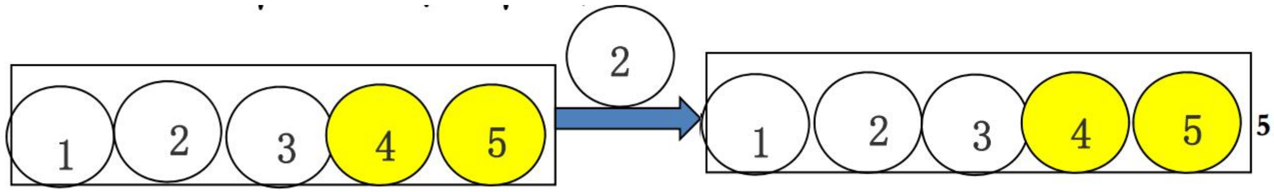
- 古典概型--超几何分布
- 几何概型--蒲丰投针
- 二项概型（贝努里概型）$b(k,n,p)$

填空题 2分

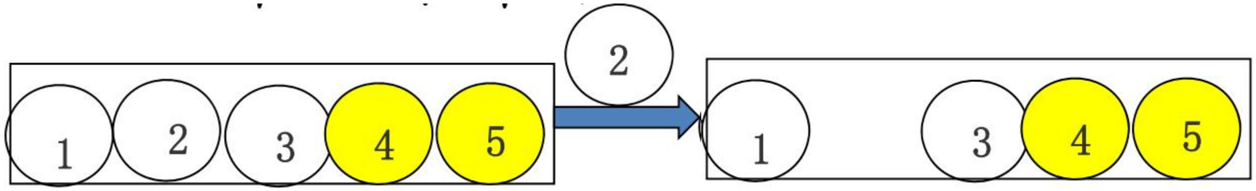
设置

袋中有a个白球，b个黄球，取k次球，

（1）有放回地取球，第i次取到白球的概率为 [填空1]



（2）无放回地取球，第i次取到白球的概率为 [填空2]



作答

思考题

- 如果把蒲丰投针中的针换成三角形，设三角形的三条边的边长分别为a,b,c，平行线的间隔是d，那么三角形与平行线相交的概率是多少？

（原始蒲丰投针相交的概率为$p = \frac{2l}{d\pi}$）

$$p = \frac{a+b+c}{d\pi}$$

Example 1.17 反欺诈（Anti-Fraud）

招聘专业品酒师，随机让他区分两种酒。每次给他一杯酒，让他品尝说出是哪一种。连续重复10次（每次后稍加休息、漱口）。如果10次中有8次正确，则聘; 否则不聘。

问： 这种做法合适否?

分析： 判断这种标准合适与否，也就是判断一下什么样的人被录用的可能性大。

如一个人水平高，区辨能力达到$p=90\%$，那么应该可以聘用；而如果一个人连蒙带唬，区辨能力为$p=50\%$，那就该拒绝。是否如此呢?

Example 1.17 反欺诈（Anti-Fraud）

解： 每次品酒要么正确，要么错误。

假设一个人每次正确判断可能为p，那么10次中有8次（包括以上）正确，其概率为：

$$b(8,10,p)+b(9,10,p)+b(10,10,p)$$

$$=p^8[45(1-p)^2+10p(1-p)+p^2]$$

如果$p=90\%$，则发生8次正确的概率为 $2.16p^8=0.929\ 809$

如果$p=80\%$，则发生8次正确的概率为 $4.04p^8=0.671\ 088$

如果$p=50\%$，则发生8次正确的概率为 $56.0p^8=0.054\ 684$

看来连蒙带唬的人8次正确的概率很小，只有约5.5%；而能力强的人（90%）8次正确的概率为约93%。

思考题

碰运气能否通过英语四级考试?

问题：早期CET-4 包括听力、语法结构、阅读理解、综合填空、写作等。除写作占15分外，其余85道题为单项选择题，每道题附有A、B、C、D四个选项。

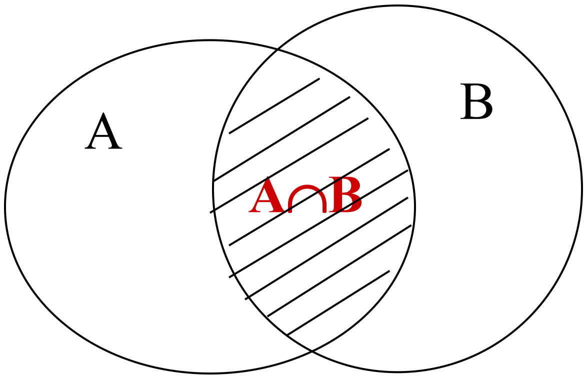
靠运气能通过CET-4考试吗?

不考虑写作所占的15分，正确率60%以上为及格，那85道选择题必须答对51题以上。

1.4 Principles for Probability

1.4.1 加法原则：

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



可以利用面积原理加以说明

Proof：

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

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

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

$$P\left(\bigcup_{i=1}^{n} A_k\right) = \sum P(A_i) - \sum P(A_iA_j) + \sum P(A_iA_jA_k) - \ldots + (-1)^{n-1}P(A_1A_2\ldots A_n)$$

填空题 1分

设置

在1-2000的整数中随机抽一个数，取到的整数既不能被6整除，也不能被8整除的概率是 [填空1]

作答

Example 1.4.1 四人打牌，以每人拿到红桃为例。

A＝东家拿到4张红桃

B＝西家拿到4张红桃

这两件事都是随机的，在抓牌前均是未知的。

我们考虑这样一件事：

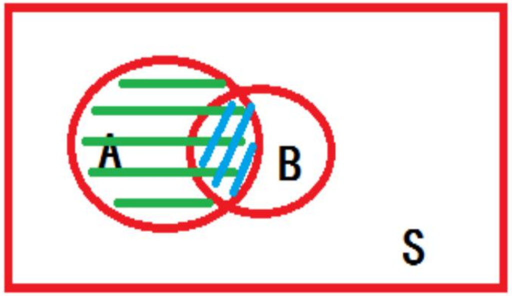
当东家真的抓到4张红桃后，他要判断一下西家拿到4张红桃的可能性是多大？

（他计算此种可能性的关心程度远超过在摸牌之前计算"西家拿到4张红桃"的可能性！）

注：卡内基梅隆大学和 Facebook 联合打造的史上最强德州扑克 AI "Pluribus" 在六人德州扑克这项复杂游戏中击败了顶级人类玩家，登上了Science 封面。

1.4.2条件概率与乘法定理

[定义1.4] 概率是对随机事件发生可能性大小的描述，在一事件A发生的前提下，另一事件B是否发生依然是随机的，其发生的可能性大小与A的发生有关系，记这样的概率为 P(B|A).



How to calculate P(B|A)

- 定义： 设**P(A)>0，**在A发生的前提下，看B发生的可能性

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

- 特例：古典概型，假设总样本数是n个，A有m个，AB有k个

$$P(B|A) = k/m$$

1.4.2条件概率与乘法定理

Example 1.4.1：家里有两个小孩，已知至少一个为女孩，

问：两个都是女孩的概率

解：A—至少一个为女孩；B—两个都是女孩。

S={(兄弟) (姐弟) (兄妹) (姐妹)}

A={(姐弟) (兄妹) (姐妹)}

B={(姐妹)}

$$P(B \mid A)=\frac{P(AB)}{P(A)}=\frac{1/4}{3/4}=\frac{1}{3}$$

$$P(B)=\frac{1}{4}$$

这里$P(B) \neq P(B \mid A)$.

其原因在于事件A的发生改变了样本空间，

使它由原来的S缩减为新的样本空间$S_A=A$.

1.4.2条件概率与乘法定理

Remark 1: 条件概率P(·| A)依然是概率，它满足概率的三个原则：

(1) 非负性： $P(B \mid A) \geq 0;$

(2) 规范性： $P(S \mid A) = 1;$

(3) 可列可加性： $B_1, B_2, \ldots, B_i B_j = \varnothing, i \neq j,$

则 $$P\left(\bigcup_{i=1}^{\infty} B_i \mid A\right) = \sum_{i=1}^{\infty} P(B_i \mid A).$$

Remark 2: P(AB）与 P（A|B）的区别

乘法定理

假设以下条件概率都存在：

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

$P(ABC)=P(C|AB)P(AB)=P(C|AB)P(B|A)P(A)$*

$P(A_1A_2\ldots A_n)=P(A_n|A_1A_2\ldots A_{n-1})P(A_1A_2\ldots A_{n-2}A_{n-1})$

$=P(A_n|A_1A_2\cdots A_{n-1})P(A_{n-1}|A_1A_2\cdots A_{n-2})\,P(A_1A_2\ldots A_{n-2})$

⋮

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

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

积的概率分解为多个条件概率的积，而这些条件概率相对是容易计算的

Example 1.4.2 袋中有r只红球，t只白球。每次自袋中任取一球，观察其颜色后放回，并再放入a只与所取球同色的球。

问：连续取球4次，第一、二次为红球且第三、四次为白球的概率

解：以$A_i(i=1,2,3,4)$表示“第 i 次取得红球”（从解题角度，难点在于此处的定义一个合理的事件）

则 $\overline{A}_3,\overline{A}_4$分别表示事件第三、四次取到白球。

所描述的事件为 $A_1A_2\overline{A}_3\overline{A}_4$

$$P(A_1A_2\overline{A}_3\overline{A}_4)=P(\overline{A}_4|A_1A_2\overline{A}_3)P(\overline{A}_3|A_1A_2)P(A_2|A_1)P(A_1)$$

$$=P(A_1)P(A_2|A_1)P(\overline{A}_3|A_1A_2)P(\overline{A}_4|A_1A_2\overline{A}_3)$$

$$=\frac{r}{r+t}\bullet\frac{r+a}{r+t+a}\bullet\frac{t}{r+t+a+a}\bullet\frac{t+a}{r+t+a+a+a}$$

Remark 1: 在处理条概的问题（和后续的贝叶斯问题）入门的难点在于，合理的设定被研究事件是什么！

1.4.3 独立事件

[定义1.2]: A、B为两事件，如果满足

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

\- 则称A、B相互独立。

[定理1.1] A、B为两事件，如果P(A)＞0，则

\- A、B相互独立 $\Leftrightarrow$ $P(B|A)=P(B)$。

填空题 2分

设置

一盒子装4只产品，3只合格1只次品。取两次球。在第一次取得是合格品的情况下，第二次依然取得合格品的概率

（1）无放回地取球 [填空1]

（2）有放回地取球 [填空2]

作答

一盒子装4只产品，3只合格1只次品。取两次。

问：在第一次取得是合格品的情况下，第二次依然取得合格品的概率

（1）无放回

A：第一次取得合格品；B：第二次取得合格品。

S—共有4×3种可能

A—共有3×3种可能， B—共有3×3种可能

AB—共有3×2种可能

$$P(B \mid A)=\frac{P(AB)}{P(A)}=\frac{6/12}{9/12}=\frac{2}{3}$$

A和B不独立

一盒子装4只产品，3只合格1只次品。取两次。

问：在第一次取得是合格品的情况下，第二次依然取得合格品的概率

（2）有放回

A：第一次取得合格品；B：第二次取得合格品。

S—共有4×4种可能

A—共有3×4种可能，B—共有4×3种可能

AB—共有3×3种可能

$$P(B \mid A) = \frac{P(AB)}{P(A)} = \frac{9/16}{12/16} = \frac{3}{4}$$

A和B独立

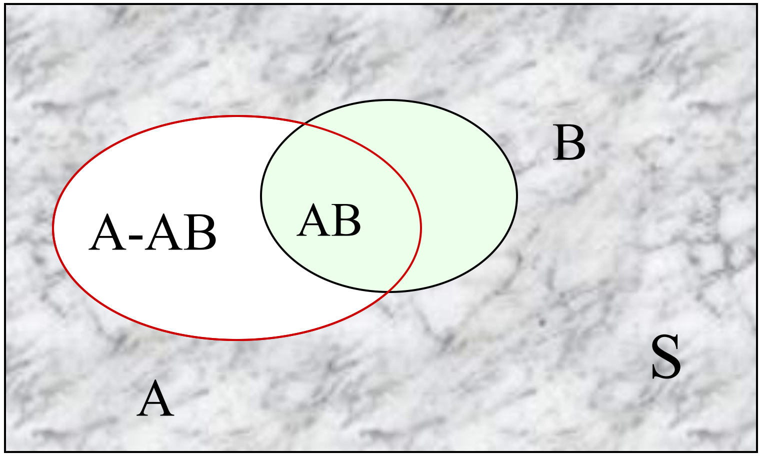
1.4.3 独立事件

➢ 事件A,B是独立的, 则A与$\overline{B}$ 独立

如果证明了此，我们就知道:

B与 $\overline{A}$ 独立

$\overline{A}$与 $\overline{B}$ 独立



Proof:

$$P(A\overline{B}) = P(A - AB)$$
$$= P(A) - P(AB)$$
$$= P(A) - P(A)P(B)$$
$$= P(A)\{1 - P(B)\}$$
$$= P(A)P(\overline{B})$$

1.4.3 独立事件

[定义1.3] 三个事件的相互独立，如果满足以下所有等式

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

$$P(BC)=P(B)P(C)$$

$$P(AC)=P(A)P(C)$$

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

一般，设A_1，A_2，……，A_n是n个事件，如果其中的任意多个事件的积事件的概率都等于各个事件概率的积，则称A_1，A_2，……，A_n相互独立。

Remark 3：两两独立未必三个独立

例 四张卡片分别标以数字1,2,3,4，今任取一张。

A：取得的是1或2，$P(A)=1/2$

B：取得的是1或3，$P(B)=1/2$

C：取得的是1或4，$P(C)=1/2$

$P(AB)=1/4 \quad P(BC)=1/4 \quad P(CA)=1/4$

$P(AB)=P(A)P(B),\ P(AC)=P(A)P(C),\ P(BC)=P(B)P(C)$

但是:

$P(ABC)=1/4, \quad P(A)P(B)P(C)=1/8$

引申1：此时$P(A|B)$，$P(AB|C)$是什么意思，又是多少?

引申2**：当5张数字时，事件不变，还两两独立吗?

Example 1.4.8 设某型号的计算机服务器可靠性为85%，现在采用冗余配置（并行工作），欲使得系统的可靠性达到99.99%，问至少应配多少台这样的服务器?

解： 假设共需要n台这样的服务器才能使得系统的可靠性达到99.99%。

A系统工作正常，A_i＝第i台服务器工作正常

$A = \cup A_i$

$P(A) = P(\cup A_i) \geq 99.99\%$

注意到A_i是可以同时发生，不具备互斥性，无法直接分解成概率的和。但是$\overline{A_i}$是独立的，也就是n台服务器是否down机是独立的。考虑如何把A_i和事件表达、分解为$\overline{A_i}$积事件。

Example 1.4.8

$$P(A) = 1 - P(\bigcap_{i=1}^{n} \overline{A}_i)$$

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

$$= 1 - (1 - 85\%)^n$$

$$= 1 - 0.15^n$$

$\geq 99.99\%$

加法原则：针对互斥的情况使用
乘法原则：针对独立的情况使用

$0.000\,1 \geq 0.15^{n}$

$n \geq 4.85$ 至少需要5台机器。

Example 1.4.9 若每个人的呼吸道中有感冒病毒的概率为0.002

求：在有1500人看电影的剧场中有感冒病毒的概率。

解：A_i－事件“第i个人带有感冒病毒”（$i=1,2,\ldots,1500$）

假定每个人是否带有感冒病毒是相互独立的，则所求概率为

$$P\left(\bigcup_{i=1}^{1500} A_i\right)=1-P\left(\overline{A_1}\,\overline{A_2}\cdots\overline{A_{1500}}\right)$$

$$=1-P\left(\overline{A_1}\right)P\left(\overline{A_2}\right)\cdots P\left(\overline{A_{1500}}\right)$$

$$=1-(1-0.002)^{1500}$$

$$=1-0.998^{1500}$$

$$\approx 0.95$$

可见：虽然每人带感冒病毒的可能性很小，但许多人聚集在一起时，空气中含有感冒病毒的概率可能会很大

- 这种现象称为**小概率事件的聚众效应**。
- 特殊时期，不聚众，戴口罩。

思考题

甲乙两人进行乒乓球比赛，每局甲胜的概率为p，$p \geq 1/2$。问：对甲而言，采用三局两胜制有利，还是采用五局三胜制有利。**假设各局的胜负相互独立**。

1.4.3 独立事件

Remark 1：区分互斥、对立、独立三个概念；

互斥：$A \cap B = \Phi$

对立：互斥$A \cap B = \Phi$，并$A \cup B = S$

独立：$P(AB) = P(A)P(B)$

互斥必定不独立

Remark 2: 独立、互斥往往是根据实际意义去判断