概率统计1——概率

基本概念

- **随机试验**:对一个不确定现象通过观察，测量，实验等手段获取其结果的过程
- **样本空间**:所有随机试验结果的总和，如掷色子的样本空间为$\{1,2,3,4,5,6\}$,记作Ω。**样本空间可以是无限集**
- **随机事件**: 样本空间的子集
- **样本点**:样本空间中的元素，也就是一次随机试验的单个结果，记作ω
- **互不相容事件**:事件A,B不同时发生，即$A\cap B=\phi$
- **对立事件**:A,B互不相容且$A\cup B=\Omega$,记作$A=\bar{B},B=\bar{A}$
- **完备事件组**:$A_1,A_2,\cdots,A_n$两两互不相容，且$\bigcup_{i=1}^{n}A_i=\Omega$

> Ω——必然事件——样本空间——全集
> ϕ——不可能事件——空集

概率的公理化描述

概率测度（Probability Measure）是一个从事件映射到实数域的实函数，其定义域为样本空间Ω，上域为实数域R，值域为$[0,1]$.任何一个概率测度都必须满足以下公理：

基本公理

- **非负性公理**:任何一个事件A发生的概率不可能为负，即$0\le P(A)$
- **规范性公理**:样本空间作为一个事件必然发生，即$P(\Omega)=1$
- **完全可加性公理**:若事件$A_1,A_2,A_3,\cdots,A_n$两两互不相容，则
  $P(A_1+A_2+A_3+\cdots+A_n)=P(A_1)+P(A_2)+P(A_3)+\cdots+P(A_n),n\in N^+$

基本公理的推论

1. $P(A)=1-P(\bar{A})$

> **证明**：因为事件 A 与 $\bar(A)$ 两两互不相容，即 $A\cap\bar{A}=\phi,A\cup\bar{A}=\Omega$,故$P(\Omega)=P(A+\bar{A})=P(A)+P(\bar{A})=1$,证毕

2. $P(A-B)=P(A)-P(AB)$(事件的减法性原理)

> **证明**：$A=(A-B)\cup AB$,又因为显然 $A-B$ 与AB 互不相容，则根据公理3，有$P(A)=P(A-B)+P(AB)$,证毕

3. $P(A+B)=P(A)+P(B)-P(AB)$(事件的加法性原理)

> **证明**：因为$A\cup B=A\cup(B-AB)$,且A与 $(B-AB)$ 互不相容，则显然根据完全可加性公理，有$P(A+B)=P(A)+P(B-AB)$.根据事件的减法性原理，有$P(B-AB)=P(B)-P(B\cap AB)=P(B)-P(AB)$.因此，有$P(A+B)=P(A)+P(B)-P(AB)$，证毕.

三个事件的加法公式：$P(A+B+C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)$

例题1: $P(A)=P(B)=P(C)=\frac{1}{4},P(AB)=0,P(AC)=P(BC)=\frac{1}{16}$,求：

1. ABC中至少发生一个
2. ABC都不发生

条件概率

有的时候，两个事件之间不是相互独立的，往往某个事件的发生与否会对另一个事件的发生可能性产生影响，这个时候，我们就需要使用

条件概率公式去描述这种现象了。

定义1: 设Ω为样本空间，A, B为两个事件且$P(B) > 0$,则称在B已经发生的条件下A发生的概率为A对 B的条件概率，记作$P(A|B)$

条件概率的本质是将样本空间从所有事件都发生的Ω缩小到原来仅仅B发生的事件的集合Ω_B中去，然后在这个新集合中计算A发生的概率。因此条件概率的计算公式就是：

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

例题2: 编号为$1-6$的球，随机取一个观察其编号，现设A_1表示取到1号，A_2取到2号，A_3取到大于4号的球；B表示取到偶数的球。求：

1. 在取到偶数的情况下取到2号球的概率
2. 在取到偶数的情况下A_3发生的概率

条件概率的性质：

- $P(A|B) \geq 0$
- $P(\Omega|B) = 1$
- 设$A_1, A_2, \cdots, A_n, n \in N^+$互不相容，则$P(\sum_{i=1}^{\infty} A_i|B) = \sum_{i=1}^{\infty} P(A_i|B)$

条件概率的乘法公式：

- $P(AB) = P(B)P(A|B) = P(A)P(B|A)$
- $P(ABC) = P(A)P(B|A)P(C|AB)$（超过三个可依次类推）

例题3: 如果天气多云时有雨的概率是0.3, 且天气多云的概率是0.2，那么天气多云且有雨的概率是?

课后作业1： 假设一共有100个产品，其中10个不合格，现在从中不放回抽取3个，问第一次第二次抽到次品并且第三次抽到的才是正品的概率有多大?

全概率公式

定义2： 设$A_1, A_2, A_3 ... A_n$是某个完备事件组,且$P(A_i) > 0, B$是某个同一样本空间中的其他事件，则有

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B|A_i)$$

全概率公式的意义在于，当需要求某个事件的发生概率时，我们可以去求该事件在各种不同情况下发生的概率，最终求其总和。

例题4: 假设现有10个零件，其中7个是合格品，3个是次品，现已经取走2个零件，则问当取第三个零件时，是正品的概率有多大?

例题5: 盒子中有3个红球和1个蓝球，无重复的抽取两个，则第二个球是红色的概率是多少?

课后作业2: 假设现在有10个产品，其中分别有$0,1,2$个次品的概率均为$\frac{1}{3}$，并且由于质检存在误差，将正品误检为次品的概率为0.02，将次品误检为正品的概率为0.05。现在随机从10个产品中抽取1个送检，则问：该产品检验通过的概率有多少?

贝叶斯公式

贝叶斯公式是一种已知结果反求起因的数学模型，也就是导致某个事件发生的起因有多种，现在已知该事件发生了，那么请问引发该事件发生的各个起因的概率分别是多少?

其实贝叶斯公式就是一种条件概率公式的变形，其具体定义如下所示：

定义3:设$A_1,A_2,A_3...A_n$某个完备事件组，且$P(A_i)>0,B$是同一样本空间中可由$A_1,A_2,A_3...A_n$引发的某个已知事件,则

$$P(A_i|B)=\frac{P(A_iB)}{P(B)}=\frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)}$$

例题6: 设某种疾病的发病概率为0.0004，其中由于医院的检测手段存在误差，其中真正的患者被诊断为有病的概率为0.99，健康的人被诊断为有病的概率为0.001。现某人经医院检查被确诊，则该人确实有病的概率有多大?

课后作业3: 设全国有0.4的人玩LOL,玩LOL的群体中,0.7是男性,不玩LOL的群体中,0.2是男性。已知Bob是男性，问：他是LOL玩家的概率是多大?

事件的独立性

事件的独立性的概念是指：已知一个事件发生不能为我们提供另一个事件是否发生的信息，或一个事件的发生不影响另一个事件的发生概率。即$P(A)=P(A|B)$.进一步地，由于$P(AB)=P(B)P(A|B)$,因此，当两个事件互相独立时，有

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

上式子便是两个事件独立的充分必要条件

独立事件的性质：

- ϕ和Ω与任意事件A都独立
- A与B独立，则A与$\bar{B}$, $\bar{A}$与B，$\bar{A}$与$\bar{B}$都互相独立

我们讨论事件多于两个时的独立性问题。我们首先给出事件**两两独立**(pairwise independent)和**互相独立**(mutually independent)的概念。

所谓两两独立，是指在同一样本空间Ω中有某事件集$A_1,A_2,\cdots,A_n$，从其中任意选取两个$A_i,A_j\quad i,j\in 1,2,3\cdots n$,都有$P(A_iA_j)=P(A_i)P(A_j)$成立。

所谓互相独立，是指在同一样本空间Ω中有某事件集$A_1,A_2,\cdots,A_n$，其任意的子集 $A_{i1},A_{i2},\cdots,A_{im}\quad i1,i2\cdots im\in 1,2,3\cdots n$满足$P(A_{i1}A_{i2}\cdots A_{im})=P(A_{i1})P(A_{i2})\cdots P(A_{im})$

很显然地，互相独立的定义已经包含了两两独立，若一个事件组是互相独立的，则其必然也是两两独立的。那么反过来的情况如何呢？下例将给出否定的答案：

> 抛掷两次骰子，令A表示第一次得到正面，B表示第二次得到正面，C表示总共得到一次正面。则易知

> $P(A)=P(B)=P(C)=\frac{1}{2}, P(AB)=P(A)P(B)=\frac{1}{4}$
> $P(AC)=P(A)P(C|A)=\frac{1}{4}=P(A)P(C), P(BC)=P(B)P(C|B)=\frac{1}{4}=P(B)P(C)$
> 所以显然此时ABC是两两独立的，但是显然$P(ABC)=0\neq P(A)P(B)P(C)=\frac{1}{8}$

究其原因，我们可以发现$P(ABC)=P(A)P(B|A)P(C|AB)$,当ABC 两两独立时有$P(B)=P(B|A)$,但是却无法保证$P(C)=P(C|AB)$。因此我们可以发现，互相独立的概率意义在于:**给出一个事件组并从该事件组中任选一个子集，则该子集中的事件同时发生时，事件组中剩余的任一事件对其保持独立**

例题7: 一个电路中，第一个元件与第二三个串联起来的元件形成一个并联电路，设A_i表示第i个元件保持正常工作，且$P(A_i)=p, i=1,2,3$。问该电路正常工作的概率是多少?

课后作业4: 设有一套密码，每个技术员能够将其破译的概率为0.6，并且每个技术员的破译工作都是相互独立的.则问：如果需要以0.99以上的可能性将其破译，那么至少需要几个技术员?