概率论与数理统计 第2次作业

Name: 宋昊原 　　 Student ID: 2022010755

October 15, 2023

1 三元容斥

$$
\begin{aligned}
A+B+C &= (A+B)^cC+(A+B)C^c+(A+B)C \\
&= A^cB^cC+(A^cB+AB^c+AB)C^c+(A^cB+AB^c+AB)C \\
&= A^cB^cC+A^cBC^c+AB^cC^c+ABC^c+A^cBC+AB^cC+ABC
\end{aligned}
$$

将上述七项分别记作F_1至F_7，则类似地有

$$A = F_3+F_4+F_6+F_7$$

$$
\begin{aligned}
B &= F_2+F_4+F_5+F_7 \\
A &= F_1+F_5+F_6+F_7 \\
AB &= F_4+F_7 \\
BC &= F_5+F_7 \\
AC &= F_6+F_7 \\
ABC &= F_7
\end{aligned}
$$

F_1至F_7两两互斥，故有

$$P(A+B+C)=\sum_{i=1}^{7}P(F_i)=P(A)+P(B)+P(C)-P(AB)-P(BC)-P(AC)+P(ABC)$$

2 条件概率

分别证明三条概率需要满足的三条原理.

2.1 $P(A|B)>=0$

$\forall A\in\mathcal{F}$，有$P(AB)\geq 0$，又由题设，$P(B)>0$，故有

$$P(A|B)=\frac{P(AB)}{P(B)}\geq 0$$

2.2 $P(\Omega|B)=1$

$$P(\Omega|B)=\frac{P(\Omega B)}{P(B)}=\frac{P(B)}{P(B)}=1$$

2.3 加法原理

$\forall A_i, A_j, A_iA_j = \emptyset$，有$(A_iB)(A_jB) = A_iA_jB = \emptyset B = \emptyset$，于是任何$A_iB, A_jB$互斥. 进而有

$$
\begin{aligned}
\sum_{i=1}^{\infty} P(A_i|B) &= \sum_{i=1}^{\infty} \frac{P(A_iB)}{P(B)} \\
&= \frac{\sum\limits_{i=1}^{\infty} P(A_iB)}{P(B)} \\
&= \frac{P(\sum\limits_{i=1}^{\infty} A_iB)}{P(B)} \\
&= \frac{P((\sum\limits_{i=1}^{\infty} A_i)B)}{P(B)} \\
&= P(\sum_{i=1}^{\infty} A_i|B)
\end{aligned}
$$

3 互斥事件和独立事件

3.1

不正确. 当$P(A) < 1$且$A = B$时，$P(A|B) = P(A|A) = 1 > P(A)$.

3.2

不正确. 当$B = \emptyset$时，显然A与B互斥，而$P(AB) = P(A)P(B) = 0$.

3.3

不正确. 当$A = \emptyset$且$P(B)P(C) \neq P(BC)$时，也有$P(ABC) = P(A)P(B)P(C) = 0$.

4 反直觉的独立性

掷两个骰子可能的点数和及对应的样本点数如下表:

点数和	2	3	4	5	6	7	8	9	10	11	12
样本点数	1	2	3	4	5	6	5	4	3	2	1

据表有

$$
\begin{aligned}
P(A_2) &= \frac{1+3+5+5+3+1}{36} = \frac{1}{2} \\
P(A_3) &= \frac{2+5+4+1}{36} = \frac{1}{3} \\
P(A_5) &= \frac{4+3}{36} = \frac{7}{36} \\
P(A_2A_3) = P(A_6) &= \frac{5+1}{36} = \frac{1}{6} \\
P(A_{10}) &= \frac{3}{36} = \frac{1}{12}
\end{aligned}
$$

故

$$P(A_2A_3) = P(A_2)P(A_3)$$

而

$$P(A_2A_5) \neq P(A_2)P(A_5)$$

于是A_2和A_3独立，A_2和A_5不独立.

5 独立和条件独立

5.1 独立并不意味着条件独立

掷一个骰子，设A_1表示点数为偶数，A_2表示点数为3的倍数，B表示点数大于等于4. 我们有：

$$P(A_1) = \frac{1}{2}$$
$$P(A_2) = \frac{1}{3}$$
$$P(A_1A_2) = \frac{1}{6}$$
$$P(A_1|B) = \frac{2}{3}$$
$$P(A_2|B) = \frac{1}{3}$$
$$P(A_1A_2|B) = \frac{1}{3}$$

可见A_1与A_2独立，但并不关于B条件独立.

5.2 条件独立并不意味着独立

设$0 < P(B) < 1$，则B与B总是不独立的，但B与B关于B独立.

6 小概率事件

设A_i表示第i次试验中事件A不发生. 则

$$P(A_i) = 1 - \epsilon$$

前n次试验中发生一次的概率为

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

故试验迟早发生一次的概率为

$$\lim_{n\to\infty} 1 - (1-\epsilon)^n = 1 - 0 = 1$$

7 红黑卡片

设A为向上的面为红色，B_1为选中全黑卡片，B_2为选中全红卡片，B_3为选中双色卡片.
则B_1，B_2，B_3构成对样本空间Ω的一个等概率划分.
我们有

$$P(A|B_1) = 0$$
$$P(A|B_2) = 1$$
$$P(A|B_3) = \frac{1}{2}$$

根据Bayes公式，我们有

$$\begin{aligned} P(B_3|A) &= \frac{P(A|B_3)P(B_3)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3)} \\ &= \frac{\frac{1}{2} \times \frac{1}{3}}{0 \times \frac{1}{3} + 1 \times \frac{1}{3} + \frac{1}{2} \times \frac{1}{3}} \\ &= \frac{1}{3} \end{aligned}$$

即此时另一面是黑色（即B_3发生）的概率是$\frac{1}{3}$

8 抓阄

两种方案都公平.

8.1 抓完同时打开的方案

这种方案下，每个人抓到“中”的概率都是先验概率$\frac{1}{n}$，公平.

8.2 边抓边打开的方案

这种方案下，每个人抓到“中”的概率会随着之前阄的打开变成后验概率，下面我们根据这个逻辑讨论每个人中签的全概率. 设第i个人中签为事件A_i.
首先，第一个人中签与先验后验无关，$P(A_1)=\frac{1}{n}$
第二个人中签的条件是第一个人没中签，而此时的条件概率为$\frac{1}{n-1}$，则有

$$P(A_2)=P(A_1^c)P(A_2|A_1^c)=\frac{8}{9}\times\frac{1}{8}=\frac{1}{9}$$

类似地

$$\begin{aligned}P(A_i)&=P(A_1^c)P(A_2^c|A_1^c)...P(A_{i-1}^c|A_1^c...A_{i-1}^c)P(A_i|A_1^cA_2^c...A_{i-1}^c)\\&=\frac{n-1}{n}\times\frac{n-2}{n-1}\times...\times\frac{n-i+1}{n-i+2}\times\frac{1}{n-i+1}\\&=\frac{1}{n}\end{aligned}$$

因此每个人中签的概率均为$\frac{1}{n}$，方案公平.

9 手术诊断

设事件A表示小明患病，B表示检验阳性，则先验概率$P(A)=0.6$，条件概率$P(B|A)=1$，$P(B|A^c)=0.3$. 根据Bayes公式，我们有

$$\begin{aligned}P(A|B)&=\frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|A^c)P(A^c)}\\&=\frac{1\times0.6}{1\times0.6+0.3\times0.4}\\&=\frac{5}{6}>0.8\end{aligned}$$

故医生应该建议手术.

10 不要赌博

本题我的做法参考了William Feller著《概率论及其应用》中对随机徘徊和破产问题的讨论，加以整理得到.

10.1 破产的概率

设此人在初始资金为r时会输光离场的概率为$q(r)$（为了严谨起见做此声明：这里的“会破产的概率”是m轮内会破产的概率对$m\to\infty$的极限，m轮内会破产的概率是单调不减的，故此极限存在）.
可以写出q_r满足的方程:

$$q_r=\begin{cases}1 & r=0\\ pq_{r+1}+(1-p)q_{r-1} & 0<r<n\\ 0 & r=n\end{cases}\tag{1}$$

考虑当$0<r<n$时的差分方程:

$$q_r=pq_{r+1}+(1-p)q_{r-1}\tag{2}$$

假如我们已经找到了方程(2)的两个线性无关的特解$q_{1,r}$和$q_{2,r}$，由于q_r可以由q_0和q_1确定，方程(2)的解空间维数为2，故$q_r = Aq_{1,r} + Bq_{2,r}$是全部可能的解，进而可以待定系数法根据边界条件求出对应的概率解.

考虑$p \neq \frac{1}{2}$的情形，我们可以找到(2)的两个线性无关的特解$q_r = 1$和$q_r = (\frac{1-p}{p})^r$，故可能的解表达为：

$$q_r = A + B(\frac{1-p}{p})^r$$

代入边界条件$q_0 = 1$和$q_n = 0$解得：

$$q_r = \frac{(\frac{1-p}{p})^n - (\frac{1-p}{p})^r}{(\frac{1-p}{p})^n - 1}$$

考虑$p = \frac{1}{2}$的情形，两个线性无关特解为$q_r = 1$和$q_r = r$，于是类似手段可以得到满足边界条件的解：

$$q_r = 1 - \frac{r}{n}$$

综上所述，破产的概率为：

$$q_k = \begin{cases} \frac{(\frac{1-p}{p})^n - (\frac{1-p}{p})^k}{(\frac{1-p}{p})^n - 1} & p \neq \frac{1}{2} \\ 1 - \frac{k}{n} & p = \frac{1}{2} \end{cases} \tag{3}$$

10.2 极限情况讨论

$p \leq 0.5$时，对$n \to \infty$取极限，有$q_k = 1$，与k无关.

11 生物灭亡

设一个此种生物最终灭亡的概率是p.
在此说明：这里p指的是m分钟后这一个生物及其全部可能后代全部死亡的概率对$m \to \infty$的极限. 由于这个概率应当对m单调不减而有上界1，故此极限存在.
考虑接下来等可能发生的三个事件：生物死亡，存活1个，或分裂成2个，其概率均为$\frac{1}{3}$，而这三种情形下的该生物及其可能后代全部死亡的概率分别为1，p，p^2（这里已经利用了每个此生物个体在每一时刻的命运相互独立的条件）.
故我们有

$$p = \frac{1}{3} + \frac{1}{3}p + \frac{1}{3}p^2$$

解得唯一解$p = 1$，故此生物经过足够长的时间后必然灭亡.

注记
若三个概率分别改为$\frac{1}{3}, 0, \frac{2}{3}$，则可以解出$p = \frac{1}{2}$和$p = 1$两个解，此时应取小者. 具体原因有待进一步讨论.

12 治疗方案

设两种治疗方案治愈该患者分别为事件W_1和W_2，患者患A、B、C病设为事件A、B、C. （注：这里W_1和W_2应当认为属于不同的概率空间，它们不会同时被选择）
A、B、C构成对样本空间的一个分割，则有

$$\begin{aligned} P(W_1) &= P(W_1|A)P(A) + P(W_1|B)P(B) + P(W_1|C)P(C) \\ &= 0.8 \times 0.8 + 0.05 \times 0.1 + 0.1 \times 0.1 = 0.655 \\ P(W_2) &= P(W_2|A)P(A) + P(W_2|B)P(B) + P(W_2|C)P(C) \end{aligned}$$

$$= 0.6 \times 0.8 + 0.9 \times 0.1 + 0.9 \times 0.1 = 0.66$$

故$P(W_1) < (W_2)$，纯从治愈率高的角度应当采取乙方案.
而考虑到该患者大概率患的是A病，而甲方案对A病治愈率很高，也可以推荐甲方案.
用p_{ij}表示该患者依次经过i，j两次治疗后治愈的概率，考虑两次治疗后治愈的概率：

$$p_{11} = 0.655 + 0.16 \times 0.8 + 0.095 \times 0.05 + 0.09 \times 0.1 = 0.797$$

$$p_{12} = 0.655 + 0.16 \times 0.6 + 0.095 \times 0.9 + 0.09 \times 0.9 = 0.918$$

$$p_{21} = 0.66 + 0.32 \times 0.8 + 0.01 \times 0.05 + 0.01 \times 0.1 = 0.918$$

$$p_{22} = 0.66 + 0.32 \times 0.6 + 0.01 \times 0.9 + 0.01 \times 0.9 = 0.870$$

这可以看出，实际上无论先使用甲方案还是乙方案，为了让两次完成治愈的概率较大，如果第一次没有治愈，第二次都应该选择另一种方案.

13 摸球

13.1

$$P(B_1) = \frac{1}{2}P(B_1|U_1) + \frac{1}{2}P(B_1|U_2) = \frac{6}{10} = \frac{3}{5}$$
$$P(U_1|B_1) = \frac{\frac{1}{2}P(B_1|U_1)}{P(B_1)} = \frac{2}{3}$$
$$P(U_1) = \frac{1}{2}$$

$P(U_1|B_1) > P(U_1)$，这是因为U_1中黑球所占比例大于U_2，因此如果摸出了黑球，更有理由认为是从U_1中摸出的.

13.2

若将第一个球放回，第二个球与第一个球无区别，故

$$P(B_2) = P(B_1) = \frac{3}{5}$$

若不放回，则需要考虑选择U_1还是U_2.

$$P(B_2|U_1) = P(B_2|U_1B_1)P(B_1|U_1) + P(B_2|U_1B_1^c)P(B_1^c|U_1)$$
$$= \frac{3}{4} \times \frac{4}{5} + \frac{4}{4} \times \frac{1}{5} = \frac{4}{5}$$
$$P(B_2|U_2) = P(B_2|U_2B_1)P(B_1|U_2) + P(B_2|U_2B_1^c)P(B_1^c|U_2)$$
$$= \frac{1}{4} \times \frac{2}{5} + \frac{2}{4} \times \frac{3}{5} = \frac{2}{5}$$

于是

$$P(B_2) = P(B_2|U_1)P(U_1) + P(B_2|U_2)P(U_2) = \frac{4}{5} \times \frac{1}{2} + \frac{2}{5} \times \frac{1}{2} = \frac{3}{5}$$

故仍有$P(B_2) = P(B_1)$.
事实上，无论最开始选取的袋子是哪个袋子，由于不知道第一次摸球的结果，无论放回还是不放回，第一次摸球和第二次摸球摸到黑球的概率都是一样的. 因此无论放回还是不放回，都有$P(B_2) = P(B_2)$.

13.3

此时B_1已经发生，根据第一问，有

$$P(U_1|B_1)=\frac{2}{3}$$

$$P(U_2|B_1)=1-P(U_1|B_1)=\frac{1}{3}$$

$$\begin{aligned}P(B_2|B_1)&=P(B_2|U_1B_1)P(U_1|B_1)+P(B_2|U_2B_1)P(U_2|B_1)\\&=\frac{4}{5}\times\frac{2}{3}+\frac{2}{5}\times\frac{1}{3}=\frac{2}{3}\end{aligned}$$

于是，$P(B_2|B_1)>P(B_2)$，此时因为已知第一个球摸出黑球，第一问中已做解释，此时实际提高了取到U_1袋的后验概率，进一步也就提高了第二次摸到黑球的概率.

13.4

$$\begin{aligned}P(U_1|B_1B_2...B_n)&=\frac{P(B_1B_2...B_n|U_1)P(U_1)}{P(B_1B_2...B_n|U_1)P(U_1)+P(B_1B_2...B_n|U_2)P(U_2)}\\&=\frac{(\frac{4}{5})^n\times\frac{1}{2}}{(\frac{4}{5})^n\times\frac{1}{2}+(\frac{2}{5})^n\times\frac{1}{2}}=\frac{(\frac{4}{5})^n}{(\frac{4}{5})^n+(\frac{2}{5})^n}\end{aligned}$$

于是

$$P(U_2|B_1B_2...B_n)=1-P(U_1|B_1B_2...B_n)=\frac{(\frac{2}{5})^n}{(\frac{4}{5})^n+(\frac{2}{5})^n}$$

故

$$\begin{aligned}P(B_{n+1}|B_1...B_n)&=P(B_{n+1}|B_1...B_nU_1)P(U_1|B_1...B_n)+P(B_{n+1}|B_1...B_nU_2)P(U_2|B_1...B_n)\\&=\frac{4}{5}\times\frac{(\frac{4}{5})^n}{(\frac{4}{5})^n+(\frac{2}{5})^n}+\frac{2}{5}\times\frac{(\frac{2}{5})^n}{(\frac{4}{5})^n+(\frac{2}{5})^n}=\frac{(\frac{4}{5})^{n+1}+(\frac{2}{5})^{n+1}}{(\frac{4}{5})^n+(\frac{2}{5})^n}\end{aligned}$$

令$n\to\infty$，有

$$\lim_{n\to\infty}P(B_{n+1}|B_1...B_n)=\frac{4}{5}$$

事实上，随着连续摸到黑球的次数增多，对选到的袋子是U_1的后验概率就越来越大，如果连续无数次摸到黑球，则此后验概率就是1，于是摸到黑球的概率就是$P(B_1|U_1)=\frac{4}{5}$.

13.5

上一问已经进行了计算：

$$P(U_1|B_1...B_n)=\frac{(\frac{4}{5})^n}{(\frac{4}{5})^n+(\frac{2}{5})^n}$$

同时

$$\lim_{n\to\infty}P(U_1|B_1...B_n)=1$$

这个理解在上一问也有提到，因为1号袋中黑球比例高于2号袋，因此如果连续摸到黑球的数目足够多，则选到1号袋的后验概率会逐渐增加直到增加到1.

14 对赌

14.1

假设100元中用x元跟甲对赌，用$100-x$元跟乙对赌，则若A获胜，会获利$-x+(100-x)\times\frac{15}{10}=150-\frac{5}{2}x$元，若B获胜，会获利$x\times\frac{20}{5}-(100-x)=5x-100$元. 要想必定获利，要求上述两个获利值都为正，即

$$150-\frac{5}{2}x>0$$

且

$$5x-100>0$$

解得

$$20<x<60$$

故要想必定获利，不能只和一家对赌，而要出20至60元与甲赌，出40至80元与乙赌.

14.2

甲、乙的主观概率应满足两人的期望收益为0. 故

$$-20P_1(B)+5(1-P_1(B))=0$$

$$-15P_2(A)+10(1-P_2(A))=0$$

解得

$$P_1(B)=\frac{1}{5}$$

$$P_2(A)=\frac{2}{5}$$

于是

$$P_1(B)+P_2(A)=\frac{3}{5}<1$$

这说明，两人都对自己支持的队伍的胜率有过高的估计，这导致他们愿意出大金额对赌，从而使得存在与他们同时对赌而必定获利的方式. 若$P_1(B)+P_2(A)\geq 1$，则不会有这种情况出现.

15 计算机实验：掷硬币

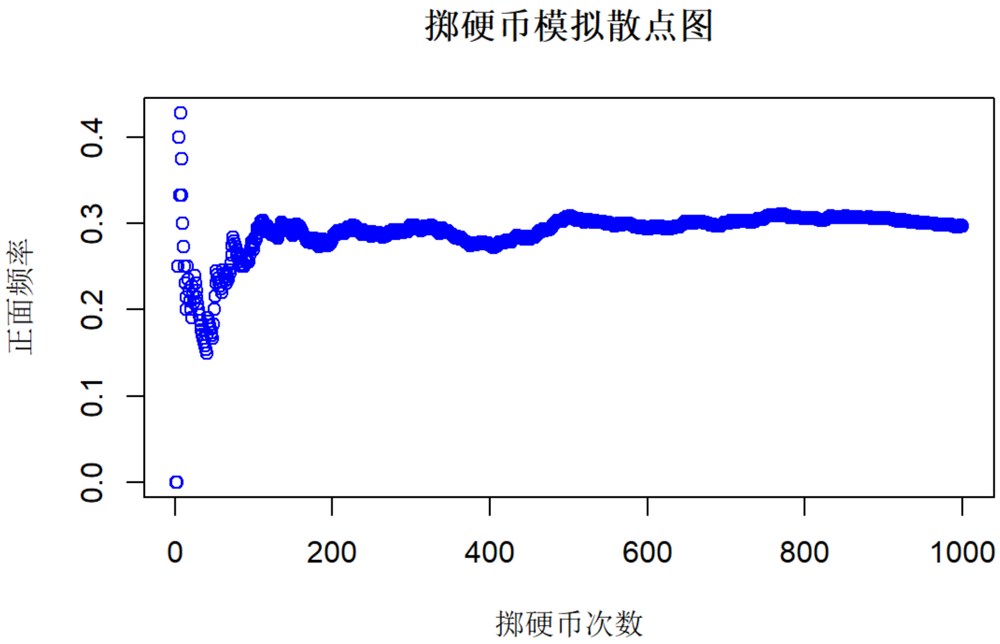

Figure 1: 一次试验的相对频数散点图

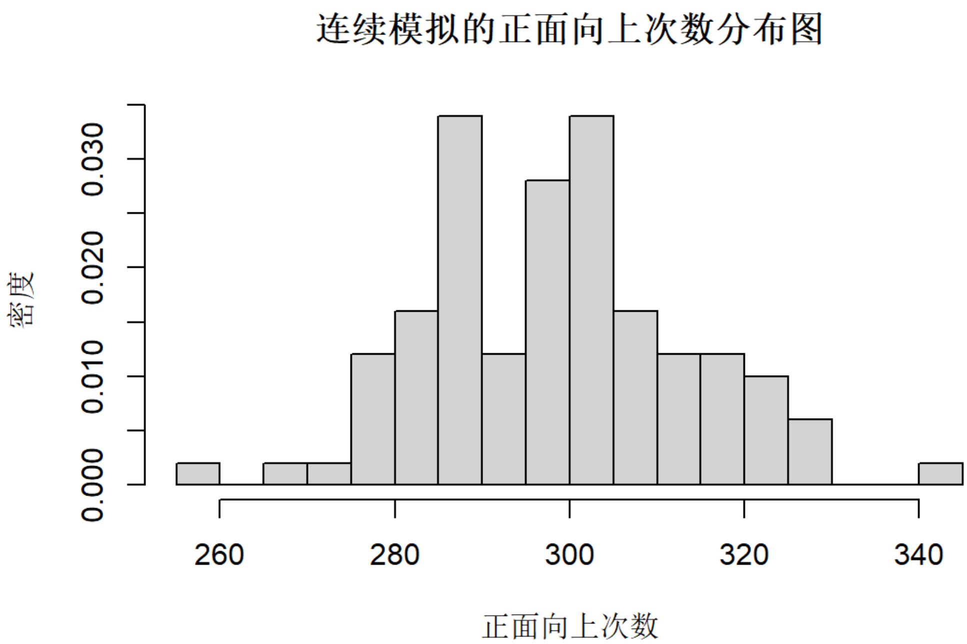

Figure 2: 多次试验的正面向上次数直方图

```r
#进行单次试验并画出散点图
n <- 1000
p <- 0.3
coin <- sample(c("H", "T"), size = n, replace = TRUE, prob=c(p,1-p))
count <- cumsum(coin == "H")
freq <- count / (1:n)
plot(1:n, freq, type = "p", col = "blue",
     xlab = "掷硬币次数", ylab = "正面频率", main = "掷硬币模拟散点图")
#进行多次试验并画出直方图
m <- 100
counts <- vector()
for(i in 1:m){
  n<-1000
  p<-0.3
  coin<-sample(c("H","T"),size=n,replace=TRUE,prob=c(p,1-p))
  counts[i]<-sum(coin=="H")
}
hist(counts,breaks=20,freq=FALSE,xlab="正面向上次数",ylab="密度",main="连续模拟的正面向上次数分布图")
```

Figure 3: 完成实验的R语言代码