概率论第 3 次习题课讲义 (国庆特辑)

宗语轩

2022 秋, USTC

1　作业讲评

3.2.5　这里只考虑离散型:

$$\begin{aligned}\mathbb{P}\left(\sum_{r=1}^{n} X_r \geqslant x\right) &= \sum_{a_1+\cdots+a_n \geqslant x} \mathbb{P}(X_1 = a_1, X_2 = a_2, \cdots, X_n = a_n)\\ &= \sum_{a_1+\cdots+a_n \geqslant x} \mathbb{P}(X_1 = a_1)\mathbb{P}(X_2 = a_2)\cdots\mathbb{P}(X_n = a_n)\\ &= \sum_{a_1+\cdots+a_n \geqslant x} \mathbb{P}(X_1 = -a_1)\mathbb{P}(X_2 = -a_2)\cdots\mathbb{P}(X_n = -a_n)\\ &= \sum_{-a_1-\cdots-a_n \leqslant -x} \mathbb{P}(X_1 = -a_1)\mathbb{P}(X_2 = -a_2)\cdots\mathbb{P}(X_n = -a_n)\\ &= \sum_{-a_1-\cdots-a_n \leqslant -x} \mathbb{P}(X_1 = -a_1, X_2 = -a_2, \cdots, X_n = -a_n)\\ &= \mathbb{P}\left(\sum_{r=1}^{n} X_r \leqslant -x\right).\end{aligned}$$

补充: 连续型类似 (只不过变成 joint density function 以及转化为积分形式表示). 一般情况较难处理, 这里不作要求. 大致思路: 只需考虑 $n = 2$, 一般情形归纳即可. 而

$$\{X_1 + X_2 > x\} = \bigcup_{\substack{q_1, q_2 \in \mathbb{Q}\\ q_1+q_2 > x}} \{X_1 \geqslant q_1, X_2 \geqslant q_2\}$$

将 $-x$ 替换成 x, 先分别说明两者对应的集合概率测度相等, 然后再说明并起来后集合概率测度相等. 有限并无需赘言, 无限并可用集合升链的概率测度收敛的性质 (反复拆分 + 合并).

3.3.3 (a)　使用方差的线性运算需要验证独立性.

---

[0]个人主页: http://home.ustc.edu.cn/~zyx240014/.
发现错误欢迎联系: zyx240014@mail.ustc.edu.cn.

补充题 1 $X \sim P(\lambda)$, 求 k 为何值时, $\mathbb{P}(X=k)$ 最大?

解. 注意到

$$\frac{\mathbb{P}(X=k+1)}{\mathbb{P}(X=k)} = \frac{\lambda}{k+1}, \qquad \frac{\mathbb{P}(X=k)}{\mathbb{P}(X=k-1)} = \frac{\lambda}{k}.$$

可知 $k=[\lambda]$ 时, $\mathbb{P}(X=k)$ 最大. (若 λ 为整数, 则两头可同时取到) □

补充题 2 先掷一次均匀骰子, 再抛均匀硬币, 硬币抛掷次数为掷出的骰子点数. 求硬币正面个数 H 的分布列.

解: $P(X=k) = \sum_{i=1}^{6} P(\text{掷出点数为}i)P(\text{抛}i\text{次抛出}k\text{个 H 朝上}) = \frac{1}{6}\sum_{i=k}^{6} \mathrm{C}_i^k(\frac{1}{2})^i$, 于是分布列为

X	0	1	2	3	4	5	6
P	$\frac{63}{384}$	$\frac{120}{384}$	$\frac{99}{384}$	$\frac{64}{384}$	$\frac{29}{384}$	$\frac{8}{384}$	$\frac{1}{384}$

补充题 3 随机变量 $X \sim B(n,p)$, 求 $\mathbb{E}[X^3]$.

解. 利用

$$\mathbb{E}[X] = np, \qquad \mathbb{E}[X(X-1)] = n(n-1)p^2, \qquad \mathbb{E}[X(X-1)(X-2)] = n(n-1)(n-2)p^3$$

可得

$$\mathbb{E}[X^3] = \mathbb{E}[X(X-1)(X-2)] + 3\mathbb{E}[X(X-1)] + \mathbb{E}[X] = n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np.$$

□

另解. 先做分解:

$$X = X_1 + X_2 + \cdots + X_n, \qquad X_r \text{ i.i.d } \sim B(1,p),\ r = 1, \cdots, n.$$

因此, 我们有

$$\begin{aligned}
\mathbb{E}[X^3] = \mathbb{E}[(X_1+X_2+\cdots+X_n)^3] &= \sum_{1\leqslant i,j,k\leqslant n} \mathbb{E}[X_iX_jX_k] \\
&= \sum_{1\leqslant i=j=k\leqslant n} \mathbb{E}[X_i^3] = \sum_{\substack{1\leqslant i,j\leqslant n \\ i\neq j}} \mathbb{E}[X_i^2X_j] + \sum_{\substack{1\leqslant i,j,k\leqslant n \\ i\neq j\neq k}} \mathbb{E}[X_iX_jX_k] \\
&= \sum_{k=1}^{n} \mathbb{E}[X_k^3] + 6\sum_{1\leqslant i<j\leqslant n} \mathbb{E}[X_i^2]\mathbb{E}[X_j] + 6\sum_{1\leqslant i<j<k\leqslant n} \mathbb{E}[X_i]\mathbb{E}[X_j]\mathbb{E}[X_k] \\
&= np + 3n(n-1)p^2 + n(n-1)(n-2)p^3.
\end{aligned}$$

□

2 专题选讲

2.1 组合恒等式

引理 2.1. 我们有

(1) $\binom{n}{k} = \binom{n}{n-k}$ $(0 \leqslant k \leqslant n)$.

(2) $k\binom{n}{k} = n\binom{n-1}{k-1}$.

(3) $\binom{n}{m} + \binom{n}{m+1} = \binom{n+1}{m+1}$

提示. (3) 即为 $n+1$ 个盒子选出 $m+1$ 个的组合总数, 考虑最后一个是否被选出, 两种情况的组合总数相加即可.

定理 2.1 (二项式定理)**.** 对 $\forall x \in \mathbb{R}$ 及 $n \in \mathbb{N}^*$, 有

$$(1+x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i.$$

提示. 归纳. 利用**引理 2.1(3)**, 我们有

$$(1+x)^n = (1+x)\cdot(1+x)^{n-1} = (1+x)\sum_{i=0}^{n-1}\binom{n-1}{i}x^i = \sum_{i=1}^{n-1}\left(\binom{n-1}{i}+\binom{n-1}{i-1}\right)x^i+1+x^n = \sum_{i=0}^{n}\binom{n}{i}x^i.$$

引理 2.2. 利用上述引理和定理, 我们得到

(1) $\sum\limits_{k=0}^{n}\binom{n}{k} = 2^n$.

(2) $\sum\limits_{\substack{0\leqslant k\leqslant n \\ k\text{为偶数}}}\binom{n}{k} = \sum\limits_{\substack{0\leqslant k\leqslant n \\ k\text{为奇数}}}\binom{n}{k} = 2^{n-1}$.

(3) $\binom{m+n}{k} = \sum\limits_{i=0}^{k}\binom{m}{i}\binom{n}{k-i}$. 特别地, $\binom{2n}{n} = \sum\limits_{i=0}^{n}\binom{n}{i}\binom{n}{n-i} = \sum\limits_{i=0}^{n}\binom{n}{i}^2$.

(4) $\sum\limits_{k=0}^{n} k\binom{n}{k} = n2^{n-1}$.

(5) $\sum\limits_{k=m}^{n}\binom{k}{m} = \binom{n+1}{m+1}$.

(6) $\sum\limits_{k=0}^{n}\binom{m+k}{k} = \binom{m+n+1}{n}$.

提示. (1) (2) 对 $(1+x)^n$ 运用二项式定理, 分别取 $x=1,-1$ 即可. (3) 注意到 $(1+x)^{m+n}=(1+x)^m(1+x)^n$, 分别运用二项式定理, 考察等式两边 x^k 项的系数. (4)$(1+x)^n=\sum_{i=0}^{n}\binom{n}{i}x^i$ 两侧对 x 求导, 并取 $x=1$ 即可. (5) (6) 利用**引理 2.1(3)** 归纳即可.

例 2.1. 设 $n\geq m$. 证明:

$$\sum_{k=0}^{m}\binom{m}{k}\binom{n+k}{m}=\sum_{k=0}^{m}\binom{m}{k}\binom{n}{k}2^k.$$

证明.

$$\begin{aligned}\sum_{k=0}^{m}\binom{m}{k}\binom{n+k}{m}&=[x^m](1+x)^n\sum_{k=0}^{m}\binom{m}{k}(1+x)^k=[x^m](1+x)^n(2+x)^m\\&=[x^m](1+x)^n\sum_{k=0}^{m}\binom{m}{k}2^kx^{m-k}\\&=\sum_{k=0}^{m}\binom{m}{k}\binom{n}{k}2^k.\end{aligned}$$

其中 “$[x^m]f(x)$” 表示 $f(x)$ 中项 x^m 的系数. □

注. 也可以用组合方法 (~~实质是通过带有二项式定理的代数方法来编故事 (bushi~~):

已知 m 个男孩和 n 个女孩, 并满足条件:

(1) 给一些 (数量未知) 男孩吃苹果;

(2) 从所有人中选出 m 个人吃梨;

(3) 吃了梨的男孩必须也吃了苹果.

假设每个苹果 (梨) 不作区分, 每个人至多拿 1 个苹果及 1 个梨. 问一共有多少种符合条件的分法?

一方面, 设 (1) 中给 k 个男孩苹果, 其分法有 $\binom{m}{k}$ 种, 由 (3) 知, 另外没吃苹果的 $m-k$ 个男孩不可能吃梨, 结合 (2), 剩下梨的分法有 $\binom{n+k}{m}$ 种. 对 k 累加得

$$\sum_{k=0}^{m}\binom{m}{k}\binom{n+k}{m}.$$

另一方面, 设 (2) 中给 k 个女孩梨, 则给 $m-k$ 个男孩梨. 分法分别为 $\binom{n}{k}$ 和 $\binom{m}{m-k}=\binom{m}{k}$ 种. 由 (3) 知吃了梨的男孩一定吃了苹果. 剩下 k 个男孩可以吃苹果也可以不吃苹果, 分法共 2^k 种. 对 k 累加得

$$\sum_{k=0}^{m}\binom{m}{k}\binom{n}{k}2^k.$$

2.2 示性函数与概率论

示性函数: $I_A(\omega)=\begin{cases}1, & \omega\in A,\\ 0, & \omega\notin A.\end{cases}$, 是一种随机变量. 例子: $B(1,p)$. 先从集合的观点看起:

引理 2.3. 对集合 $A,B,A_j(j\in J)$, 有

(1) $I_A^k=I_A(k\in N^*)$.

(2) $I_{A\cap B}=I_AI_B,\ I_{A^c}=1-I_A,\ I_{A\cup B}=I_A+I_B-I_AI_B,\ I_{A\setminus B}=I_A(1-I_B)$.

(3) $I_{A\Delta B}=|I_A-I_B|$.

(4) $I_{\bigcup_j A_j}=1-\prod\limits_{j\in J}(1-I_{A_j})$.

提示. (4) 利用 De.Morgan 法则: $\bigcup\limits_{i\in I}A_i=\left(\bigcap\limits_{i\in I}A_i^c\right)^c$.

转化方式:

(1) 对事件 A, $\mathbb{P}(A)=\mathbb{E}[I_A]$, 且有 $\mathbb{E}[I_A]=\mathbb{E}[I_A^k](k\in N^*)$, 算方差时取 $k=2$.

(2) 分解成其他的随机变量之和. 例如: 对离散型随机变量 X, 有 $X=\sum\limits_x xI_{\{X=x\}}$ 及

$$\mathbb{E}[X]=\sum_x x\mathbb{P}(X=x)=\sum_x x\mathbb{E}[I_{\{X=x\}}]=\mathbb{E}[\sum_x xI_{\{X=x\}}].$$

也可以分解成若干示性函数的和. 由此可用来计算方差或其他 k 阶矩 (参考作业题 **3.3.3(a)**).

动机: 利用期望的性质, 尤其是线性性.

引理 2.4. 对非负整值随机变量 X, 有

$$\mathbb{E}[X]=\sum_{n=0}^{\infty}\mathbb{P}(X>n).$$

证明. 利用 Fubini 定理, 我们有

$$\mathbb{E}[X]=\mathbb{E}[\sum_{n=0}^{X-1}1]=\mathbb{E}[\sum_{n=0}^{\infty}I_{\{X>n\}}]\xlongequal{\text{Fubini}}\sum_{n=0}^{\infty}\mathbb{E}[I_{\{X>n\}}]=\sum_{n=0}^{\infty}\mathbb{P}(X>n).$$

□

注. 求分布列的常见转化方法:

$$\mathbb{P}(X=n)=\mathbb{P}(X\geqslant n)-\mathbb{P}(X\geqslant n+1)=\mathbb{P}(X\leqslant n)-\mathbb{P}(X\leqslant n-1)$$

$$\begin{aligned}\mathbb{P}(X=m,Y=n)&=\mathbb{P}(X\geqslant m,Y\geqslant n)-\mathbb{P}(X\geqslant m+1,Y\geqslant n)-\mathbb{P}(X\geqslant m,Y\geqslant n+1)\\&\quad+\mathbb{P}(X\geqslant m+1,Y\geqslant n+1)\\&=\mathbb{P}(X\geqslant m,Y\leqslant n)-\mathbb{P}(X\geqslant m+1,Y\leqslant n)-\mathbb{P}(X\geqslant m,Y\leqslant n-1)\\&\quad+\mathbb{P}(X\geqslant m+1,Y\leqslant n-1).\end{aligned}$$

例 2.2. 利用**引理 2.3 (4)**, 证明**Jordan 公式**:

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right)=\sum_{k=1}^n(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}\mathbb{P}(A_{i_1}\cdots A_{i_k}).$$

证明. 利用

$$I_{\bigcup_j A_j}=1-\prod_{j\in J}(1-I_{A_j})=\sum_{k=1}^n(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}I_{A_{i_1}}\cdots I_{A_{i_k}}=\sum_{k=1}^n(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}I_{A_{i_1}\cap\cdots\cap I_{A_{i_k}}},$$

两边同时取期望并利用期望的线性性即得. □

例 2.3. 一座楼共 n 层, 有一个载有 m 人的升降电梯在这 n 层楼里停靠若干次, 每个人随机选择一层楼停靠. 求电梯平均停靠的总次数.

解. 记 T 为电梯停靠的次数, 直接求 T 的分布非常困难, 但我们可以将其分解成 n 个示性函数之和:

$$T=I_{A_1}+\cdots+I_{A_n},$$

其中事件 A_i 表示第 i 层楼有人停. 故有

$$\mathbb{E}[T]=\sum_{i=1}^n\mathbb{E}[I_{A_i}]=\sum_{i=1}^n\mathbb{P}(A_i)=\sum_{i=1}^n\left(1-\left(\frac{n-1}{n}\right)^m\right)=n\left(1-\left(1-\frac{1}{n}\right)^m\right).$$

□

2.3 条件期望的应用

定义 2.1. 设

$$\psi(x):=\mathbb{E}[Y|X=x]=\sum_y yf_{Y|X}(y|x).$$

称 $\psi(X)$ 为 Y 关于 X 的**条件期望**, 记 $\mathbb{E}[Y|X]$.

注. $\mathbb{E}[Y|X]$ 为一个与 X 有关的**随机变量**. 由定义知, 求条件期望 $\mathbb{E}[Y|X]$ 时往往先求 $\mathbb{E}[Y|X=x]$ 的表达式, 再用 X 代替表达式里的 x, 得到的新表达式即为 $\mathbb{E}[Y|X]$.

定理 2.2. 设 X,Y 为离散型随机变量, 则有

$$\mathbb{E}[\mathbb{E}[Y|X]]=\mathbb{E}[Y].$$

例 2.4. 多项分布: $\boldsymbol{X}=(X_1,\cdots,X_r)$. $\mathbb{P}(X_1=k_1,\cdots,X_r=k_r)=\frac{n!}{k_1!\cdots k_r!}p_1^{k_1}\cdots p_r^{k_r}$, $\sum_{i=1}^r k_i=n$, $\sum_{i=1}^r p_i=1$ 且 $p_i>0$. 现对 $i\neq j$,

(1) 求 $Cov(X_i,X_j)$, $\rho(X_i,X_j)$.

(2) 求 $\mathbb{E}[X_j|X_i>0]$.

解. (1) 由题意得, $X_i\sim B(n,p_i), X_i+X_j\sim B(n,p_i+p_j)$. 故有

$\mathbb{E}[X_i]=np_i,\quad \mathbb{E}[X_i+X_j]=n(p_i+p_j),\quad Var(X_i)=np_i(1-p_i),\quad Var(X_i+X_j)=n(p_i+p_j)(1-p_i-p_j).$

再利用 $Var(X_i+X_j)=Var(X_i)+Var(X_j)+2Cov(X_i,X_j)$, 得

$$Cov(X_i,X_j)=-np_ip_j,\quad \rho(X_i,X_j)=\frac{-np_i\cdot np_j}{\sqrt{np_i(1-p_i)np_j(1-p_j)}}=-\sqrt{\frac{p_ip_j}{(1-p_i)(1-p_j)}}.$$

(2) 利用条件期望的性质, 我们有

$$\mathbb{E}[X_j]=\mathbb{E}[\mathbb{E}[X_j|X_i]]=\mathbb{E}[X_j|X_i>0]\mathbb{P}(X_i>0)+\mathbb{E}[X_j|X_i=0]\mathbb{P}(X_i=0).$$

而

$$\begin{aligned}\mathbb{P}(X_j=k|X_i=0)&=\frac{\mathbb{P}(X_j=k,X_i=0)}{\mathbb{P}(X_i=0)}=\frac{n!}{k!0!(n-k)!}p_j^kp_i^0(1-p_i-p_j)^{n-k}\cdot\frac{1}{(1-p_i)^n}\\&=\binom{n}{k}\left(\frac{p_j}{1-p_i}\right)^k\left(1-\frac{p_j}{1-p_i}\right)^{n-k}.\end{aligned}$$

因此,

$$\mathbb{E}[X_j|X_i=0]=\frac{np_j}{1-p_i}.$$

利用 $\mathbb{E}[X_j]=np_j$, $\mathbb{P}(X_i=0)=(1-p_i)^n$, $\mathbb{P}(X_i>0)=1-(1-p_i)^n$, 有

$$\mathbb{E}[X_j|X_i>0]=\frac{np_j(1-(1-p_i)^{n-1})}{1-(1-p_i)^n}.$$

□

注. 补充协方差中两个很有用的结论:

1. $Var\left(\sum_{i=1}^n X_i\right)=\sum_{i=1}^n Var(X_i)+2\sum_{1\leqslant i<j\leqslant n}Cov(X_i,X_j).$
2. 双线性性: $Cov(aX+bY,Z)=aCov(X,Z)+bCov(Y,Z), Cov(X,aY+bZ)=aCov(X,Y)+bCov(X,Z)$.

我们可以从 “投影” 的角度理解条件期望, 下面事实说明 $\mathbb{E}[Y|X]=\arg\min_{g(x)}\mathbb{E}[(Y-g(x))^2]$. 换言之, 假设 “$X$” 是你已知的信息, 此时 “$Y$” 可以做最好的估计:

例 2.5. 设 (X,Y) 是联合离散随机向量, X 与 Y 的二阶矩存在, 记 $\psi(X) := \mathbb{E}[Y|X]$. 若 g 为可测函数且 $g(X)$ 的二阶矩存在, 证明:

$$\mathbb{E}[(Y-\psi(X))^2] \leqslant \mathbb{E}[(Y-g(X))^2]$$

解. 我们有

$$\begin{aligned}\mathbb{E}[(Y-g(X))^2] &= \mathbb{E}[(Y-\psi(X)+\psi(X)-g(X))^2] \\ &= \mathbb{E}[(Y-\psi(X))^2] + 2\mathbb{E}[(Y-\psi(X))(\psi(X)-g(X))] + \mathbb{E}[(Y-g(X))^2].\end{aligned}$$

利用 $\mathbb{E}[(Y-g(X))^2] \geqslant 0$ 以及

$$\begin{aligned}\mathbb{E}[(Y-\psi(X))(\psi(X)-g(X))] &= \mathbb{E}[\mathbb{E}[(Y-\psi(X))(\psi(X)-g(X))|X]] \\ &= \mathbb{E}[(\psi(X)-g(X))\mathbb{E}[(Y-\psi(X))|X]] \\ &= \mathbb{E}[(\psi(X)-g(X))(\mathbb{E}[Y|X]-\psi(X))] \\ &= 0.\end{aligned}$$

□

注. 类似地, 我们亦可以从 "投影" 的角度理解方差: $Var(X) = \min\limits_{a\in\mathbb{R}} \mathbb{E}[(X-a)^2]$. 因为

$$\mathbb{E}[(X-a)^2] = a^2 - 2a\mathbb{E}[X] + \mathbb{E}[X^2] = (a-\mathbb{E}[X])^2 + Var(X) \geqslant Var(X),$$

等号当且仅当 $a = \mathbb{E}[X]$ 时取等, 此时即为 $Var(X)$ 的形式.

例 2.6. 独立重复地掷一个均匀骰子, 按序记录投掷的点数. 记 τ_{ij} 表示首次出现 "ij" 时骰子已投掷的总次数, 其中 $1 \leqslant i,j \leqslant 6$. 求 $\mathbb{E}[\tau_{11}], \mathbb{E}[\tau_{12}]$.

解. 记 N_1, N_2 分别表示第 $1,2$ 次骰子投掷的点数, τ_i 表示首次出现 "i" 时骰子已投掷的总次数, 则有

$$\begin{aligned}\mathbb{E}[\tau_{11}] &= \mathbb{E}[\mathbb{E}[\tau_{11}|N_1]] = \mathbb{E}[\tau_{11}|N_1=1]\mathbb{P}(N_1=1) + \mathbb{E}[\tau_{11}|N_1\neq 1]\mathbb{P}(N_1\neq 1) \\ &= \mathbb{E}[\mathbb{E}[\tau_{11}|N_1=1,N_2]]\cdot\frac{1}{6} + (1+\mathbb{E}[\tau_{11}])\cdot\frac{5}{6} \\ &= (\mathbb{E}[\tau_{11}|N_1=1,N_2=1]\mathbb{P}(N_2=1) + \mathbb{E}[\tau_{11}|N_1=1,N_2\neq 1]\mathbb{P}(N_2\neq 1))\cdot\frac{1}{6} + (1+\mathbb{E}[\tau_{11}])\cdot\frac{5}{6} \\ &= \left(2\cdot\frac{1}{6} + (2+\mathbb{E}[\tau_{11}])\cdot\frac{5}{6}\right)\cdot\frac{1}{6} + (1+\mathbb{E}[\tau_{11}])\cdot\frac{5}{6} \\ &= \frac{35}{36}\mathbb{E}[\tau_{11}] + \frac{7}{6}\end{aligned}$$

因此

$$\mathbb{E}[\tau_{11}] = 42.$$

类似地, 有

$$\begin{aligned}\mathbb{E}[\tau_{12}] &= \mathbb{E}[\mathbb{E}[\tau_{12}|N_1]] = \mathbb{E}[\tau_{12}|N_1=1]\mathbb{P}(N_1=1) + \mathbb{E}[\tau_{12}|N_1\neq 1]\mathbb{P}(N_1\neq 1)\\ &= \mathbb{E}[\tau_{12}|N_1=1]\cdot\frac{1}{6} + (1+\mathbb{E}[\tau_{12}])\cdot\frac{5}{6}.\end{aligned}$$

同时, 我们有

$$\begin{aligned}\mathbb{E}[\tau_{12}|N_1=1] &= \mathbb{E}[\mathbb{E}[\tau_{12}|N_1=1,N_2]]\\ &= \mathbb{E}[\tau_{12}|N_1=1,N_2=1]\mathbb{P}(N_2=1) + \mathbb{E}[\tau_{12}|N_1=1,N_2=2]\mathbb{P}(N_2=2)\\ &+ \mathbb{E}[\tau_{12}|N_1=1,N_2\neq 1,2]\mathbb{P}(N_2\neq 1,2)\\ &= (1+\mathbb{E}[\tau_{12}|N_1=1])\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + (2+\mathbb{E}[\tau_{12}])\cdot\frac{4}{6}\\ &= \frac{1}{6}\mathbb{E}[\tau_{12}|N_1=1] + \frac{2}{3}\mathbb{E}[\tau_{12}] + \frac{11}{6}.\end{aligned}$$

将上述两式联立可得, $\mathbb{E}[\tau_{12}] = 36$. □

2.4 概率方法举例

例 2.7 (涂色问题)**.** 平面上的 n 个点和连接各点之间的连线叫做一个完全图, 记作 G. 点称作图的顶点, 顶点之间的连线叫做边, 共有 $\binom{n}{2}$ 条. 给定一个整数 k, G 中任意 k 个顶点连同相应的边构成一个有 k 个顶点的完全子图, G 中共有 $\binom{n}{k}$ 个这样的子图, 记作 $G_i, i=1,\cdots,\binom{n}{k}$. 现将图 G 的每条边涂成红色或蓝色. 当 $\binom{n}{k} < 2^{\frac{k(k-1)}{2}-1}$ 时, 问: 是否有一种涂色方法, 使得没有一个子图 G_i 的 $\binom{k}{2}$ 条边是同一颜色的.

解. 我们对 G 的边进行随机涂色, 每条边为红色和蓝色的概率是 $\frac{1}{2}$. 记事件 A_i 表示子图 G_i 各边的颜色相同, 则 $\bigcap_{i=1}^{\binom{n}{k}} A_i^c$ 表示没有一个子图 G_i 的 $\binom{k}{2}$ 条边是同一颜色的. 利用

$$\mathbb{P}(A_i) = \mathbb{P}(G_i\text{的各边均为红色}) + \mathbb{P}(G_i\text{的各边均为蓝色}) = \frac{2}{2^{\binom{k}{2}}} = 2^{-\frac{k(k-1)}{2}+1}$$

及 $\binom{n}{k} < 2^{\frac{k(k-1)}{2}-1}$, 有

$$\mathbb{P}\left(\bigcup_{i=1}^{\binom{n}{k}} A_i\right) \leqslant \sum_{i=1}^{\binom{n}{k}} \mathbb{P}(A_i) = \binom{n}{k}2^{-\frac{k(k-1)}{2}+1} < 1$$

故 $\mathbb{P}\left(\bigcap_{i=1}^{\binom{n}{k}} A_i^c\right) > 0$. 从而存在一种涂色方法, 使得没有一个子图 G_i 的 $\binom{k}{2}$ 条边是同一颜色的. □

例 2.8 (Balancing vectors)**.** 设向量 $\boldsymbol{v_1}, \boldsymbol{v_2}, \cdots, \boldsymbol{v_n} \in \mathbb{R}^n$.

(1) 若对任意 $1 \leqslant i \leqslant n$ 满足 $|\boldsymbol{v_i}| = 1$. 证明: 存在 $\varepsilon_i \in \{-1, 1\}$, 使得

$$\left|\sum_{i=1}^{n} \varepsilon_i \boldsymbol{v_i}\right| \leqslant \sqrt{n} \text{ (这里换成 } \geqslant \text{ 亦成立).}$$

(2) 若对任意 $1 \leqslant i \leqslant n$ 满足 $|\boldsymbol{v_i}| \leqslant 1$. 对 $\boldsymbol{w} = \sum_{i=1}^{n} p_i \boldsymbol{v_i}$, 其中 $p_i \in [0, 1]$. 证明: 存在 $\varepsilon_i \in \{0, 1\}$, 使得

$$\left|\sum_{i=1}^{n} \varepsilon_i \boldsymbol{v_i} - \boldsymbol{w}\right| \leqslant \frac{\sqrt{n}}{2}.$$

证明. (1) 令 ε_i i.i.d 且满足 $\mathbb{P}(\varepsilon_i = 1) = \mathbb{P}(\varepsilon_i = -1) = \frac{1}{2}$. 记 $X = \left|\sum_{i=1}^{n} \varepsilon_i \boldsymbol{v_i}\right|$. 则有

$$\begin{aligned}\mathbb{E}[X^2] = \sum_{1 \leqslant i,j \leqslant n} \boldsymbol{v}_i \cdot \boldsymbol{v}_j \mathbb{E}[\varepsilon_i \varepsilon_j] = \sum_{i=1}^{n} \boldsymbol{v}_i \cdot \boldsymbol{v}_i \mathbb{E}[\varepsilon_i^2] + \sum_{1 \leqslant i \neq j \leqslant n} \boldsymbol{v}_i \cdot \boldsymbol{v}_j \mathbb{E}[\varepsilon_i]\mathbb{E}[\varepsilon_j] = \sum_{i=1}^{n} |\boldsymbol{v}_i|^2 \\ = n.\end{aligned}$$

因此, 一定存在 $\varepsilon_i \in \{-1, 1\}$, 使得

$$\left|\sum_{i=1}^{n} \varepsilon_i \boldsymbol{v_i}\right| \leqslant \sqrt{n}.$$

(2) 令 ε_i 相互独立且满足 $\mathbb{P}(\varepsilon_i = 1) = p_i, \mathbb{P}(\varepsilon_i = 0) = 1 - p_i$. 记 $X = \left|\sum_{i=1}^{n} \varepsilon_i \boldsymbol{v_i} - \boldsymbol{w}\right|$. 注意到

$$X^2 = \left|\sum_{i=1}^{n} (\varepsilon_i - p_i) \boldsymbol{v_i}\right|^2 = \sum_{1 \leqslant i,j \leqslant n} \boldsymbol{v}_i \cdot \boldsymbol{v}_j (\varepsilon_i - p_i)(\varepsilon_j - p_j).$$

因此有

$$\begin{aligned}\mathbb{E}[X^2] &= \sum_{1 \leqslant i,j \leqslant n} \boldsymbol{v}_i \cdot \boldsymbol{v}_j \mathbb{E}[(\varepsilon_i - p_i)(\varepsilon_j - p_j)] \\ &= \sum_{i=1}^{n} \boldsymbol{v}_i \cdot \boldsymbol{v}_i \mathbb{E}[(\varepsilon_i - p_i)^2] + \sum_{1 \leqslant i \neq j \leqslant n} \boldsymbol{v}_i \cdot \boldsymbol{v}_j \mathbb{E}[(\varepsilon_i - p_i)]\mathbb{E}[(\varepsilon_j - p_j)] \\ &= \sum_{i=1}^{n} |\boldsymbol{v}_i|^2 Var(\varepsilon_i) \leqslant \sum_{i=1}^{n} \frac{|\boldsymbol{v}_i|^2}{4} \\ &\leqslant \frac{n}{4}.\end{aligned}$$

故一定存在 $\varepsilon_i \in \{0, 1\}$, 使得

$$\left|\sum_{i=1}^{n} \varepsilon_i \boldsymbol{v_i} - \boldsymbol{w}\right| \leqslant \frac{\sqrt{n}}{2}.$$

□

例 2.9 (**Erdös, 1965. 组合数论**)**.** 证明: 对每个非零整数集合 $B=\{b_1,\cdots,b_n\}$, 都存在一个 sum-free 子集 A, 使得 $|A|>\frac{n}{3}$. 其中 **sum-free** 集 A 指的是: 不存在 $a_1,a_2,a_3\in A$, 使得 $a_1+a_2=a_3$.

证明. 令 p 为 $3k+2$ 型素数且满足 $p>2\max\limits_{1\leqslant i\leqslant n}|b_i|$, 设 $C=\{k+1,k+2,\cdots,2k+1\}$, 可知 C 是 $\mathbb{Z}_p$ 的 sum-free 子集且满足

$$\frac{|C|}{p-1}=\frac{k+1}{3k+1}>\frac{1}{3}.$$

现在 $\{1,2,\cdots,p-1\}$ 中随机均匀选取整数 x, 并定义 $d_i\equiv xb_i(\bmod p)$, $1\leqslant d_i\leqslant p-1$. 我们知道, 对固定的 b_i, 让 x 取遍 $1,2,\cdots,p-1$, 则 d_i 取遍 $\mathbb{Z}_p$ 中的非零元, 因此

$$\mathbb{P}(\{d_i\in C\})=\frac{|C|}{p-1}>\frac{1}{3}.$$

所以 $|B|$ 中满足 $\{d_i\in C\}$ 所对应的元素 b_i 的总数的期望值大于 $\frac{n}{3}$. 因此, 存在 $1\leqslant x\leqslant p-1$, 使得 B 中有一个子集 A 满足 $|A|>\frac{n}{3}$ 且

$$xa\equiv(\bmod p)\in C,\ \forall a\in A.$$

最后证明 A 是 sum-free 的. 否则, 若存在 $a_1,a_2,a_3\in A$, 使得 $a_1+a_2=a_3$, 则有

$$xa_1+xa_2\equiv xa_3(\bmod p),$$

这与 C 是 $\mathbb{Z}_p$ 的 sum-free 子集相矛盾. 这样我们就完成了证明. □

3 补充习题

1 自行求解第 2 次习题课讲义**例 2.4** 的条件期望 $\mathbb{E}[X|X+Y]$.

2 用示性函数的方法证明 **Waring 公式**(详见 **Grimmett 1.8.13**)

3 小朋友 A 有 N 块积木, N 服从参数为 λ 的泊松分布, 每块积木由小朋友 B 独立地以 $\frac{1}{2}$ 的概率拿走. 记小朋友 B 拿走的积木块数为 K, 求 $\mathbb{E}[K],\mathbb{E}[N|K]$.

解. 我们有

$$\mathbb{E}[K]=\mathbb{E}[\mathbb{E}[K|N]]=\mathbb{E}[\frac{1}{2}N]=\frac{1}{2}\mathbb{E}[N]=\frac{\lambda}{2}.$$

以上也可利用泊松分布在随机选择下的不变性 (见第 2 次习题课讲义) 得到 $K\sim P\left(\frac{\lambda}{2}\right)$. 而

$$\mathbb{P}(N=n)=\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda},\quad \mathbb{P}(K=k|N=n)=\frac{n!}{k!(n-k)!}\left(\frac{1}{2}\right)^n.$$

所以

$$\mathbb{P}(N=n|K=k)=\frac{\mathbb{P}(K=k|N=n)\mathbb{P}(N=n)}{\mathbb{P}(K=k)}=\frac{\frac{n!}{k!(n-k)!}\left(\frac{1}{2}\right)^n\cdot\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}}{\frac{(\frac{\lambda}{2})^k}{k!}\mathrm{e}^{-\lambda/2}}=\frac{(\frac{\lambda}{2})^{n-k}}{(n-k)!}\mathrm{e}^{-\lambda/2}.$$

因此

$$\mathbb{E}[N|K=k]=\sum_{n=k}^{\infty}n\mathbb{P}(N=n|K=k)=\sum_{n=k}^{\infty}(n-k+k)\frac{(\frac{\lambda}{2})^{n-k}}{(n-k)!}\mathrm{e}^{-\lambda/2}=k+\frac{\lambda}{2}.$$

故 $\mathbb{E}[N|K]=K+\dfrac{\lambda}{2}$. □

4 一个瓮包括了 b 个蓝球和 r 个红球.

(1) 在瓮中随机取球, 直到第一个蓝球被取出后停止. 证明取出的总球数的期望值是 $\dfrac{b+r+1}{b+1}$.

(2) 在瓮中随机取球, 直到某一种颜色的球都被取出后停止. 求瓮中取出的总球数的期望值.

解. (1) 记 X 表示取出的总球数, 在装有 b 个蓝球和 r 个红球的瓮下记 $M_{b,r}=\mathbb{E}[X]$, 并令 $Y=I_{\{\text{取出的第一个球是蓝球}\}}$, 则有

$$\begin{aligned}M_{b,r}=\mathbb{E}[X]=\mathbb{E}[\mathbb{E}[X|Y]]=\mathbb{E}[X|Y=0]\mathbb{P}(Y=0)+\mathbb{E}[X|Y=1]\mathbb{P}(Y=1)&=\frac{r}{b+r}(1+M_{b,r-1})+\frac{b}{b+r}\\&=1+\frac{r}{b+r}M_{b,r-1},\end{aligned}$$

利用 $M_{b,0}=1$, 可归纳得

$$M_{b,r}=\frac{b+r+1}{b+1}\text{ (细节请自行补充)}.$$

(2) 令 M 为过程停止时瓮中取出的总球数, B_r 为移走所有红球时剩下的蓝球数量, B_b 为移走所有蓝球时剩下的红球数量. 则

$$B_r+B_b+M=b+r.$$

假设我们把瓮中所有的球取出并按序排好, 记到第一个蓝球被取出后取出的红球数为 T_r, 由 (1) 知 $\mathbb{E}[T_r]=\dfrac{b+r+1}{b+1}-1=\dfrac{r}{b+1}$, 利用头尾对称性可知

$$\mathbb{E}[B_b]=\mathbb{E}[T_r]=\frac{r}{b+1}.$$

同理, $\mathbb{E}[B_r]=\dfrac{b}{r+1}$. 所以

$$\mathbb{E}[M]=b+r-\mathbb{E}[B_r]-\mathbb{E}[B_b]=\frac{br}{b+1}+\frac{br}{r+1}.$$

□

注. (1) 也可以用 $\mathbb{E}[X]=\sum_{n=0}^{\infty}\mathbb{P}(X>n)$ 来求解, 或者同 (2) 把所有的球取出并按序排好然后考虑插空.

5 证明:

$$Var(Y) = \mathbb{E}[Var(Y|X)] + Var(\mathbb{E}[Y|X])$$

注. 这里定义条件方差:

$$Var(Y|X) = \mathbb{E}[Y^2|X] - (\mathbb{E}[Y|X])^2.$$

解. 一方面, 有

$$\mathbb{E}[Var(Y|X)] = \mathbb{E}[\mathbb{E}[Y^2|X]] - \mathbb{E}[(\mathbb{E}[Y|X])^2] = \mathbb{E}[Y^2] - \mathbb{E}[(\mathbb{E}[Y|X])^2].$$

另一方面, 有

$$Var(\mathbb{E}[Y|X]) = \mathbb{E}[(\mathbb{E}[Y|X])^2] - (\mathbb{E}[\mathbb{E}[Y|X]])^2 = \mathbb{E}[(\mathbb{E}[Y|X])^2] - (\mathbb{E}[Y])^2.$$

上述两式相加即得证. □

6 Grimmett 3.11.15, 3.11.36

7 给定 $b > a > 0$, 离散随机变量 X 取值于区间 $[a, b]$.

(1) 证明: $Var(X) \leqslant \dfrac{(b-a)^2}{4}$.

(2) 求 $\mathbb{E}[X]\mathbb{E}[\frac{1}{X}]$ 的取值范围.

解. (1) 利用方差的 “投影” 性质及 $\left|X - \dfrac{a+b}{2}\right| \leqslant \dfrac{b-a}{2}$, 我们有

$$Var(X) = \min_{t\in\mathbb{R}} \mathbb{E}[(X-t)^2] \leqslant \mathbb{E}[(X - \frac{a+b}{2})^2] \leqslant \frac{(b-a)^2}{4}.$$

或者考虑中心化: $Y = X - \dfrac{b+a}{2}$, 则 $-\dfrac{b-a}{2} \leqslant Y \leqslant \dfrac{b-a}{2}$. 因此

$$Var(X) = Var(Y) \leqslant \mathbb{E}[Y^2] \leqslant \frac{(b-a)^2}{4}.$$

(2) 一方面, 由 Cauchy-Schwarz 不等式, 我们有

$$\mathbb{E}[X]\mathbb{E}[\frac{1}{X}] \geqslant (\mathbb{E}[\sqrt{X} \cdot \frac{1}{\sqrt{X}}]) = 1,$$

当 X 为常值时等号成立. 另一方面, 注意到 $Cov\left(X, \dfrac{1}{X}\right) = 1 - \mathbb{E}[X]\mathbb{E}[\dfrac{1}{X}]$, 而

$$\left|Cov\left(X, \frac{1}{X}\right)\right| \leqslant \sqrt{Var(X)Var\left(\frac{1}{X}\right)} \leqslant \sqrt{\frac{(b-a)^2}{4} \cdot \frac{(\frac{1}{a}-\frac{1}{b})^2}{4}} = \frac{(b-a)^2}{4ab},$$

因此

$$1 - \mathbb{E}[X]\mathbb{E}[\frac{1}{X}] \geqslant -\frac{(b-a)^2}{4ab} \implies \mathbb{E}[X]\mathbb{E}[\frac{1}{X}] \leqslant \frac{(a+b)^2}{4ab},$$

当 $\mathbb{P}(X=a) = \mathbb{P}(X=b) = \dfrac{1}{2}$ 时等号成立. □

致谢: 本习题课讲义的一部分内容参考了 2022 春学期概率论范惟和张国宇两位助教的习题课讲义, 感谢两位认真负责的前助教！