Homework 4

姓名: 方嘉聪 学号: 2200017849

Problem 1. 一个盒子中有 n 个小球, 编号分别为 $1, 2, \cdots, n$. 从盒子中取出 $k \le n$ 个小球, 每次等概率从盒子中剩余的小球中取出一个, 且每次取完后均不放回. 也即, 第一次取小球时, 每个小球被取出的概率均为 $1/n$. 第二次取小球时, 剩余的 $n-1$ 个小球各自被取出的概率均为 $1/(n-1)$. 以此类推, 直至一共取出 k 个小球.

定义随机变量 $X_1, X_2, \cdots, X_k$, 其中 $X_i (1 \le i \le k)$ 表示第 i 次取出小球的编号.

(1) 对于 $1 \le i < j \le k$, 判断 X_i 是否与 X_j 相互独立.

(2) 计算 $X_1, X_2, \cdots, X_k$ 的联合分布列.

(3) 对于 $1 \le i \le k$, 计算 X_i 的边际分布列.

(4) 对于任意 $1 \le i < j \le k$ 和 $1 \le a_i, a_j \le n$, 计算 $\mathbb{P}(X_i = a_i \cap X_j = a_j)$.

(5) 利用恒等式

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$

对于 $1 \le i \le k$, 计算 $\mathbb{E}(X_i)$ 和 $\mathrm{Var}(X_i)$.

(6) 对于 $1 \le i < j \le k$, 计算 $\mathrm{Cov}(X_i, X_j)$.

◀

Solution. (1) 注意到对于任意 $x_i, x_j \in \{1, \cdots, n\}$, 相当于第 i/j 位已经确定, 余下排列即可. 故有

$$\mathbb{P}(X_i = x_i) = \frac{(n-1)!}{n!} = \frac{1}{n}, \text{ 同理 } \quad \mathbb{P}(X_j = x_j) = \frac{1}{n}.$$

而当 $x_i = x_j$ 时, $\mathbb{P}(X_i = x_i, X_j = x_j) = 0 \neq \mathbb{P}(X_i = x_i) \cdot \mathbb{P}(X_j = x_j)$, 故 X_i 与 X_j 不相互独立.

(2) 对于 $\{x_i\}_{i=1}^k \in \{1, 2, \cdots, n\}$, 若 $\exists i \neq j$ 使得 $x_i = x_j$, 则 $\mathbb{P}(X_1 = x_1, X_2 = x_2, \cdots, X_k = x_k) = 0$. 否则, 有

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, \cdots, X_k = x_k) = \prod_{i=1}^{k} \frac{1}{n-i+1} = \frac{(n-k)!}{n!}.$$

(3) 对于 $1 \le i \le k$, 由 (1) 可知 $\mathbb{P}(X_i = x_i) = 1/n, \forall x_i \in \{1, 2, \cdots, n\}$. 否则, 有 $\mathbb{P}(X_i = x_i) = 0$.

(4) 若 $a_i = a_j$, 则 $\mathbb{P}(X_i = a_i \cap X_j = a_j) = 0$. 否则, 有

$$\mathbb{P}(X_i = a_i \cap X_j = a_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}.$$

综上,

$$\mathbb{P}(X_i = a_i \cap X_j = a_j) = \begin{cases} 0, & a_i = a_j, \\ \frac{1}{n(n-1)}, & a_i \neq a_j. \end{cases}$$

(5) 先计算 $\mathbb{E}(X_i)$, 有

$$\mathbb{E}(X_i) = \sum_{x_i=1}^{n} x_i \cdot \mathbb{P}(X_i = x_i) = \frac{1}{n} \sum_{x_i=1}^{n} x_i = \frac{n+1}{2}.$$

再计算 $\mathbb{E}(X_i^2)$, 有

$$\mathbb{E}(X_i^2) = \sum_{x_i=1}^{n} x_i^2 \cdot \mathbb{P}(X_i = x_i) = \frac{1}{n}\sum_{x_i=1}^{n} x_i^2 = \frac{(n+1)(2n+1)}{6}$$

那么

$$\mathrm{Var}(X_i) = \mathbb{E}(X_i^2) - \mathbb{E}(X_i)^2 = \frac{n^2-1}{12}.$$

(6) 我们来计算 $\mathbb{E}(X_iX_j)$, 其中 $1 \le i < j \le k$,

$$\mathbb{E}(X_iX_j) = \sum_{a_i=1}^{n}\sum_{a_j=1}^{n} a_ia_j\mathbb{P}(X_i = a_i, X_j = a_j) = \frac{1}{n(n-1)}\sum_{a_i \ne a_j} a_ia_j = \frac{(n+1)(3n+2)}{12}$$

那么

$$\mathrm{Cov}(X_i, X_j) = \mathbb{E}(X_iX_j) - \mathbb{E}(X_i)\mathbb{E}(X_j) = -\frac{n+1}{12}$$

◁

Problem 2. 将 n 个编号为 $1, 2, \cdots, n$ 的小球随机打乱, 每一种排列等概率出现. 用 $\pi_1, \cdots, \pi_n$ 表示随机打乱后每个位置上的小球编号, 也即 π_i 表示随机打乱后, 位置为 i 的小球的原始编号. 对于 $1 < i < n$, 我们称 i 是一个局部极大值, 当且仅当 $\pi_i > \pi_{i-1}$ 且 $\pi_i > \pi_{i+1}$. 令随机变量 X 表示所有 $1 < i < n$ 中局部极大值的总数量. 计算 $\mathbb{E}(X)$. ◀

Solution. 记随机变量 $X_i = \mathbb{1}_{i\text{为局部极大值}}$, 那么 $X = \sum_{i=2}^{n-1} X_i$. 而

$$\mathbb{E}(X_i) = \mathbb{P}(i\text{为局部极大值}) = \frac{\sum_{i=3}^{n}\binom{i-1}{2}\cdot 2!\cdot(n-3)!}{n!} = \frac{\sum_{i=1}^{n}(i-1)(i-2)}{n(n-1)(n-2)} = \frac{1}{3}.$$

那么

$$\mathbb{E}(X) = \sum_{i=2}^{n-1}\mathbb{E}(X_i) = \frac{n-2}{3}.$$

◁

Problem 3. 令随机变量 $X \sim G(p)$, 即随机变量 X 服从参数为 p 的几何分布. 证明

$$\begin{aligned}\mathbb{E}\left(X^2\right) &= p + \mathbb{E}\left((X+1)^2\right)(1-p),\\ \mathbb{E}\left(X^3\right) &= p + \mathbb{E}\left((X+1)^3\right)(1-p).\end{aligned}$$

并计算 $\mathbb{E}\left(X^2\right)$ 和 $\mathbb{E}\left(X^3\right)$. ◀

Solution. (1) 按照定义展开

$$\begin{aligned}\mathbb{E}\left(X^2\right) &= 1^2\cdot\mathbb{P}(X=1) + \sum_{k=2}^{+\infty} k^2\cdot\mathbb{P}(X=k) = p + \mathbb{P}(X>1)\sum_{k=2}^{+\infty} k^2\cdot\mathbb{P}(X=k|X>1)\\ &= 1 + (1-p)\sum_{k=2}^{+\infty} k^2\cdot\frac{p(1-p)^{k-1}}{1-p} = 1 + (1-p)\sum_{k=2}^{+\infty} k^2\cdot\mathbb{P}(X=k-1)\\ &= 1 + (1-p)\sum_{k=1}^{+\infty}(k+1)^2\cdot\mathbb{P}(X=k) = p + \mathbb{E}\left((X+1)^2\right)(1-p).\end{aligned}$$

对于 $\mathbb{E}\left(X^3\right)$, 同理可得 $\mathbb{E}\left(X^3\right) = p + \mathbb{E}\left((X+1)^3\right)(1-p)$.
那么 ($\mathbb{E}(X) = 1/p$)

$$\mathbb{E}\left(X^2\right) = p + (1-p)\cdot\left(\mathbb{E}\left(X^2\right) + 2\mathbb{E}(X) + 1\right) \implies \mathbb{E}\left(X^2\right) = \frac{2-p}{p^2}.$$

同理可得

$$\mathbb{E}\left(X^3\right) = p + (1-p)\left(\mathbb{E}(X^3) + 3\mathbb{E}(X^2) + 3\mathbb{E}(X) + 1\right) \implies \mathbb{E}\left(X^3\right) = \frac{p^2 - 6p + 6}{p^3}.$$

$\triangleleft$

Problem 4. 令 $X_1, X_2, \cdots$ 为一列同分布的离散随机变量. 离散随机变量 N 取正整数值, 且 $N, X_1, X_2, \cdots$, 相互独立. 在课上我们证明了

$$\mathbb{E}\left(\sum_{i=1}^{N} X_i\right) = \mathbb{E}(N)\cdot\mathbb{E}(X_1).$$

回答下列问题:

(1) 给出例子使得

$$\mathrm{Var}\left(\sum_{i=1}^{N} X_i\right) \neq \mathbb{E}(N)\cdot\mathrm{Var}(X_1).$$

(2) 证明

$$\mathrm{Var}\left(\sum_{i=1}^{N} X_i\right) = \mathbb{E}(N)\cdot\mathrm{Var}(X_1) + \mathrm{Var}(N)\cdot\mathbb{E}(X_1)^2.$$

$\blacktriangleleft$

Solution. (1) 任意 $i \in \mathbb{N}^+$, 令 X_i 同分布于单点分布, 即 $\mathbb{P}(X_i = c) = 1$, 其中 $c \in \mathbb{R}_{\neq 0}$ 为给定常数. 那么 $\mathrm{Var}(X_i) = 0$. 令 N 服从两点分布, 特别地令 $\mathbb{P}(N=1) = \mathbb{P}(N=2) = 1/2$. 那么

$$\mathbb{P}\left(\sum_{i=1}^{N} X_i = c\right) = \frac{1}{2}, \quad \mathbb{P}\left(\sum_{i=1}^{N} X_i = 2c\right) = \frac{1}{2}.$$

进而

$$\mathrm{Var}\left(\sum_{i=1}^{N} X_i\right) = \frac{1}{2}\cdot\frac{2c^2}{4} = \frac{c^2}{4} \neq 0 = \mathbb{E}(N)\cdot\mathrm{Var}(X_1).$$

(2) 由课上我们证明的结论, 可知

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{i=1}^{N} X_i\right)^2\right] &= \mathbb{E}\left[\sum_{i=1}^{N} X_i^2 + 2\sum_{1\le i<j\le N} X_iX_j\right] = \mathbb{E}\left[\sum_{i=1}^{N} X_i^2\right] + \mathbb{E}\left[\sum_{1\le i<j\le N} 2X_iX_j\right] \\ &= \mathbb{E}(N)\cdot\mathbb{E}(X_1^2) + \mathbb{E}(N^2 - N)\cdot\mathbb{E}(X_1)^2 = \mathbb{E}(N)\cdot\mathrm{Var}(X_1) + \mathbb{E}(N^2)\cdot E(X_1)^2\end{aligned}$$

那么

$$
\begin{aligned}
\operatorname{Var}\left(\sum_{i=1}^{N} X_i\right) &= \mathbb{E}\left[\left(\sum_{i=1}^{N} X_i\right)^2\right] - \mathbb{E}\left[\sum_{i=1}^{N} X_i\right]^2 \\
&= \mathbb{E}(N)\operatorname{Var}(X_1) + \mathbb{E}(N^2)E(X_1)^2 - \mathbb{E}(N)^2\mathbb{E}(X_1)^2 \\
&= \mathbb{E}(N)\cdot\operatorname{Var}(X_1) + \operatorname{Var}(N)\cdot\mathbb{E}(X_1)^2.
\end{aligned}
$$

证毕.

◁

Problem 5. (1) 对于正整数 r 和实数 $p \in (0,1)$, 给定 $X \sim NB(1,p), Y \sim NB(r,p)$. 若 X 和 Y 相互独立, 证明 $X+Y \sim NB(r+1,p)$. 提示: 使用恒等式

$$
\binom{n}{m} = \binom{n-1}{m-1} + \binom{n-1}{m}.
$$

(2) 对于正整数 r, 随机变量 $X_1, \cdots, X_r \overset{\text{i.i.d}}{\sim} G(p)$, 证明 $\sum_{i=1}^{r} X_i \sim NB(r,p)$.

◀

Solution. (1) 首先证明一个组合恒等式

$$
\sum_{k=1}^{n-1}\binom{n-k-1}{r-1} = \sum_{k=1}^{n-1}\binom{n-k}{r} - \binom{n-k-1}{r} = \binom{n-1}{r}.
$$

又注意到 $X, Y \in \mathbb{N}^+$, 那么对于任意 $n \in \mathbb{N}_{\geq 2}$:

$$
\begin{aligned}
\mathbb{P}(X+Y=n) &= \sum_{k=1}^{n} \mathbb{P}(X+Y=n|X=k)\mathbb{P}(X=k) \\
&= \sum_{k=1}^{n-1} \mathbb{P}(Y=n-k)\mathbb{P}(X=k) = \sum_{k=1}^{n-1} p(1-p)^{k-1}\cdot\binom{n-k-1}{r-1}p^r(1-p)^{n-k-r} \\
&= p^{r+1}(1-p)^{n-(r+1)}\sum_{k=1}^{n-1}\binom{n-k-1}{r-1} \\
&= \binom{n-1}{r}p^{r+1}(1-p)^{n-(r+1)}.
\end{aligned}
$$

故 $X+Y \sim NB(r+1,p)$, 证毕.

(2) 注意到分布 $G(p)$ 即为 $NB(1,p)$. 我们用数学归纳法, 当 $r=2$ 时由 (1) 中结论有

$$
X_1 + X_2 \sim NB(2,p)
$$

成立. 假设当 $r=k$ 其中 $k \in \mathbb{N}_{\geq 2}$ 时有 $\sum_{i=1}^{k} X_i \sim NB(k,p)$.
考虑 $r=k+1$ 那么由于 $\sum_{i=1}^{k} X_i \sim NB(k,p)$ 及 (1) 中结论有

$$
\sum_{i=1}^{k+1} X_i = \left(\sum_{i=1}^{k} X_i\right) + X_{k+1} \sim NB(k+1,p)
$$

故由数学归纳法, 对于正整数 r, 随机变量 $X_1, \cdots, X_r \overset{\text{i.i.d}}{\sim} G(p)$, $\sum_{i=1}^{r} X_i \sim NB(r,p)$. 证毕.

◁