1. 设随机变量 $X(\omega):\Omega\to\{1,2,3,\cdots,n,\cdots\}$，称 X 服从**几何分布**，若

$$P\{X=k\}=(1-p)^{k-1}p\ (k=1,2,\cdots,n,\cdots,\ 0<p<1).$$

同时记 $X\sim \mathrm{Ge}(p)$. 对这种随机变量的取值有可列个整数的情形，定义 $\mathrm{E}(X)=\sum_{k=1}^{\infty}x_kp_k=\lim_{n\to\infty}\sum_{k=1}^{n}x_kp_k$.

(1) 验证这是一个合法的分布列，即 $\sum_{k=1}^{\infty}P\{X=k\}=\lim_{n\to\infty}\sum_{k=1}^{n}P\{X=k\}=1$;

(2) 先待定 a,b 满足 $k(1-p)^{k-1}=[a(k+1)+b](1-p)^k-[ak+b](1-p)^{k-1}$，再证明 $\mathrm{E}(X)=1/p$;

(3) 用 $\mathrm{D}(X)=\mathrm{E}(X^2)-\mathrm{E}^2(X)$ 来证明 $\mathrm{D}(X)=(1-p)/p^2$;

(4) 几何分布有一个重要的性质：**无记忆性**。也就是对 $m,n\in\mathbb{N}^*$ 有

$$\boxed{P\{X=m+n\mid X>n\}=P\{X=m\}.\rightsquigarrow P\{X>m+n\mid X>n\}=P\{X>m\}.}$$

为了证明上述性质你需要一点准备工作，先验证几何级数 $\sum_{k=1}^{\infty}a_1q^{n-1}=\lim_{n\to\infty}\sum_{k=1}^{n}a_1q^{n-1}=a_1/(1-q)\ (|q|<1)$. 这一点在你求解 $P\{X>n\}=\sum_{k=n+1}^{\infty}P\{X=k\}$ 时会使用到。

(5) 求解几何分布的期望还可以使用下面的方法：

$$\begin{aligned}
\mathrm{E}(X)=\sum_{k=1}^{\infty}kP\{X=k\}&=1\cdot P\{X=1\}+\sum_{k=2}^{\infty}kP\{X=k\}\\
&=1\cdot P\{X=1\}+\sum_{k=2}^{\infty}kP\{X=k\mid X>1\}P\{X>1\}\\
&=1\cdot P\{X=1\}+P\{X>1\}\sum_{k=2}^{\infty}kP\{X=k\mid X>1\}\\
&=1\cdot P\{X=1\}+P\{X>1\}\sum_{k=2}^{\infty}(k-1+1)P\{X=k-1\}\\
&=1\cdot P\{X=1\}+P\{X>1\}\left[\sum_{k=2}^{\infty}(k-1)P\{X=k-1\}+\sum_{k=2}^{\infty}P(X=k-1)\right] \quad (1)\\
&=1\cdot P\{X=1\}+P\{X>1\}\left[\sum_{m=1}^{\infty}mP\{X=m\}+1\right] \quad (2)\\
&=1\cdot P\{X=1\}+P\{X>1\}(\mathrm{E}(X)+1)=1\cdot P\{X=1\}+(1-P\{X=1\})(\mathrm{E}(X)+1).
\end{aligned}$$

这种方法巧妙利用无记忆性得到一个关于 $\mathrm{E}(X)$ 的方程，化简问题，请你验证其中的每一步，并指出从式 (2) 到式 (1) 利用了**分布列的什么性质**？

2. 设 $X_1(\omega), X_2(\omega)$ 都是定义在样本空间 Ω 上的随机变量，且 $X_i : \Omega \to \{1, 2, \cdots, n\}$. 我们称

$$P\{X_1 = i, X_2 = j\} = p_{ij}\ (1 \le i, j \le n,\ p_{ij} \ge 0).$$

为二维随机变量 (X_1, X_2) 的**联合分布列**，自然有 $\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} p_{ij} = P(\Omega) = 1$.

(1) 验证

$$P\{X_1 = i\} = \sum_{j=1}^{n} P\{X_1 = i, X_2 = j\},\ P\{X_2 = j\} = \sum_{i=1}^{n} P\{X_1 = i, X_2 = j\}.$$

并利用这个结论证明 $E(X_1 + X_2) = E(X_1) + E(X_2)$ **恒成立**。

(2) 我们称 X_1, X_2 这两个**随机变量是独立的**，若

$$P\{X_1 = i, X_2 = j\} = P\{X_1 = i\} \cdot P\{X_2 = j\}\ (\forall\, 1 \le i, j \le n).$$

请验证在 X_1, X_2 独立的情况下有，$E(X_1X_2) = E(X_1)E(X_2)$.

(3) 利用 $D(X) = E[(X - E(X))^2]$ 验证在 X_1, X_2 独立的情况下有

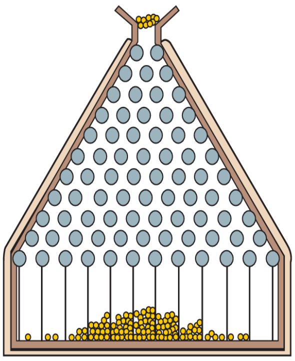
$$D(X_1 \pm X_2) = D(X_1) + D(X_2).$$

3. 一个袋子中有 10 个大小相同的球，其中有 4 个黄色球，6 个白球，从中随机地抽出 5 个球作为样本。用 X 表示样本中黄色球的个数。

   (1) 分别对有放回摸球和不放回摸球，求 X 的分布列；

   (2) 若用样本中黄色球的比例估计总体黄色球的比例，为了使得估计**尽可能准确**，应该采用哪种摸球方式？

4. 下图是一块高尔顿板的示意图。在一块木板上钉着与排列相互平行但相互错开的圆柱形小木钉，小木钉之间留有适当的空隙作为通道，前面挡有一块玻璃。将小球从顶端放入，小球下落的过程中，每次碰到小木钉后都会等概率地向左或向右落下，最后落入底部的格子中。格子从左到右分别编号为 0，1，2，…，10，用 X 表示小球最后落入格子的编号，求 X 的分布列。



5. 甲、乙两人选手进行象棋比赛，如果每局比赛甲获胜的概率为 0.6，乙获胜的概率为 0.4，那么采用 3 局 2 胜制还是采用 5 局 3 胜制对甲更有利？

6. 在概率统计中，我们常常通过观察到的实验结果应用**极大似然估计法**来估计参数的取值。设 X 为其分布列与未知参数 m 有关的离散型随机变量，其中 m 的取值范围为 S. 若对已知结果 $X=k$，有 $m_0 \in S$ 满足

$$\forall\, m_1 \in S,\ P\{X=k \mid m=m_0\} \geq P\{X=k \mid m=m_1\},$$

   则称 m_0 为 m 在 $X=k$ 下的一个极大似然估计。

   (1) 分别对下列情况求 m 在 $X=3$ 下的极大似然估计：

      i. $X \sim B(2, m)$;

      ii. $X \sim B(m, 1/2)$.

(2) 某台抽奖机上有一个按钮，参与者需要连续快速点击按钮来累积积分换取奖品。已知每次点击按钮后，获得 1 积分的概率为 p $(0<p<1)$，不获得积分的概率为 $1-p$. 小丽参加这个抽奖活动后总共获得了 k 积分，用极大似然估计的方法估计她点击按钮的总次数 m 的取值为 m_0，证明：$m_0 \le k/p$. 并指出等号成立的条件。

7. 绝大多数比赛都使用"$(2n-1)$ 局 n 胜制"的规则，但也有一些项目，比如冰壶运动，其整个比赛通常是进行偶数局。现在有甲、乙两名同学进行一项趣味项目的比赛，两人约定比赛规则为：共进行 $2n$ $(n\in\mathbb{N}^*)$ 局，谁赢的局数大于 n，谁就获得最终胜利。已知每局比赛中，甲的获胜概率均为 p $(0<p<1/2)$，乙获胜的概率均为 $1-p$，记甲赢得比赛的概率为 P_{2n}，讨论数列 $\{P_{2n}\}$ 的单调性。

8. 一个数轴上的质点在随机外力的作用下，从原点 O 出发，每隔 1 秒随机向左或向右移动一个单位，设向右移动的概率为 p $(0<p<1)$，移动 n 次后位于位置 X_n.

(1) $p=1/2$ 时，求 $P\{X_6=4\}$;

(2) 求 $E(X_n)$;

(3) 移动 n 次后质点最有可能位于哪个位置？即求使得 $P\{X_n=x\}$ 最大的 x.

9. 随机游走在空气中的烟雾扩散，股票市场的价格波动等动态随机现象中有重要应用。在平面直角坐标系中，粒子从原点出发，每秒等可能选择向左，向右，向上或向下移动一个单位。例如在 1 s 后，粒子可能到达的位置有 $(0,1)$, $(1,0)$, $(0,-1)$, $(-1,0)$.

(1) 设粒子在第 2 秒将移动到 (x,y)，记随机变量 $X=x+y$，求 $E(X)$；

(2) 记在第 n 秒粒子回到原点的概率为 p_n.

(i) 求 p_{2n}；

(ii) 已知：

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n < n! < \left(\frac{6}{\pi}\right)^{1/4}\sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

设 $b_n=p_{2n}$，记 S_n 为 $\{b_n\}$ 的前 n 项和。**称粒子为常返的**，若对任意实数 $M>0$，都存在 $n\in\mathbb{N}^*$ 使得 $S_n>M$，证明：该粒子是常返的。

10. 在空间直角坐标系中，**原点处**有一只蚂蚁可以沿 x, y, z 轴的正方向和负方向共 6 个方向**等可能**爬行.设 $2n$ 次爬行后，蚂蚁回到原点的概率为 P_{2n}. 求证：

$$P_{2n} > (C_{2n}^{n})^2/6^{2n}.$$

引例

11. 甲、乙、丙三人进行比赛，规定每局两个人比赛，胜者与第三人比赛，依次循环，直至有一人**连续两次胜利**为止，此人即为冠军。而每次比赛双方取胜的概率都是 $1/2$，现假定甲、乙两人先比，试求各人得冠军的概率。

期望递推

12. 甲、乙两人进行象棋比赛，每局甲胜的概率为 $p\ (0 < p < 1)$，乙胜的概率为 $q = 1 - p$。比赛进行到有一人**连续胜两局**为止，设比赛局数为随机变量 X，求 $E(X)$.

13. 甲、乙两口袋中各装有 1 个黑球和 2 个白球，现从甲、乙两口袋中各任取一个球交换放入另一口袋，重复进行 $n(n \in \mathbb{N}^*)$ 次这样的操作，记甲口袋中黑球的个数为 X_n，恰有 1 个黑球的概率为 p_n，恰有 2 个黑球的概率为 q_n，恰有 0 个黑球的概率为 r_n.

(1) 求 p_1, p_2 的值；

(2) 容易看出第 n 次甲口袋中黑球个数只受到 $n-1$ 次操作后口袋中黑球数量这一状态的影响，与先前的操作无关。记 $p_n = a \cdot p_{n-1} + b \cdot q_{n-1} + c \cdot r_{n-1}$，其中 $a, b, c \in [0, 1]$ 为常数，同时 $p_n + q_n + r_n = 1$，请求出 p_n；

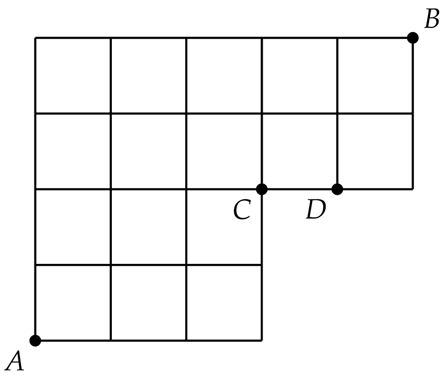
(3) 证明：X_n 的数学期望 $E(X_n)$ 为定值。

1. 从排成一排的 9 位同学中，随机选出 3 位同学，这 3 位同学互不相邻的概率是 ____.

2. 15 个人围坐在圆桌旁，从其中任选 4 人，两两不相邻的概率是 ____.

A. $\frac{30}{91}$ B. $\frac{25}{91}$ C. $\frac{10}{91}$ D. $\frac{31}{90}$

3. 在某城市中，A、B 两地之间有如图所示的道路网，甲沿道路随机关选一条最短路径，从 A 地出发去往 B 地，若甲途经 C 地，且不经过 D 地，则不同的路径共有 ____条。



4. 将一个圆环分成 n $(n \in \mathbb{N}, n \geq 3)$ 个区域，用 m $(m \in \mathbb{N}, m \geq 3)$ 种颜色给这个 n 个区域染色，要求相邻区域不使用同一种颜色，但同一种颜色可重复使用，则不同的染色方案有多少种？

5. 已知 $n \geq 2$，且平面内有 n 条直线，其中任意两条不平行，任意三条不共点，证明：这些直线的交点的个数为

$$f(n) = \frac{n(n-1)}{2}.$$

6. 求证：平面上 n 个圆最多把平面划分成 $n^2 - n + 2$ 个区域，其中 $n \geq 1$.

7. 在 $n \times n$ 网格上从点 0 走到点 n，要求每一步只能走 ↗ 或者 ↘，称为长为 $2n$ 的格路。不走到 x 轴下方的格路称为迪克路 (Dyck Path)。图 1 就展示了一个长度为 12 的迪克路。编号为 $1, 2, 3, \cdots, 2n$ 的点按顺序组成一个圈，用直线段将他们连接且线段不能出现交叉，这样连接的方法总数称为 Catalan 数，记作 C_n。图 2 展示了一个 $n = 6$ 的连接方法。

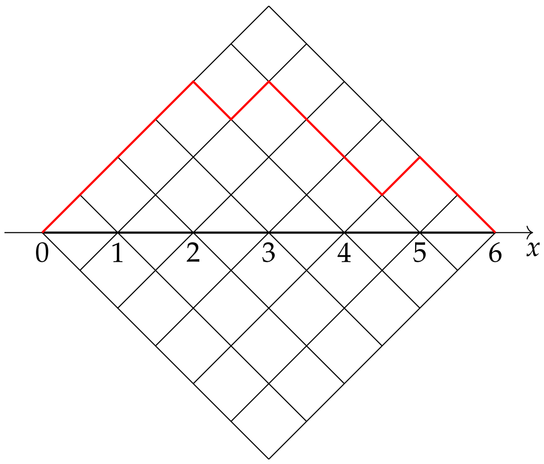


图 1: Dyck Path

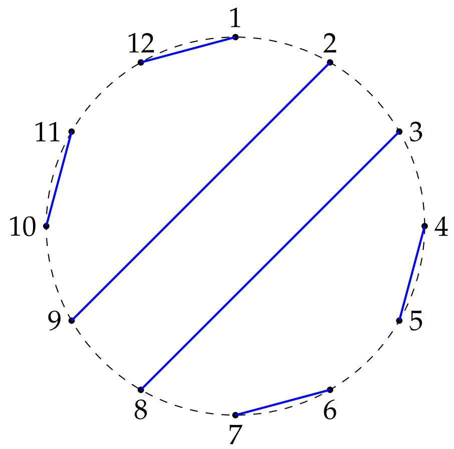
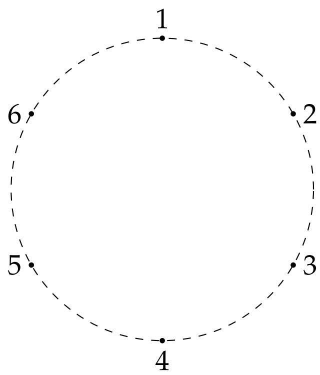
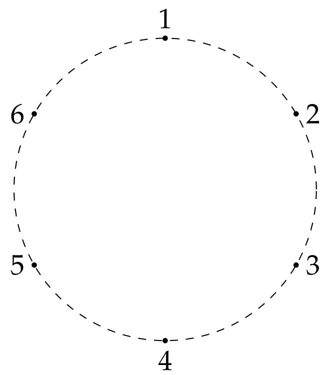


图 2: Catalan 配对

(1) 求 C_3 并在下面的图中绘制你找到的其中两种连接方法。





(2) 为了不出现交叉，其实相互配对的两个点的奇偶性必定相反，请你依据这一点，证明下列恒等式：

$$C_n = C_0 C_{n-1} + C_1 C_{n-2} + \cdots + C_{n-1} C_0 = \sum_{i=0}^{n-1} C_i C_{n-1-i}.$$

(3) 为了求解 C_n 的通项公式，用一一对应原理，请你尝试建立 Dyck Path 和 Catalan 配对之间的一一对应关系。事实上，图 1 和图 2 在这种对应关系下是等价的。

(4) 长度为 $2n$ 的 Dyck Path 的数目可以通过下列思路求解。利用“正难则反”的思想，对于每一个走到 x 轴下方的格路，都可以将其转化为**起点在 A 点，只在绿色阴影区域内走到点 n 的格路**。请你建立上述一一对应

关系并得到

$$C_n = \frac{1}{n+1} C_{2n}^n.$$

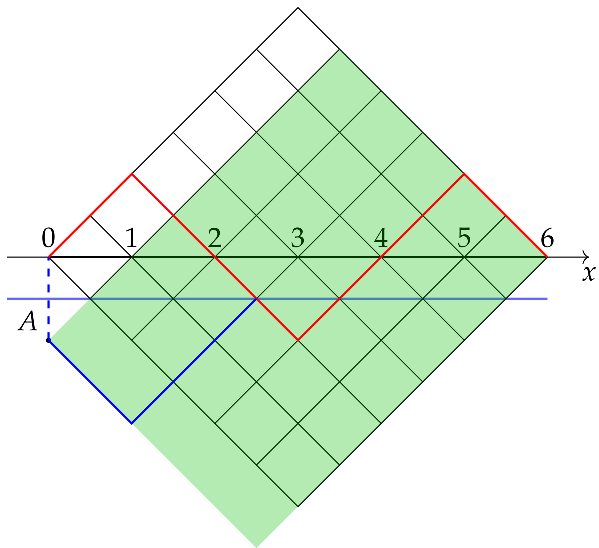


图 3: 一个非 Dyck Path 的示例