1. 算法说明

1.1 问题背景

“囚犯开箱问题”是一个著名的概率智力游戏。情境如下：有 N 名囚犯（编号从 0 到 N-1）和 N 个箱子（编号从 0 到 N-1）。每个箱子内随机放入一张写有唯一囚犯编号的纸条。囚犯们的目标是在不互相交流的情况下，所有囚犯都能找到写有自己编号的纸条。每位囚犯最多可以打开 N/2 个箱子。

1.2 策略详解

报告中主要探讨了两种策略：

1.2.1 循环策略 (Cyclical Strategy)

基本思想：利用排列的循环结构。每个囚犯从与自己编号相同的箱子开始查找，然后根据箱子里的编号继续查找下一个箱子，形成一条链。

执行步骤：

1.囚犯 P 首先打开编号为 P 的箱子。

2.如果箱子中纸条的编号是 P，则该囚犯成功找到。

3.如果箱子中纸条的编号是 X ($X \neq P$)，则囚犯 P 接下来打开编号为 X 的箱子。

4.重复步骤 2 和 3，直到找到自己的编号或者达到最大尝试次数（N/2）。

成功条件：所有囚犯都能在 N/2 次尝试内找到自己的编号。这等价于整个排列中，所有循环的长度都不超过 N/2。

1.2.2 随机策略 (Random Strategy)

基本思想：每个囚犯独立地随机选择箱子打开。

执行步骤：

1.每个囚犯随机选择 N/2 个箱子。

2.逐一打开这些箱子，查看纸条上的编号。

成功条件:所有囚犯都在他们随机选择的 N/2 个箱子中找到了自己的编号。

2. 实验结果

报告通过 Python 模拟了上述两种策略，并计算了它们的成功率。默认参数为：囚犯数 N=100，最大尝试次数 N/2=50。

2.1 成功率模拟

循环策略 (cyclical_strategy) 模拟结果：

在 10000 次模拟中，成功率为 0.3223。

使用向量化方法进行 100000 次大规模模拟时，成功率为 0.3119。

改变参数进行模拟：

囚犯数 N=50，最大尝试 25 次，成功率为 0.3226。

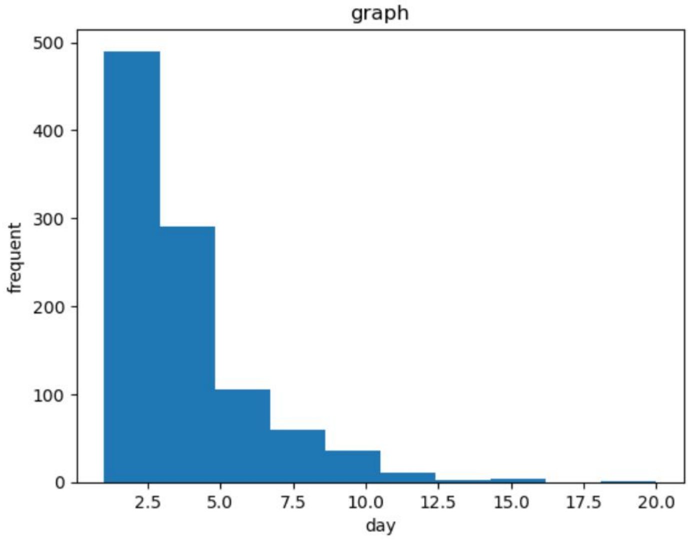
囚犯数 N=200，最大尝试 100 次，成功率为 0.3141。 这些结果与理论计算值（约 30.7%）非常接近，表明循环策略是一种有效的策略。

随机策略 (random_strategy) 模拟结果：

在 10000 次模拟中，成功率为 0.0。 随机策略的成功率极低，几乎为 0，远低于循环策略。这验证了随机选择并非解决此问题的有效方法。

2.2 循环策略成功所需时间分布

循环策略下，模拟 1000 次直到游戏成功所需天数（轮次）的直方图。该图展示了在不同天数下成功完成游戏的频率分布，有助于理解该策略的收敛特性。



3. 优化思路

为了应对大规模模拟的需求，报告中引入了向量化技术，显著提升了模拟效率。

3.1 向量化模拟 (box_game_vectorized)

核心思想：利用 NumPy 库的数组操作能力，将原本需要通过 Python 循环逐个模拟的逻辑，转化为针对整个模拟批次并行计算的矩阵运算。

实现方法：

箱子初始化：一次性生成 (num_simulations, num_prisoners) 大小的随机排列数组，代表所有模拟批次的箱子布局。

查找过程：通过 NumPy 的 np.take_along_axis 和 np.where 等函数，在每次尝试中同时更新所有模拟批次中所有囚犯的查找状态和下一个要打开的箱子，避免了显式的 for 循环。

成功判定：通过 np.all(found_self，axis=1) 快速判断每个模拟批次是否所有囚犯都已成功找到自己的编号。

优势：这种方法极大减少了 Python 解释器的开销，使得数万甚至数十万次模拟能在短时间内完成，非常适合大规模的统计分析。

4. 成功概率的数学推导

报告对循环策略的成功概率进行了严谨的数学推导，将其转化为一个纯粹的排列组合问题。

问题转化：囚犯集体获释的条件是，在随机生成的 N 个元素的排列中，不存在长度超过 N/2 的循环。因为如果存在一个长度大于 N/2 的循环，那么处于该循环中的囚犯将无法在 N/2 次尝试内找到自己的编号。

失败概率的近似：

在一个 N 元素的随机排列中，包含一个长度为 k 的循环的概率是 1/k。

因此，当最长循环长度 L＞N/2 时，失败的概率 P(失败) 可以近似表示为所有可能的长循环长度的概率之和： P(失败)＝P(L=N/2+1)＋P(L=N/2+2)＋... ＋P(L=N) P(失败) ≈ 求和(1/k)，k 从 N/2+1 到 N

对于 N=100： P(失败) ≈ 1/51 + 1/52 + ... + 1/100

调和级数与自然对数：这个求和是调和级数的一部分，可以近似地用自然对数表示： 求和(1/k) ≈ ln(N) - ln(N/2) = ln(N / (N/2)) = ln(2) 所以，对于 N=100，P(失败) ≈ ln(2) ≈ 0.693。

最终成功概率： P(成功) = 1 - P(失败) ≈ 1 - ln(2) ≈ 1 - 0.693 = 0.307 这表明，采用循环策略时，囚犯们集体获释的理论概率约为 30.7%。这一理论值与模拟结果高度一致，进一步证明了循环策略的有效性和其背后的数学原理。