24-25 应随实验班第二次期中 11.26

2024 年 11 月 29 日

一、X_t 是一个状态空间为 S 转移速率为 Q 的跳过程，假设其正常返，不变分布为 π. S_n 是一个与之独立的参数为 λ 的 Poisson 流，即 $S_n = Z_1 + \cdots + Z_n$ 其中 Z_i 是 i.i.d. 的参数为 λ 的指数随机变量. 求证：$Y_n := X_{S_n}$ 是马氏链并求其不变分布.

二、3-正则树上从根出发的跳过程，每条边的转移速率都是 1，求其在根结点总停留时长的分布.

三、考虑 $[-N, N]^2$ 网格，把除了 x 轴之外的所有横向的边删去（变成某种狼牙棒），求该图上从原点出发的简单随机游走首次抵达 $(N, 0)$ 的平均时间.

四、设在概率测度 $\mathbb{P}$ 下 X_t 是状态空间 Ω 上的一个马氏链，转移概率为 $p(x, y)$，不变分布为 π. 考虑定义在 Ω 的子集上的一个马氏链 S_t，转移规则如下：给定 $S_t = S \subset \Omega$，sample 一个独立的 U 服从 $[0, 1]$ 上的均匀分布，然后令

$$S_{t+1} = \left\{ y \in \Omega : \frac{\sum_{x \in S} \pi(x) p(x, y)}{\pi(y)} \geq U \right\}.$$

记上述马氏链的概率测度为 P. 记 t-step 转移矩阵 $P^t(x, y) := \mathbb{P}(X_t = y | X_0 = x)$. 求证：

$$P^t(x, y) = \frac{\pi(y)}{\pi(x)} \cdot P(y \in S_t | S_0 = \{x\}).$$

五、设 $G(N, \lambda/N)$ 是一个 N 个点的随机图，每条边独立的以 λ/N 的概率连接. 记 $\mathcal{C}$ 为该图中的最大连通分支，$|\mathcal{C}|$ 是其所含顶点数. 求证如下的相变：

(1)$\lambda < 1$ 是常数时，存在常数 $\alpha > 0$ 使得 $\lim\limits_{N \to \infty} \mathbb{P}(|\mathcal{C}| > \alpha \log N) = 0$.

(2)$\lambda > 1$ 是常数时，存在常数 $\alpha > 0$ 使得 $\lim\limits_{N \to \infty} \mathbb{P}(|\mathcal{C}| > \alpha N) = 1$.

六、考虑一个带“计划生育政策”的 Galton-Watson tree：子代分布是均值大于 1 的 Poisson 分布，但每层子代 sample 出来之后要随机删去其中 100 个（若只有不到 100 个那就视为灭绝）. 换句话说设 $\{Z^i_j\}_{i,j \geq 0}$ 是 i.i.d. 的 $Poisson(\lambda)$ 变量，$\lambda > 1$. 递归定义第 n 代个体个数 X_n：$X_0 = 1$，然后

$$X_{n+1} = \left(\sum_{k=1}^{X_n} Z^n_k - 100 \right)^+.$$

问：是否有正概率不灭绝（即 $\mathbb{P}(\forall n\ X_n \neq 0) > 0$）？

七、 常数 $\alpha \geq 0$，X_t 是 $\mathbb{Z}^3$ 上的跳过程，其嵌入链是简单随机游走，且在点 v 处的转移速率为 $|v|^\alpha$ 其中 $|v|$ 为点 v 到原点的欧氏距离. 对不同的 α 试讨论 X_t 是否爆炸.

八、 考虑如下的 majority vote 模型：T 是一个高度为 N 的三叉树，有 3^N 个叶结点. 每个叶结点处独立地放一个对称 Bernoulli 变量（即以 0.5 的概率取 +1 以 0.5 的概率取 -1）. 然后对于上面的每个结点它服从它的三个子结点中多数的那一方，我们关心根结点是 +1 还是 -1. 一个“开票策略”τ 是指我们按照某种“算法”依次查看一些叶结点的状态使得最终能确定根结点的状态（当然把 3^N 个叶结点全开了是一种办法，但非常浪费，因为假设你在某时刻知道一个结点的两个子结点都是 +1 那就不需要开第三个子结点了）. 记 $\mathbb{E}|\tau|$ 为策略 τ 查看的叶结点个数的期望.

(1) 求证：存在 τ 使得 $\mathbb{E}|\tau| \leq (2.5)^N$.

(2) 求证：存在 $0 < \alpha < 2.5$ 和一个 τ 使得对充分大的 N 有 $\mathbb{E}|\tau| \leq \alpha^N$.

一、显然是马氏链，下证 π 是不变分布：设初值 $Y_0 = X_0\pi$，则 $Y_1 = X_{Z_1}$ 的分布就是 X_t 的分布 μ_t 乘上指数分布在 t 处的密度函数 $f(t)$ 再对 t 积分. 然而 π 是 X 的不变分布所以每个 $\mu_t = \pi$，于是积分得到 $Y_1\pi$.

二、每次从根出发都有 1/2 的概率跑到无穷远不再回来，所以回到根的次数服从参数 1/2 的几何分布（可取 0 值的）. 每次待的时间是独立的均值 1/3 指数随机变量. 这样所求分布就是独立的 *Geometric*(1/2) 个（不取 0 值，因为要加上一开始在根停的那次，众所周知几何分布有两种定义（狗头））均值 1/3 指数随机变量之和.

三、设随机游走为 $Z_t = (X_t, Y_t)$. 取出来横坐标变了的时刻：$\tau_0 = 0, \tau_n = \inf\{t > \tau_{n-1} : X_t \neq X_{t-1}\}$. 那么 $W_n := X_{\tau_n}$ 是 $[-N, N]$ 上的随机游走. 记 $\sigma_N = \inf\{n : W_n = N\}$. 经过一些由一维简单随机游走的基本计算可得 $\mathbb{E}\sigma_N = 3N^2 - 2N$(笨人考场上此处算错了). 接下来看 τ_n 之间是什么样的，观察到 $Z_{\tau_{n-1}}$ 和 $Z_{\tau_n - 1}$ 之间实际上是这个随机游走在某条竖线上跑，也就由若干段“游弋”构成，每段“游弋”就是 $[0, N]$ 上从 0 出发直到再回到 0，游弋结束即回到 x 轴，那么有 1/2 的概率接着竖着走，所以 $\mathbb{E}(\tau_n - 1 - \tau_{n-1})$ 等于两倍的“游弋”长度期望. 同样又由一维简单随机游走的基本计算可得“游弋”长度的均值是 $2N$(笨人考场上此处又算错了). 组装在一起利用强马氏性和 Wald's identity 得所求为：

$$\mathbb{E}\tau_{\sigma_N} = \mathbb{E}\sigma_N \cdot \mathbb{E}(\tau_n - \tau_{n-1}) = (3N^2 - 2N) \cdot (4N + 1).$$

四、我们对 t 归纳. $t = 1$ 的情况我们验证定义：

$$P(y \in S_1 | S_0 = \{x\}) = P\left(\frac{\pi(x)p(x,y)}{\pi(y)} \geq U\right) = \frac{\pi(x)p(x,y)}{\pi(y)} = P^1(x,y) \cdot \frac{\pi(x)}{\pi(y)}.$$

归纳步骤 t 推 $t+1$ 先用归纳假设然后继续验证定义：

$$\begin{aligned}
P^{t+1}(x,y) &= \sum_{z\in\Omega} p(z,y) \cdot P^t(x,z) \\
&= \sum_{z\in\Omega} p(z,y) \cdot \frac{\pi(z)}{\pi(x)} \cdot P(z \in S_t | S_0 = \{x\}) \\
&= \sum_{z\in\Omega} \sum_{z\in S\subset\Omega} p(z,y) \cdot \frac{\pi(z)}{\pi(x)} \cdot P(S_t = S | S_0 = \{x\}) \\
&= \frac{1}{\pi(x)} \sum_{S\subset\Omega} P(S_t = S | S_0 = \{x\}) \cdot \sum_{z\in S} p(z,y)\pi(z) \\
&= \frac{1}{\pi(x)} \sum_{S\subset\Omega} P(S_t = S | S_0 = \{x\}) \cdot \pi(y) P(y \in S_{t+1} | S_t = S) \\
&= \frac{\pi(y)}{\pi(x)} \cdot P(y \in S_{t+1} | S_0 = \{x\}).
\end{aligned}$$

五、啊啊啊啊啊怎么考 Erdős–Rényi Graph 的 phase transition...（看到这题心态爆炸）. 鉴于敝人并没在考场做（默）出此题，请移步维基词条 Erdős–Rényi Model 中的 *Properties of* $G(n,p)$ 部分.

六、首先我们肯定让第一个人生超多孩子，然后观察到每层大概翻了 λ 倍，所以哪怕去掉 100 个也还应当是指数增长的. 取 $1 < \lambda_1 < \lambda_0 < \lambda$. 取充分大的正整数 N_0 使得 $\lambda_0 N_0 - 100 > \lambda_1 N_0$. 并且存在 $c > 0$ 使得对任意 $N > N_0$，若 $X_1, \cdots, X_N$ 是独立的 *Poisson*(λ) 变量，那么 $\mathbb{P}(X_1 + \cdots + X_N < \lambda_0 N) < \exp(-cN)$. 这里可以直接用 *Poisson*($N\lambda$) 的具体估计（见 https://www.math.uci.edu/ rvershyn/papers/HDP-book/HDP-book.pdf 的 Exercise2.3.5）或者直接引用大偏差理论得到指数 tail（见 Durrett 2.7 节）.

有了这些工具我们就可以完成这道题. 首先以正概率 p 第一个人有多于 $N_0 + 100$ 个子代. 然后条件在第 n 代有至少 $\lambda_1^n N_0$ 个人存活，下一代以至少 $1 - \exp(-c\lambda_1^n N_0)$ 的概率有不小于 $\lambda_0 \lambda_1^n N_0 - 100 > \lambda_1^{n+1} N_0$ 个人存活. 每层之间是独立的，相乘得到至少以如下概率不灭绝：

$$p \cdot \prod_{n=1}^{\infty} (1 - \exp(-c\lambda_1^n N_0)) > 0.$$

七、我直接忽略了原点是吸收态这件事... 所以得到 $\alpha > 2$ 以概率 1 爆炸，反之以概率 1 不爆炸.

Condition on 随机游走的一条轨迹 X_n 上，跳过程跳的时间就是独立的 θ_n 之和，θ_n 服从均值为 $|X_n|^{-\alpha}$ 的指数分布. 概率论的计算可以得到这些指数分布变量之和以概率 1 有限（或发散）当且仅当和的均值 $\sum_n |X_n|^{-\alpha}$ 收敛（或发散）. 所以我们只需要讨论 $\sum_n |X_n|^{-\alpha}$ 是否收敛. 计算期望：X_n 的分布约等于 $\mathcal{N}(0,n)$，引用一些集中不等式的结果我们得到 $|X_n| - \sqrt{n}$ 是 sub-gaussian 的，所以 $\mathbb{E}|X_n|^{-\alpha} \approx n^{-\alpha/2}$. 因此对于 $\alpha > 2$ 有：

$$\mathbb{E}\sum_{n=1}^{\infty} |X_n|^{-\alpha} \approx \sum_{n=1}^{\infty} n^{-\alpha/2} < \infty.$$

因此 a.s. 有 $\sum_n |X_n|^{-\alpha} < \infty$，于是 a.s. 跳过程爆炸. 而对于 $\alpha \leq 2$，因为 $|X_n| - \sqrt{n}$ 是 sub-gaussian 的，存在 $c > 0$ 使得

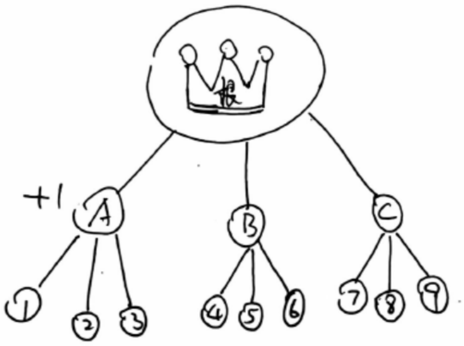
$$\mathbb{P}(|X_n| > 2\sqrt{n}) < exp(-cn^2).$$

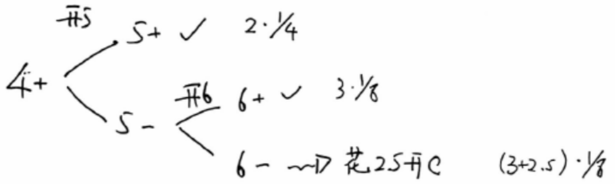
是可和的，于是由 Borel-Cantelli 引理知 $\mathbb{P}(|X_n| > 2\sqrt{n}\ \text{infinitelyoften}) = 0$. 所以 a.s. 有: 对充分大的 n，$|X_n| \leq 2\sqrt{n}$，此时 $\sum_n |X_n|^{-\alpha} = \infty$. 从而跳过程 a.s. 不爆炸. 对于原点是吸收态的时候上述估计仍然成立，只需要讨论是否击中零点即可，$\alpha > 0$ 时只要击中原点就不爆炸.

八、(1) 观察到如果是 $N = 1$ 的情况，三个叶子可以先开两个，它们不一样才需要开第三个. 所以平均来说可以少开半个，期望等于 5/2. 对更高的树我们归纳给出策略：假设可以用平均 $(5/2)^N$ 次开一个 N 层树的根结点，对于 $N+1$ 层树，我们看它第一层的三个结点分别是三个独立 N 层子树的根. 用平均 $2 * (5/2)^N$ 次开其中两个，然后有一半的概率需要开第三个，这样得到 $(2.5)^{N+1}$.

(2) 根据 (1) 中的归纳步骤，我们只需给出一个开两层树的策略平均次数比 $(2.5)^2 = 6.25$ 就可以. 我们用决策树给出如下策略：

先用(1)中办法把A开了，花费2.5
不妨设开出来A为+1，去开4





这里"花1.5开B"因为4已知，5和6只需开1.5个.

平均花费：$2.5+\frac{2}{4}+\frac{1}{8}+\frac{5.5}{8}$

$+\frac{3}{8}+\frac{4}{16}+\frac{5.5}{16}+\frac{4}{16}+\frac{5.5}{16}+\frac{4.5}{8}$

$=\frac{95}{16}<6<6.25$