马尔可夫链

请在 12 月 26 日课前提交纸质作业.

1. (6 分) 考虑 $\mathbb{Z}$ 上的随机游走. 马尔可夫核是

$$P(x+1|x) = p, \quad P(x-1|x) = 1-p$$

其中 $p \in (0,1)$ 是参数. 请计算这个马尔可夫链返回初始点的概率.

$$\Pr[\exists i > 0 \text{ such that } X_i = X_0].$$

2. (6 分) 考虑 $\mathbb{Z}^d$ 上的随机游走. 马尔可夫核是

$$P(y_1, \ldots, y_d | x_1, \ldots, x_d) = \begin{cases} 1/3^d, & \text{if } \forall i, |y_i - x_i| \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

这个马尔可夫核在各个维度上独立，便于分析. 证明

$$\Pr[\exists i > 0 \text{ such that } X_i = X_0] = \begin{cases} 1, & \text{if } d = 2 \\ 1 - \Omega(1), & \text{if } d > 2 \end{cases}$$

提示：考虑

$$\mathbb{E}[\text{number of } i > 0 \text{ such that } X_i = X_0].$$

3. (6 分) 有 n 种不同卡片. 可以从一个抽卡机中，每次独立地获得一张随机卡片.
   (1) 期望需要抽多少次卡，才能收集到每种卡片至少一张.
   (2) 需要抽多少次卡，才能以至少 $1 - 1/n$ 的概率收集到每种卡片至少一张. （给出一个尽量紧的上界即可. 可以有常数倍的放松.）

4. (10 分) 简单图 G 中有 n 个点, 最大度数记为 Δ. 用 $C > 5\Delta$ 种颜色对 G 随机点染色, 要求任意一对相邻点的染色不同. 为了均匀采样一个随机染色, 我们使用 MCMC 方法. 马尔可夫核是:
   - 假设当前染色为 $f : V \to C$.
   - 随机选取一个点 $v \in C$, 随机选取一个颜色 $c \in C$. (TODO 明年改成随机选取一个邻居没有的颜色)
   - 如果 v 的邻居的颜色都不是 c, 就将 v 的染色修改为 c; 否则保持染色不变.

   请估算混合时间 $\tau(\varepsilon)$, 给出一个尽量好的上届.

$$\tau(\varepsilon) = \text{smallest } t \text{ s.t. } d(t) \leq \varepsilon$$
$$d(t) = \max_x \Delta_{\text{TV}}(P^t(x, \cdot), \pi)$$

注. 如果想用 *coupling* 分析 $C > 2\Delta$ 的情形, 建议用 $S_t \subseteq V$ 表示 t 时刻 *coupling* 中两个染色一致的点集. 考虑被 S_t 切的边 (一个端点在 S_t 中, 另一个端点在 S_t 外) 有怎样的影响.

5. (10 分) 随机图 $G(n,p)$ 是连通的概率是多少?
   (1) 证明存在常数 $\alpha > 0$，使得对所有充分大的 n，$G(n, \alpha \ln n/n)$ 是连通图的概率不高于 $1/n$.
   (2) 证明存在常数 $\alpha > 0$，使得对所有充分大的 n，$G(n, \alpha \ln n/n)$ 是连通图的概率不低于 $1 - 1/n$.
   提示：$G(n,p)$ 可以如此采样：先从二项分布 $\mathrm{Binom}(\binom{n}{2}, p)$ 中采样 m，再从 $G(n,m)$ 中采样. $G(n,m)$ 可以如此采样：从没有边的图开始，每次随机添加一条新边，重复 m 次.