Homework #3

Due: 2024-6-18 23:59 | 5 Questions, 100 Pts

Name: XXX, ID: XXX

Question 1 (15') (Coupon Collector).

假设有 n 种抽奖券。每次抽取的过程中抽中任一种奖券的概率均相同。

a (5') 假设抽取了 X 次后第一次抽到第一种抽奖券，求关于 X 的期望 $\mathbb{E}[X]$。

b (5') 假设抽取了 Y 次后集齐了全部的奖券，求关于 Y 的期望 $\mathbb{E}[Y]$。

c (5') 为了以高于 $1-\epsilon$ 的概率集齐全部的奖券，求证最少进行抽取的次数

$$m = \mathcal{O}\left(n \log \frac{n}{\epsilon}\right).$$

◄

Question 2 (30') (独立性).

在概率论的实践中，我们强调变量间的独立性。对于随机变量 X, Y，我们称他们统计上独立，当且仅当他们的联合概率等于它们概率的乘积，即

$$\Pr[X \cap Y] = \Pr[X] \Pr[B].$$

a (5') 请证明对于统计上独立的随机变量 X, Y，对于期望算子 $\mathbb{E}$ 满足

$$\mathbb{E}[X^n Y^m] = \mathbb{E}[X^n]\mathbb{E}[Y^m].$$

两个变量 X, Y 间的协方差 cov 被定义为

$$\operatorname{cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

b (5') 从 $[-1, 1]$ 上随机均匀采样，将采样结果作为随机变量 X，定义依赖于 X 的随机变量 $Y = X^2$。请计算 X 与 Y 的协方差。

c (5') 设随机变量 X, Y 的联合分布密度为

$$f(x, y) = \begin{cases} \frac{1}{4}(1 + xy), & |x| < 1, |y| < 1, \\ 0, & \text{else}. \end{cases}$$

请计算判断 X 与 Y 是否独立，X^2 与 Y^2 是否独立。

我们指出，对于随机变量 X, Y，若对于任意给定的常数 a, b，两者的线性组合 $aX + bY$ 均可表示为单变量正态分布，那么如果 X, Y 不相关，那么它们是独立的。但是实践中，有些人会认为两个线性不相关、正态分布的随机变量一定是统计独立的；有些人会认为正态分布关于随机变量的线性组合是正态分布的。以下两题为两个反例。

d (5') 取 X 为一个期望为 0、方差为 1 的满足正态分布的变量。W 独立于 X，以相同的概率取 1 或者 -1。计算随机变量 $Y = WX$ 的协方差，并给出例子来说明 X, Y 不独立。

e (5') 取 X 为一个期望为 0、方差为 1 的满足正态分布的变量。取

$$Y = \begin{cases} X, & \text{if } |X| \le c, \\ -X, & \text{if } |X| > c, \end{cases}$$

其中 c 为某一常数。观察可知若 $c \to 0$，那么 $\mathrm{cov}[X, Y] \to -1$；若 $c \to \infty$，那么 $\mathrm{cov}[X, Y] \to 1$。由于相关性关于 c 连续，那么必然存在一个值 c 使得 $\mathrm{cov}[X, Y] = 0$。请证明 Y 为一个正态分布。

对于一列相同概率空间上的随机变量 $\{X_i\}_{i=1}^n$，我们称每 k 个元素独立为对于每一个大小不大于 k 的子集 $I \subseteq \{1, \ldots, n\}$，$|I| \le k$，对于任一取值列 $\{a_i\}$，满足

$$\Pr\left[\bigwedge_{i \in I} X_i = a_i\right] = \prod_{i \in I} \Pr[X_i = a_i].$$

若 $k = n$，则这些 $\{X_i\}$ 相互独立。

f (5') 设三维随机向量 (X, Y, Z) 的联合密度函数为

$$f(x, y, z) = \begin{cases} \dfrac{1}{8\pi^3}(1 - \sin x \sin y \sin z), & 0 < x, y, z < 2\pi, \\ 0, & \text{else}. \end{cases}$$

请求出 X, Y, Z 各自的边际分布，并判断是否两两独立？是否相互独立？

◀

Question 3 (10') (大数定律).
尽管很多情形下大数定理可以得到满足，但是我们也要注意不满足的情形。

a (5') 令 $\{X_n, n \ge 2\}$ 为一列独立的随机变量序列，满足

$$\Pr[X_n = \pm n] = \frac{1}{2n \log n}, \Pr[X_n = 0] = 1 - \frac{1}{n \log n}, n = 2, 3, \ldots$$

请证明 $\{X_n, n \ge 2\}$ 满足弱大数律，不满足强大数律。

b (5') 令 $\{X_n, n \ge 2\}$ 为一列独立的随机变量序列，密度函数为

$$f_n(x) = \frac{1}{\sqrt{2}\sigma_n} \exp\left\{-\frac{\sqrt{2}|x|}{\sigma_n}\right\}, \quad x \in \mathbb{R},$$

其中 $\sigma_n^2 = 2n^2/(\log n)^2, n \ge 2$。请证明 $\{X_n, n \ge 2\}$ 满足强大数律，不满足弱大数律。

◀

Question 4 (15') (鞅).
随机过程中一个重要的基础概念是鞅（Martingale）。此处仅作简单介绍，图形学中的应用将在微分方程一节的习题中展示。
取 $(Z_i)_{i=1}^n$ 与 $(X_i)_{i=1}^n$ 为共同的概率空间上的一列随机变量，若是对于所有的 i，$\mathbb{E}[X_i \mid Z_1 \ldots Z_{i=1}] = X_{i-1}$，那么 (X_i) 被称为关于 (Z_i) 的鞅。进一步地，$Y_i = X_i - X_{i-1}$ 被称为鞅差序列，满足对于所有的 i 而言，$\mathbb{E}[Y_i \mid Z_1 \ldots Z_{i=1}] = 0$。
鞅无处不在。事实上，对于任意的随机变量我们都可以得到一个鞅。

a (5') 令 A 与 (Z_i) 为共同概率空间上的随机变量。请证明

$$X_i = \mathbb{E}[A \mid Z_1 \dots Z_i]$$

是一个鞅。

以上定义对应的鞅被称为关于 A 的 Doob martingale。
选取 $\mathcal{F}_i = \{Z_1, \dots, Z_i\}$，我们称一个随机变量 $T \in \{0, 1, 2, \dots\} \cup \{\infty\}$ 为一个停时，则事件 $\{T = i\}$ 关于 $\mathcal{F}_i$ 可测。即，已知 $\mathcal{F}_i$ 以后，$\{T = i\}$ 是否成立便可知晓，而不依赖此后的历史。例如第 1 次抛硬币朝上为一个停时，而第一次硬币朝下前的最后一次朝上便不构成一个停时。
若是停时满足 $\mathbb{E}[T] < \infty$ 且对于所有的 i 与某一指定常数 c 满足 $\mathbb{E}[|X_i - X_{i-1} \mid \mathcal{F}_i|] \leq c$，那么可以得到 $\mathbb{E}[X_T] = \mathbb{E}[X_0]$。

b (5') 一个赌徒一开始身无分文，他每局以相同的概率赢得一块钱或者输掉一块钱。如果他输了 a 块钱或者赢了 b 块钱便离开，其中 a, b 均为整数。求问他赢得 b 块钱的概率。

c (5') 求问他需要多少时间才会离开。请使用 $Y_i = X_i^2 - i$ 做变量代换。

◀

Question 5 (30') (代码填空).
请完成代码包中给出的任务。

◀