Homework #6

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

Name: JS

Question 1 (15') (Coupon Collector).

a) 每次抽中第一种抽奖券的概率为 $\frac{1}{n}$。因此，X 服从参数为 $\frac{1}{n}$ 的几何分布。
几何分布的期望为 $\frac{1}{p}$，其中 p 是成功的概率。因此，

$$\mathbb{E}[X] = \frac{1}{\frac{1}{n}} = n.$$

b) 集齐所有奖券需要依次抽到第1种、第2种、⋯⋯、第n种奖券。设第i种奖券在前 $i-1$ 种奖券都已经集齐的情况下抽到的次数为 Y_i。
第i种奖券的抽中概率为 $\frac{n-(i-1)}{n} = \frac{n-i+1}{n}$，因此抽到第$i$种奖券的期望次数为 $\frac{n}{n-i+1}$。
总次数 Y 为

$$Y = Y_1 + Y_2 + \cdots + Y_n.$$

于是，

$$\mathbb{E}[Y] = \mathbb{E}[Y_1] + \mathbb{E}[Y_2] + \cdots + \mathbb{E}[Y_n] = n\left(\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n}\right) = nH_n,$$

其中 H_n 为第 n 项调和级数 $H_n = \sum_{k=1}^{n} \frac{1}{k}$。
利用调和级数的近似公式 $H_n \approx \ln n + \gamma$ （其中 γ 是欧拉-马歇罗尼常数），可以得到

$$\mathbb{E}[Y] \approx n(\ln n + \gamma).$$

c) 为了以至少 $1-\epsilon$ 的概率集齐全部的奖券，我们需要估计 Y 的上界。
利用 Chernoff 界，我们有

$$\Pr(Y > (1+\delta)\mathbb{E}[Y]) < \exp\left(-\frac{\delta^2 \mathbb{E}[Y]}{2+\delta}\right).$$

令 $1+\delta = k$，则

$$\Pr(Y > k\mathbb{E}[Y]) < \exp\left(-\frac{(k-1)^2 \mathbb{E}[Y]}{2k}\right).$$

为了满足 $\Pr(Y \le k\mathbb{E}[Y]) > 1-\epsilon$，令

$$\exp\left(-\frac{(k-1)^2 \mathbb{E}[Y]}{2k}\right) < \epsilon.$$

取对数得到

$$-\frac{(k-1)^2 \mathbb{E}[Y]}{2k} \le \ln \epsilon,$$

即

$$(k-1)^2 \mathbb{E}[Y] \ge -2k \ln \epsilon.$$

由于 $\mathbb{E}[Y] \approx n \ln n$，

$$(k-1)^2 n \ln n \ge -2k \ln \epsilon.$$

取 $k \approx \frac{2 \ln n}{\ln \epsilon}$，得到

$$m = k\mathbb{E}[Y] \approx \mathcal{O}\Big(n \log \frac{n}{\epsilon}\Big).$$

因此，最少进行 $\mathcal{O}\big(n \log \frac{n}{\epsilon}\big)$ 次抽取以高于 $1-\epsilon$ 的概率集齐全部的奖券。 ◀

Question 2 (30') (独立性).

a) 证明：
首先，回顾独立随机变量的定义。两个随机变量 X 和 Y 独立，意味着对于任意的事件 A 和 B，有

$$\Pr[X \in A, Y \in B] = \Pr[X \in A] \Pr[Y \in B].$$

对于独立随机变量 X 和 Y，其联合概率密度函数可以分解为各自概率密度函数的乘积，即

$$f_{X,Y}(x, y) = f_X(x) f_Y(y).$$

现在考虑 $\mathbb{E}[X^n Y^m]$：

$$\mathbb{E}[X^n Y^m] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^n y^m f_{X,Y}(x, y)\, dx\, dy.$$

由于 X 和 Y 独立，联合概率密度函数 $f_{X,Y}(x, y)$ 可以分解为 $f_X(x) f_Y(y)$，所以

$$\mathbb{E}[X^n Y^m] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^n y^m f_X(x) f_Y(y)\, dx\, dy.$$

将积分拆分，我们得到

$$\mathbb{E}[X^n Y^m] = \left(\int_{-\infty}^{\infty} x^n f_X(x)\, dx \right) \left(\int_{-\infty}^{\infty} y^m f_Y(y)\, dy \right).$$

这里，第一部分 $\int_{-\infty}^{\infty} x^n f_X(x)\, dx$ 是 $\mathbb{E}[X^n]$，第二部分 $\int_{-\infty}^{\infty} y^m f_Y(y)\, dy$ 是 $\mathbb{E}[Y^m]$。因此，

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

这就证明了对于统计上独立的随机变量 X 和 Y，有

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

这表明独立随机变量的协方差为零。

b) 要计算随机变量 X 和 $Y = X^2$ 的协方差，我们首先回顾协方差的定义：

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

我们分别计算 $\mathbb{E}[XY]$、$\mathbb{E}[X]$ 和 $\mathbb{E}[Y]$。
随机变量 X 在区间 $[-1, 1]$ 上均匀分布，因此其概率密度函数 $f_X(x)$ 为：

$$f_X(x) = \frac{1}{2}, \quad -1 \le x \le 1$$

$$\mathbb{E}[X] = \int_{-1}^{1} x f_X(x)\, dx = \int_{-1}^{1} x \cdot \frac{1}{2}\, dx$$

$$\mathbb{E}[X] = \frac{1}{2} \int_{-1}^{1} x\, dx$$

由于 x 在 $[-1,1]$ 上关于原点对称，且是奇函数，因此其积分为 0：

$$\int_{-1}^{1} x\,dx = 0$$

$$\mathbb{E}[X] = \frac{1}{2}\cdot 0 = 0$$

因为$\mathbb{E}[Y] = \mathbb{E}[X^2]$

$$\mathbb{E}[X^2] = \int_{-1}^{1} x^2 f_X(x)\,dx = \int_{-1}^{1} x^2\cdot\frac{1}{2}\,dx$$

$$\mathbb{E}[X^2] = \frac{1}{2}\int_{-1}^{1} x^2\,dx$$

我们计算 $\int_{-1}^{1} x^2\,dx$：

$$\int_{-1}^{1} x^2\,dx = \left[\frac{x^3}{3}\right]_{-1}^{1} = \frac{1}{3} - \left(-\frac{1}{3}\right) = \frac{1}{3}+\frac{1}{3} = \frac{2}{3}$$

$$\mathbb{E}[X^2] = \frac{1}{2}\cdot\frac{2}{3} = \frac{1}{3}$$

由于 $Y = X^2$，我们有 $XY = X\cdot X^2 = X^3$：

$$\mathbb{E}[XY] = \mathbb{E}[X^3] = \int_{-1}^{1} x^3 f_X(x)\,dx = \int_{-1}^{1} x^3\cdot\frac{1}{2}\,dx$$

$$\mathbb{E}[X^3] = \frac{1}{2}\int_{-1}^{1} x^3\,dx$$

由于 x^3 在 $[-1,1]$ 上关于原点对称，且是奇函数，因此其积分为 0：

$$\int_{-1}^{1} x^3\,dx = 0$$

$$\mathbb{E}[X^3] = \frac{1}{2}\cdot 0 = 0$$

将以上结果代入协方差公式：

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

$$\mathrm{cov}[X,Y] = 0 - 0\cdot\frac{1}{3} = 0$$

因此，随机变量 X 和 $Y = X^2$ 的协方差为：

$$\mathrm{cov}[X,Y] = 0$$

c) 要判断随机变量 X 和 Y 是否独立，我们需要检查它们的联合密度函数是否可以表示为它们各自的边缘密度函数的乘积。给定联合分布密度函数：

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

首先计算 X 和 Y 的边缘密度函数。

$$f_X(x) = \int_{-1}^{1} f(x, y)\, dy = \int_{-1}^{1} \frac{1}{4}(1 + xy)\, dy.$$

我们计算这个积分：

$$f_X(x) = \frac{1}{4}\int_{-1}^{1}(1 + xy)\, dy = \frac{1}{4}\left[\int_{-1}^{1} 1\, dy + x\int_{-1}^{1} y\, dy\right].$$

注意到 $\int_{-1}^{1} 1\, dy = 2$ 和 $\int_{-1}^{1} y\, dy = 0$：

$$f_X(x) = \frac{1}{4}\,(2 + x \cdot 0) = \frac{1}{2}.$$

所以：

$$f_X(x) = \begin{cases} \frac{1}{2}, & |x| < 1, \\ 0, & \text{otherwise.} \end{cases}$$

由于对称性，$f_Y(y)$ 与 $f_X(x)$ 的计算过程类似：

$$f_Y(y) = \int_{-1}^{1} f(x, y)\, dx = \int_{-1}^{1} \frac{1}{4}(1 + xy)\, dx.$$

同理，

$$f_Y(y) = \frac{1}{4}\int_{-1}^{1}(1 + xy)\, dx = \frac{1}{4}\left[\int_{-1}^{1} 1\, dx + y\int_{-1}^{1} x\, dx\right].$$

注意到 $\int_{-1}^{1} 1\, dx = 2$ 和 $\int_{-1}^{1} x\, dx = 0$：

$$f_Y(y) = \frac{1}{4}\,(2 + y \cdot 0) = \frac{1}{2}.$$

所以：

$$f_Y(y) = \begin{cases} \frac{1}{2}, & |y| < 1, \\ 0, & \text{otherwise.} \end{cases}$$

检查 $f(x, y)$ 是否等于 $f_X(x)f_Y(y)$：

$$f_X(x)f_Y(y) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4}.$$

然而，联合密度 $f(x, y) = \frac{1}{4}(1 + xy)$ 与 $\frac{1}{4}$ 不相等，除非 $xy = 0$ 在区间内。但 xy 不是恒等于零的。因此 X 和 Y 不是独立的。

考虑 X^2 和 Y^2 的范围，它们的联合分布密度函数 $f_{X^2,Y^2}(x, y)$ 实际上是通过变换 X 和 Y 得到的联合分布：

$$f_{X^2,Y^2}(x^2, y^2) = \begin{cases} \frac{1}{4}(1 + xy), & \sqrt{x^2} < 1, \sqrt{y^2} < 1, \\ 0, & \text{otherwise.} \end{cases}$$

由于 $\sqrt{x^2} < 1$ 和 $\sqrt{y^2} < 1$，即 $|x| < 1$ 和 $|y| < 1$，我们对 X^2 和 Y^2 的边缘密度函数进行积分：
对 X^2:

$$f_{X^2}(x^2) = \int_{-1}^{1} \delta(u - x^2)\, du = 1, \quad x^2 \in [0, 1].$$

对 Y^2:

$$f_{Y^2}(y^2) = \int_{-1}^{1} \delta(v - y^2)\, dv = 1, \quad y^2 \in [0, 1].$$

由于 X^2 和 Y^2 的边缘密度函数均匀且没有依赖关系，且 $f_{X^2,Y^2}(x^2, y^2)$ 可以表示为：

$$f_{X^2,Y^2}(x^2, y^2) = f_{X^2}(x^2) f_{Y^2}(y^2),$$

因此，X^2 和 Y^2 是独立的。

d) 协方差计算与独立性示例
设 X 为一个期望为 0、方差为 1 的正态分布随机变量，W 独立于 X，且 W 以相同的概率取 1 或者 -1。
首先计算 $\mathbb{E}[Y]$ 和 $\mathbb{E}[XY]$：

$$\mathbb{E}[Y] = \mathbb{E}[WX] = \mathbb{E}[W]\mathbb{E}[X] = 0.$$

因为 $\mathbb{E}[X] = 0$ 和 W 独立于 X，所以 $\mathbb{E}[W] = 0$。
计算 $\mathbb{E}[XY]$：

$$\mathbb{E}[XY] = \mathbb{E}[X \cdot WX] = \mathbb{E}[WX^2] = \mathbb{E}[W]\mathbb{E}[X^2] = 0 \cdot 1 = 0.$$

所以 $\mathrm{cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0 - 0 = 0$。
考虑 $W = 1$ 或 $W = -1$ 的情况下：
- 当 $W = 1$ 时，$Y = X$； - 当 $W = -1$ 时，$Y = -X$。
无论 W 取 1 还是 -1，Y 的分布和 X 的分布有完全的依赖关系，因此 X 和 Y 不是独立的。具体地，给定 X 的值，Y 只能取两个值之一：X 或 $-X$，这表明 Y 和 X 之间有依赖关系。

e) 证明 Y 为一个正态分布
设 X 为一个期望为 0、方差为 1 的正态分布随机变量，定义

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

其中 c 为某一常数。
要证明 Y 为正态分布，考虑它的概率密度函数。由于 X 是正态分布，假设其概率密度函数为 $f_X(x)$，即

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

对于 $|X| \leq c$，我们有 $Y = X$，所以在 $|X| \leq c$ 的范围内，Y 的概率密度函数为 $f_Y(y) = f_X(y)$。
对于 $|X| > c$，我们有 $Y = -X$，所以在 $|X| > c$ 的范围内，Y 的概率密度函数为 $f_Y(y) = f_X(-y)$。
所以 Y 的概率密度函数为：

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, & \text{if } |y| \leq c, \\ \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, & \text{if } |y| > c, \end{cases}$$

注意到 $f_X(y) = f_X(-y)$，所以 $f_Y(y) = f_X(y)$，即

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}.$$

这表明 Y 的分布与 X 的分布相同，因此 Y 也是正态分布。

f) 边际分布与独立性
我们需要求出三维随机向量 (X, Y, Z) 的联合密度函数：

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

并且判断 X, Y, Z 是否两两独立以及是否相互独立。
求边际分布 $f_X(x)$：

$$f_X(x) = \int_0^{2\pi} \int_0^{2\pi} f(x, y, z)\, dy\, dz.$$

计算这个积分：

$$f_X(x) = \int_0^{2\pi} \int_0^{2\pi} \frac{1}{8\pi^3}(1 - \sin x \sin y \sin z)\, dy\, dz.$$

将积分分开：

$$f_X(x) = \frac{1}{8\pi^3} \int_0^{2\pi} \int_0^{2\pi} (1 - \sin x \sin y \sin z)\, dy\, dz.$$

我们分开计算每一部分：

$$\int_0^{2\pi} \int_0^{2\pi} 1\, dy\, dz = (2\pi)(2\pi) = 4\pi^2,$$

$$\int_0^{2\pi} \int_0^{2\pi} \sin x \sin y \sin z\, dy\, dz = \sin x \int_0^{2\pi} \sin y\, dy \int_0^{2\pi} \sin z\, dz.$$

因为 $\int_0^{2\pi} \sin y\, dy = 0$ 和 $\int_0^{2\pi} \sin z\, dz = 0$，所以：

$$\int_0^{2\pi} \int_0^{2\pi} \sin x \sin y \sin z\, dy\, dz = \sin x \cdot 0 \cdot 0 = 0.$$

因此，

$$f_X(x) = \frac{1}{8\pi^3} \cdot 4\pi^2 = \frac{1}{2\pi}.$$

对边际分布 $f_Y(y)$ 和 $f_Z(z)$：
由于对称性，同样有：

$$f_Y(y) = \frac{1}{2\pi},$$

$$f_Z(z) = \frac{1}{2\pi}.$$

判断两两独立性和相互独立性

我们计算联合边际分布 $f_{X,Y}(x,y)$：

$$f_{X,Y}(x,y) = \int_0^{2\pi} f(x,y,z)\,dz = \int_0^{2\pi} \frac{1}{8\pi^3}(1 - \sin x \sin y \sin z)\,dz.$$

将积分分开：

$$f_{X,Y}(x,y) = \frac{1}{8\pi^3}\int_0^{2\pi}(1 - \sin x \sin y \sin z)\,dz.$$

我们分开计算每一部分：

$$\int_0^{2\pi} 1\,dz = 2\pi,$$

$$\int_0^{2\pi} \sin x \sin y \sin z\,dz = \sin x \sin y \int_0^{2\pi} \sin z\,dz = \sin x \sin y \cdot 0 = 0.$$

因此，

$$f_{X,Y}(x,y) = \frac{1}{8\pi^3}\cdot 2\pi = \frac{1}{4\pi^2}.$$

注意到 $f_X(x)f_Y(y) = \frac{1}{2\pi}\cdot\frac{1}{2\pi} = \frac{1}{4\pi^2}$，所以：

$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$

这表明 X 和 Y 是独立的。类似地，可以证明 X 和 Z、Y 和 Z 也是独立的。

要判断 X,Y,Z 是否相互独立，需要验证：

$$f(x,y,z) = f_X(x)f_Y(y)f_Z(z).$$

注意到：

$$f_X(x)f_Y(y)f_Z(z) = \left(\frac{1}{2\pi}\right)\left(\frac{1}{2\pi}\right)\left(\frac{1}{2\pi}\right) = \frac{1}{8\pi^3}.$$

然而，

$$f(x,y,z) = \frac{1}{8\pi^3}(1 - \sin x \sin y \sin z) \neq \frac{1}{8\pi^3}$$

因此，X,Y,Z 不是相互独立的。 ◀

Question 3 (10') (大数定律).

a) 首先 X_n 的期望和方差：

$$\mathbb{E}[X_n] = 0$$

$$\mathbb{E}[X_n^2] = \left(\frac{n^2}{2n\log n}\right) + \left(\frac{n^2}{2n\log n}\right) = \frac{n}{\log n}$$

证明 $\{X_n\}$ 满足弱大数律

弱大数律表明，如果 $\mathbb{E}[X_n] = 0$ 并且 $\mathrm{Var}(X_n)$ 收敛到 0，则 $\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{\mathbb{P}} 0$。

我们首先计算 $\frac{1}{n}\sum_{i=1}^n X_i$ 的方差：

$$\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \frac{i}{\log i}$$

我们需要证明这个方差收敛到 0。使用积分比较法，我们有：

$$\sum_{i=1}^{n} \frac{i}{\log i} \sim \int_1^n \frac{x}{\log x}\,dx$$

做变量代换 $u = \log x$，则 $du = \frac{dx}{x}$：

$$\int_1^n \frac{x}{\log x}\,dx = \int_0^{\log n} \frac{e^u}{u}\,du$$

$$\int_0^{\log n} \frac{e^u}{u}\,du$$

这个积分趋向于 $\frac{e^u}{u}$ 逐渐增长，但由于它被 $\log n$ 所限制，整体上它是 $\mathcal{O}(\frac{n}{\log n})$ 级别的，因此有：

$$\frac{1}{n^2}\sum_{i=1}^{n} \frac{i}{\log i} = \mathcal{O}\left(\frac{1}{n\log n}\right)$$

显然，当 $n \to \infty$ 时，这个表达式收敛到 0。根据切比雪夫不等式，我们可以断定：

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\mathbb{P}} 0$$

因此，随机变量序列 $\{X_n\}$ 满足弱大数律。

强大数律要求 $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\text{a.s.}} 0$。我们使用Kolmogorov强大数律的一部分，即如果 $\sum \frac{\mathrm{Var}(X_n)}{n^2} < \infty$，则序列满足强大数律。

对于我们的序列 X_n：

$$\sum_{n=2}^{\infty} \frac{\mathrm{Var}(X_n)}{n^2} = \sum_{n=2}^{\infty} \frac{\frac{n}{\log n}}{n^2} = \sum_{n=2}^{\infty} \frac{1}{n\log n}$$

这个级数是发散的，因为它与 $\sum_{n=2}^{\infty}\frac{1}{n}$ 的行为类似，这是一个调和级数，是发散的。因此，不满足 Kolmogorov 条件。

由于不满足 Kolmogorov 强大数律的条件，我们可以断定 $\{X_n\}$ 不满足强大数律。

综上所述，随机变量序列 $\{X_n\}$ 满足弱大数律，但不满足强大数律。

b) 给定的密度函数为：

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

其中 $\sigma_n^2 = \frac{2n^2}{(\log n)^2}$。

该密度函数对应的是拉普拉斯分布（Laplace distribution），期望值为零，方差为：

$$\mathrm{Var}(X_n) = 2\sigma_n^2 = \frac{4n^2}{(\log n)^2}.$$

为了证明随机变量序列 $\{X_n\}$ 满足强大数律，我们可以使用Kolmogorov强大数律。根据Kolmogorov强大数律，如果对于独立随机变量序列 $\{X_n\}$ 有：

$$\sum_{n=2}^{\infty} \frac{\mathrm{Var}(X_n)}{n^2} < \infty,$$

则 $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\text{a.s.}} 0$。
我们计算 $\sum_{n=2}^{\infty} \frac{\mathrm{Var}(X_n)}{n^2}$：

$$\sum_{n=2}^{\infty} \frac{\mathrm{Var}(X_n)}{n^2} = \sum_{n=2}^{\infty} \frac{\frac{4n^2}{(\log n)^2}}{n^2} = \sum_{n=2}^{\infty} \frac{4}{(\log n)^2}.$$

由于 $\sum_{n=2}^{\infty} \frac{1}{(\log n)^2}$ 是一个收敛级数（因为$\frac{1}{(\log n)^2}$的增长速度比$\frac{1}{n}$要快），所以：

$$\sum_{n=2}^{\infty} \frac{4}{(\log n)^2} < \infty.$$

因此，根据Kolmogorov强大数律，我们可以断定 $\{X_n\}$ 满足强大数律。
为了证明不满足弱大数律，我们需要说明 $\frac{1}{n}\sum_{i=1}^{n} X_i$ 的方差不收敛于零。首先，考虑：

$$\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \frac{4i^2}{(\log i)^2}.$$

使用积分逼近法，我们有：

$$\sum_{i=1}^{n} \frac{i^2}{(\log i)^2} \sim \int_1^n \frac{x^2}{(\log x)^2}\, dx.$$

我们可以做变量替换 $u = \log x$，则 $du = \frac{dx}{x}$，并且 $x = e^u$，因此积分变为：

$$\int_1^n \frac{x^2}{(\log x)^2}\, dx = \int_0^{\log n} \frac{e^{2u}}{u^2}\, du.$$

这个积分趋于无穷大，因为 e^{2u} 的增长速度远快于 u^2 的增长速度。因此，我们有：

$$\frac{1}{n^2}\sum_{i=1}^{n} \frac{4i^2}{(\log i)^2}$$

并不会收敛于零，而是趋向于无穷大。
因此，$\{X_n\}$ 不满足弱大数律。
综上所述，我们证明了随机变量序列 $\{X_n, n \geq 2\}$ 满足强大数律，但不满足弱大数律。 ◀

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。

◀