概率论第三次作业

Problem 1(Warm-up Problems)

211240042 林凡琪

- [**Variance (I)**] Let $X_1, X_2, \cdots, X_n$ be pairwise independent random variables. Show that $\mathbf{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbf{Var}[X_i]$.

  Proof:

  Let $Y = \sum_{i=1}^{n} X_i$. Then, we have:

  $$\begin{aligned}
  \mathbf{Var}[Y] &= \mathbf{E}[Y^2] - \mathbf{E}[Y]^2 \\
  &= \mathbf{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] - \left(\mathbf{E}\left[\sum_{i=1}^{n} X_i\right]\right)^2 \\
  &= \mathbf{E}\left[\sum_{i=1}^{n} X_i^2 + 2\sum_{i<j} X_i X_j\right] - \left(\sum_{i=1}^{n} \mathbf{E}[X_i]\right)^2 \\
  &= \sum_{i=1}^{n} \mathbf{E}[X_i^2] + 2\sum_{i<j} \mathbf{E}[X_i X_j] - \left(\sum_{i=1}^{n} \mathbf{E}[X_i]\right)^2 .
  \end{aligned}$$

  Since $X_1, X_2, \cdots, X_n$ are pairwise independent, we have $\mathbf{E}[X_i X_j] = \mathbf{E}[X_i]\mathbf{E}[X_j]$ for $i \neq j$.

  Substituting this into the above equation, we get:

  $$\begin{aligned}
  \mathbf{Var}[Y] &= \sum_{i=1}^{n} \mathbf{E}[X_i^2] + 2\sum_{i<j} \mathbf{E}[X_i]\mathbf{E}[X_j] - \left(\sum_{i=1}^{n} \mathbf{E}[X_i]\right)^2 \\
  &= \sum_{i=1}^{n} (\mathbf{E}[X_i^2] - \mathbf{E}[X_i]^2) \\
  &= \sum_{i=1}^{n} \mathbf{Var}[X_i].
  \end{aligned}$$

  Hence, we have shown that $\mathbf{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbf{Var}[X_i]$.

- [**Variance (II)**] Each member of a group of n players rolls a (fair) die. For any pair of players who throw the same number, the group scores 1 point. Find the mean and variance of the total score of the group.

  Solution:

  设 X 是这组的总分.

  $$E[X] = \frac{1}{6} \times \binom{n}{2} = \frac{n(n-1)}{12}$$

  设 $X_{i,j}$ 表示第 i 个玩家和第 j 个玩家投掷相同数字时该团队得分。则 $X = \sum_{1 \leq i < j \leq n} X_{i,j}$

  $$\mathrm{Var}[X] = E\left[X^2\right] - E[X]^2$$

$$
\begin{aligned}
E\left[X^2\right] &= E\left[\left(\sum_{1\le i<j\le n} X_{i,j}\right)^2\right] \\
&= E\left[\sum_{1\le i<j\le n} X_{i,j}^2 + 2\sum_{1\le i<j<k\le n} X_{i,j}X_{j,k}\right]
\end{aligned}
$$

因为 $X_{i,j}$ 只能取0或 1 , 所以 $X_{i,j}^2 = X_{i,j}$

因此

$$
\begin{aligned}
E\left[X^2\right] &= E\left[\sum_{1\le i<j\le n} X_{i,j} + 2\sum_{1\le i<j<k\le n} X_{i,j}X_{j,k}\right] \\
&= \sum_{1\le i<j\le n} E\left[X_{i,j}\right] + 2\sum_{1\le i<j<k\le n} E\left[X_{i,j}X_{j,k}\right] + 2\sum_{1\le i<j<k<l\le n} E\left[X_{i,j}X_{k,l}\right]
\end{aligned}
$$

由于每一对玩家得到相同数字的概率为 $\frac{1}{6}$, 则 $E\left[X_{i,j}\right] = \frac{1}{6}$, 又因为三个玩家投到相同数字的概率为 $\frac{1}{6^2}$, 则 $E\left[X_{i,j}X_{j,k}\right] = \frac{1}{36}$, 四个玩家投到相同数字的概率为 $\frac{1}{6^3}$, 则 $E\left[X_{i,j}X_{k,l}\right] = \frac{1}{216}$

因此, $E\left[X^2\right] = \binom{n}{2} \times \frac{1}{6} + 2\times\binom{n}{3}\times\frac{1}{36} + 2\times\binom{n}{4}\times\frac{1}{216} = \frac{n(n-1)(n^2+19n+174)}{2592}$

即

$\mathrm{Var}[X] = \binom{n}{2} \times \frac{1}{6} + 2\times\binom{n}{3}\times\frac{1}{36} + 2\times\binom{n}{4}\times\frac{1}{216} - \left(\frac{n(n-1)}{12}\right)^2 = \frac{n(n-1)(n^2+19n+174)}{2592} - \left(\frac{n(n-1)}{12}\right)^2$

(此题致谢陈子元同学)

- [**Variance (III)**] An urn contains n balls numbered $1, 2, \ldots, n$. We select k balls uniformly at random **without replacement** and add up their numbers. Find the mean and variance of the sum.

  Solution:

  Let's define the random variable X as the sum of the numbers on the k balls selected. We can break down X into a sum of indicator random variables X_i for $1 \le i \le n$, where $X_i = i$ if ball i is selected and 0 otherwise.

  Since we are selecting k balls uniformly at random without replacement from a total of n balls, the probability that any particular ball is selected is $\dfrac{k}{n}$. Therefore, the expected value of each X_i is:

  $\mathbf{E}[X_i] = i \cdot \frac{k}{n} + 0 \cdot \left(1 - \frac{k}{n}\right) = \frac{ik}{n}$.

  The expected value of X is:

  $\mathbf{E}[X] = \mathbf{E}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \mathbf{E}[X_i] = \sum_{i=1}^n \frac{ik}{n} = \frac{k}{n}\sum_{i=1}^n i = \frac{k(n+1)}{2}$.

  To find the variance of X, we can use the formula for the variance of a sum of pairwise independent random variables that you mentioned earlier:

  $\mathbf{Var}[X] = \sum_{i=1}^n \mathbf{Var}[X_i]$.

  Since X_i is an indicator random variable with expected value $\dfrac{ik}{n}$, its variance is:

  $\mathbf{Var}[X_i] = \mathbf{E}[X_i^2] - \mathbf{E}[X_i]^2 = i^2 \cdot \frac{k}{n} - \left(\frac{ik}{n}\right)^2 = i^2 \cdot \frac{k}{n} - i^2 \cdot \frac{k^2}{n^2} = i^2 \cdot \frac{k(n-k)}{n^2}$.

  Therefore, the variance of X is:

  $\mathbf{Var}[X] = \sum_{i=1}^n \mathbf{Var}[X_i] = \sum_{i=1}^n i^2 \cdot \frac{k(n-k)}{n^2} = \frac{k(n-k)}{n^2}\sum_{i=1}^n i^2 = \frac{k(n-k)(n+1)(2n+1)}{6n}$.

  So, the mean and variance of the sum are $\dfrac{k(n+1)}{2}$ and $\dfrac{k(n-k)(n+1)(2n+1)}{6n}$, respectively.

- **[Variance (IV)]** Let N be an integer-valued, positive random variable and let $\{X_i\}_{i=1}^{\infty}$ be indepedently identically distributed random variables that are independent of N , too. Precisely, for any finite subset $I \subseteq \mathbb{N}_+$ and N are mutually independent. Let$X = \sum_{i=1}^{N} X_i$, show that $\mathbf{Var}[X] = \mathbf{Var}[X_1]\mathbb{E}[N] + \mathbb{E}[X_1]^2\mathbf{Var}[N]$.

  Proof:

  Let's define the random variable X as $X = \sum_{i=1}^{N} X_i$. We can use the law of total variance to find the variance of X:

  $\mathbf{Var}[X] = \mathbf{E}[\mathbf{Var}[X|N]] + \mathbf{Var}[\mathbf{E}[X|N]]$.

  Since $\{X_i\}_{i=1}^{\infty}$ are independent and identically distributed, we have:

  $\mathbf{Var}[X|N] = \mathbf{Var}\left[\sum_{i=1}^{N} X_i\right] = N\mathbf{Var}[X_1]$.

  Therefore,

  $\mathbf{E}[\mathbf{Var}[X|N]] = \mathbf{E}[N]\mathbf{Var}[X_1]$.

  Also, since $\{X_i\}_{i=1}^{\infty}$ are independent and identically distributed, we have:

  $\mathbf{E}[X|N] = \mathbf{E}\left[\sum_{i=1}^{N} X_i\right] = N\mathbf{E}[X_1]$.

  Therefore,

  $\mathbf{Var}[\mathbf{E}[X|N]] = \mathbf{Var}[N]\mathbf{E}[X_1]^2$.

  Substituting these results into the law of total variance, we get:

  $$\begin{aligned}\mathbf{Var}[X] &= \mathbf{E}[\mathbf{Var}[X|N]] + \mathbf{Var}[\mathbf{E}[X|N]] \\ &= \mathbf{E}[N]\mathbf{Var}[X_1] + \mathbf{Var}[N]\mathbf{E}[X_1]^2.\end{aligned}$$

  Hence, we have shown that $\mathbf{Var}[X] = \mathbf{Var}[X_1]\mathbb{E}[N] + \mathbb{E}[X_1]^2\mathbf{Var}[N]$.

- **[Moments (I)]** Find an example of a random variable with finite j-th moments for $1 \leq j \leq k$ but an unbounded $(k+1)$-th moment. Give a clear argument showing that your choice has these properties.

  Solution:

  One example of a random variable that has finite k-th moments for $k < n$ but an unbounded n-th moment is a random variable X with a probability density function given by:

  $$f_X(x) = \begin{cases} \frac{n}{x^{n+1}} & x \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

  This is known as a Pareto distribution with shape parameter n and scale parameter 1.

  For $k < n$, the k-th moment of X is given by:

  $\mathbf{E}[X^k] = \int_{-\infty}^{\infty} x^k f_X(x)dx = \int_1^{\infty} x^k \frac{n}{x^{n+1}}dx = n\int_1^{\infty} x^{k-n-1}dx.$

  Since $k - n - 1 < -1$, this integral converges to a finite value:

  $\mathbf{E}[X^k] = n\int_1^{\infty} x^{k-n-1}dx = n\left[\frac{x^{k-n}}{k-n}\right]_1^{\infty} = \frac{n}{n-k}.$

  However, for $k = n$, the n-th moment of X is given by:

  $\mathbf{E}[X^n] = \int_{-\infty}^{\infty} x^n f_X(x)dx = \int_1^{\infty} x^n \frac{n}{x^{n+1}}dx = n\int_1^{\infty} dx = \infty.$

  Therefore, the random variable X has finite k-th moments for $k < n$ but an unbounded n-th moment.

- **[Moments (II)]** Let $X \sim \mathrm{Geo}(p)$ for some $p \in (0, 1)$. Find $\mathbb{E}[X^3]$ and $\mathbb{E}[X^4]$.

  Solution:

Let $X \sim \mathrm{Geo}(p)$ for some $p \in (0,1)$. The moment generating function of X is given by $M_X(t) = \dfrac{pe^t}{1-(1-p)e^t}$.

To find $\mathbb{E}[X^3]$, we differentiate the moment generating function thrice and evaluate it at $t=0$:

$$\begin{aligned}\mathbb{E}[X^3] &= M_X^{(3)}(0) \\ &= \frac{\mathrm{d}^3}{\mathrm{d}t^3}\frac{pe^t}{1-(1-p)e^t}\Bigg|_{t=0} \\ &= \frac{2p(1-p)^2}{(1-p)^3}\end{aligned}$$

Similarly, to find $\mathbb{E}[X^4]$, we differentiate the moment generating function four times and evaluate it at $t=0$:

$$\begin{aligned}\mathbb{E}[X^4] &= M_X^{(4)}(0) \\ &= \frac{\mathrm{d}^4}{\mathrm{d}t^4}\frac{pe^t}{1-(1-p)e^t}\Bigg|_{t=0} \\ &= \frac{6p(1-p)^3+8p(1-p)^2+2p(1-p)}{(1-p)^4}\end{aligned}$$

- [**Moments (III)**] Let $X \sim \mathrm{Pois}(\lambda)$ for some $\lambda > 0$. Find $\mathbb{E}[X^3]$ and $\mathbb{E}[X^4]$.

  Solution:

  Let $X \sim \mathrm{Pois}(\lambda)$ for some $\lambda > 0$. The moment generating function of X is given by $M_X(t) = e^{\lambda(e^t-1)}$.

  To find $\mathbb{E}[X^3]$, we differentiate the moment generating function thrice and evaluate it at $t=0$:

  $$\begin{aligned}\mathbb{E}[X^3] &= M_X^{(3)}(0) \\ &= \frac{\mathrm{d}^3}{\mathrm{d}t^3}e^{\lambda(e^t-1)}\Bigg|_{t=0} \\ &= \lambda(\lambda+1)(\lambda+2)\end{aligned}$$

  Similarly, to find $\mathbb{E}[X^4]$, we differentiate the moment generating function four times and evaluate it at $t=0$:

  $$\begin{aligned}\mathbb{E}[X^4] &= M_X^{(4)}(0) \\ &= \frac{\mathrm{d}^4}{\mathrm{d}t^4}e^{\lambda(e^t-1)}\Bigg|_{t=0} \\ &= \lambda(\lambda+1)(\lambda+2)(\lambda+3)\end{aligned}$$

- [**Covariance and correlation (I)**] Let X and Y be discrete random variables with correlation ρ. Show that $|\rho| \le 1$.

  Proof:

  设 X 和 Y 为具有相关系数 ρ 的离散随机变量。

  X 和 Y 的相关系数定义为 $\rho = \dfrac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}$。

  根据柯西-施瓦茨不等式，

  $$\begin{aligned}|\rho| &= \left|\frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}\right| \\ &\le \frac{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}{\sigma_X\sigma_Y} \\ &= 1\end{aligned}$$

  因此，$|\rho| \le 1$。

- [**Covariance and correlation (II)**]

  Proof:

$$\begin{aligned}\mathbb{E}(\max\{X^2,Y^2\}) &= \mathbb{E}\left(\frac{1}{2}(X^2+Y^2+|X^2-Y^2|)\right)\\ &= \frac{1}{2}\mathbb{E}(X^2+Y^2)+\frac{1}{2}\mathbb{E}(|X^2-Y^2|)\end{aligned}$$

由于 X 和 Y 的方差为 1，我们知道 $\mathbb{E}(X^2)=\mathbb{E}(Y^2)=1$。因此，第一项变为 $\frac{1}{2}\mathbb{E}(X^2+Y^2)=1$ 。对于第二项，我们可以使用柯西-施瓦茨不等式来获得一个上界：

$$\begin{aligned}\mathbb{E}(|X^2-Y^2|) &\le \sqrt{\mathbb{E}((X^2-Y^2)^2)}\\ &= \sqrt{\mathbb{E}(X^4-2X^2Y^2+Y^4)}\\ &= \sqrt{\mathbb{E}(X^4)+\mathbb{E}(Y^4)-2\mathbb{E}(X^2Y^2)}\end{aligned}$$

$E\left(\left|X^2-Y^2\right|\right)=E\left[\left(X^2-Y^2\right)I_{X^2\le Y^2}\right]+E\left[\left(Y^2-X^2\right)I_{X^2<Y^2}\right]$

由于 X和Y的均值都为 0 , 所以

$$\operatorname{Cov}\left(X^2,Y^2\right)=E\left(X^2Y^2\right)-E\left(X^2\right)E\left(Y^2\right)=E\left(X^2Y^2\right)-1$$

因此,

$$\begin{aligned}E\left(\max\left\{X^2,Y^2\right\}\right) &= \frac{1}{2}\left(2+E\left(\left|X^2-Y^2\right|\right)\right)\\ &\le \frac{1}{2}\left(2+2\sqrt{\left(1-\rho^2\right)}\right)\\ &= 1+\sqrt{1-\rho^2}\end{aligned}$$

因此, 得证。

(此题致谢陈子元同学.)

- [**Covariance and correlation (III)**] Construct two random variables X and Y such that their covariance $\mathbf{Cov}(X,Y)=0$ but X and Y are not independent. You should prove your construction is true.

  Solution:

  例子:两个随机变量X和Y 它们的协方差 $\mathbf{Cov}(X,Y)=0$ 但 X 和 Y 不是独立的：

  设 X 是一个随机变量，它以相等的概率取值 -1、0 和 1。设 $Y=X^2$。那么，$\mathbf{E}[X]=0$，$\mathbf{E}[Y]=\frac{1}{3}(-1)^2+\frac{1}{3}(0)^2+\frac{1}{3}(1)^2=\frac{2}{3}$，且 $\mathbf{E}[XY]=\mathbf{E}[X^3]=\frac{1}{3}(-1)^3+\frac{1}{3}(0)^3+\frac{1}{3}(1)^3=0$。因此，$\mathbf{Cov}(X,Y)=\mathbf{E}[XY]-\mathbf{E}[X]\mathbf{E}[Y]=0-0\cdot\frac{2}{3}=0$。然而，$X$ 和 Y 显然不是独立的，因为知道 Y 的值会告诉我们一些关于 X 的值的信息。例如，如果 $Y=1$，那么我们知道 X 必须是 -1 或 1。

Problem 2 (Inequalities)

- **[Reverse Markov's inequality]** Let X be a discrete random variable with bounded range $0\le X\le U$ for some $U>0$. Show that $\mathbf{Pr}(X\le a)\le\dfrac{U-\mathbf{E}[X]}{U-a}$ for any $0<a<U$.

  Proof:

  Let X be a discrete random variable with bounded range $0\le X\le U$ for some $U>0$. Let a be any value such that $0<a<U$. Define the indicator function $I_{X\le a}$ as follows: $I_{X\le a}=1$ if $X\le a$, and $I_{X\le a}=0$ otherwise. Then, we have:

  $$\mathbf{E}[I_{X\le a}]=\mathbf{Pr}(X\le a)$$

  Since $I_{X\le a}$ is either 0 or 1, we have:

  $$\mathbf{E}[I_{X\le a}]=\mathbf{Pr}(X\le a)=\frac{\mathbf{E}[I_{X\le a}]}{\mathbf{E}[I_{X\le a}]}=\frac{\mathbf{E}[I_{X\le a}]}{\mathbf{E}[1]}$$

  Now, let's consider the quantity $U-X$. Since $0\le X\le U$, we have $0\le U-X\le U$. Therefore, we can apply Markov's inequality to the random variable $U-X$ to obtain:

$$\mathbf{Pr}(U-X \geq U-a) = \mathbf{Pr}(X \leq a) = \frac{\mathbf{E}[I_{X\leq a}]}{\mathbf{E}[1]} = \frac{\mathbf{E}[U-X]}{\mathbf{E}[U-a]} = \frac{\mathbf{E}[U]-\mathbf{E}[X]}{\mathbf{E}[U]-\mathbf{E}[a]} = \frac{\mathbf{E}[U]-\mathbf{E}[X]}{\mathbf{E}[U]-a} = \frac{(U)-(U-\mathbf{E}[X])}{(U)-a} = \frac{(U-\mathbf{E}[X])}{(U)-a}$$

Thus, we have shown that for any discrete random variable X with bounded range $0 \leq X \leq U$ for some $U > 0$, and for any value a such that $0 < a < U$, we have:

$$\mathbf{Pr}(X \leq a) \leqslant \frac{(U-\mathbf{E}[X])}{(U)-a}$$

- **[Markov's inequality]** Let X be a discrete random variable. Show that for all $\beta \geq 0$ and all $x > 0$, $\mathbf{Pr}(X \geq x) \leq \mathbb{E}(e^{\beta X})e^{-\beta x}$.

  Proof:

  To prove this, let $\beta \geq 0$ and $x > 0$. Then, for any $X \geq 0$, we have:

  $e^{\beta X} \geq e^{\beta x}$ if and only if $X \geq x$.

  Therefore, $e^{\beta X} \geq e^{\beta x}\mathbf{1}_{\{X\geq x\}}$, where $\mathbf{1}_{\{X\geq x\}}$ is the indicator function that takes the value 1 if $X \geq x$ and 0 otherwise.

  Taking the expectation of both sides, we get:

  $\mathbb{E}(e^{\beta X}) \geq e^{\beta x}\mathbb{E}(\mathbf{1}_{\{X\geq x\}})$

  $= e^{\beta x}\mathbf{Pr}(X \geq x)$

  Dividing both sides by $e^{\beta x}$, we get:

  $$\mathbf{Pr}(X \geq x) \leq \frac{\mathbb{E}(e^{\beta X})}{e^{\beta x}} = \mathbb{E}(e^{\beta X})e^{-\beta x}$$

- **[Cantelli's inequality]** Let X be a discrete random variable with mean 0 and variance σ^2. Prove that for any $\lambda > 0$, $\mathbf{Pr}[X \geq \lambda] \leq \dfrac{\sigma^2}{\lambda^2+\sigma^2}$. (Hint: You may first show that $\mathbf{Pr}[X \geq \lambda] \leq \dfrac{\sigma^2+u^2}{(\lambda+u)^2}$ for all $u > 0$.)

  Proof:

  首先用 Markov 不等式来证明 $\mathbf{Pr}[X \geq \lambda] \leq \dfrac{\sigma^2+u^2}{(\lambda+u)^2}$ 对于所有 $u > 0$ 成立。

  由于 X 的期望值为 0，有 $\mathbf{E}[(X+u)^2] = \sigma^2 + u^2$。

  然后，应用 Markov 不等式，得到：

  $$\begin{aligned}\mathbf{Pr}[X \geq \lambda] = \mathbf{Pr}[X+u \geq \lambda+u] = \mathbf{Pr}[(X+u)^2 \geq (\lambda+u)^2] \\ \leq \frac{\mathbf{E}[(X+u)^2]}{(\lambda+u)^2} = \frac{\sigma^2+u^2}{(\lambda+u)^2}\end{aligned}$$

  通过最小化 $\dfrac{\sigma^2+u^2}{(\lambda+u)^2}$ 来找到最佳的 u 值。对分子求导并令其等于零，得到 $u = \lambda$。因此，

  $$\mathbf{Pr}[X \geq \lambda] \leq \frac{\sigma^2+u^2}{(\lambda+u)^2} = \frac{\sigma^2+\lambda^2}{(\lambda+\lambda)^2} = \frac{\sigma^2}{\lambda^2+\sigma^2}$$

- **[The weak law of large numbers]** Let $X_1, X_2, \cdots, X_n$ be independent and identically distributed random variables with mean μ and finite variance, use Chebyshev's inequality to show that for any constant $\epsilon > 0$ we have $\lim_{n\to\infty} \mathbf{Pr}\left(\left|\frac{X_1+X_2+\cdots+X_n}{n} - \mu\right| > \epsilon\right) = 0$.

  Proof:

  设 $S_n = X_1 + X_2 + \cdots + X_n$，则 $\mathbf{E}[S_n] = n\mu$。由于 $X_1, X_2, \cdots, X_n$ 独立且具有相同的分布，有 $\mathrm{Var}(S_n) = n\sigma^2$，其中 σ^2 是每个随机变量的方差。

  应用 Chebyshev 不等式，得：

$$\mathbf{Pr}\left(\left|\frac{S_n}{n}-\mu\right|>\epsilon\right)=\mathbf{Pr}\left(|S_n-n\mu|>n\epsilon\right)$$
$$\leq\frac{\mathrm{Var}(S_n)}{(n\epsilon)^2}=\frac{n\sigma^2}{n^2\epsilon^2}=\frac{\sigma^2}{n\epsilon^2}$$

由于 $\lim\limits_{n\to\infty}\frac{\sigma^2}{n\epsilon^2}=0$,

$$\lim_{n\to\infty}\mathbf{Pr}\left(\left|\frac{S_n}{n}-\mu\right|>\epsilon\right)=0$$

- **[Median trick]** Suppose we want to estimate the value of Z. Let $\mathcal{A}$ be a randomized algorithm that outputs $\widehat{Z}$ satisfying $\mathbf{Pr}[(1-\epsilon)Z\leq\widehat{Z}\leq(1+\epsilon)Z]\geq\frac{3}{4}$ for some fixed parameter $\epsilon>0$. We run $\mathcal{A}$ independently for $2n-1$ times, and obtain the outputs $\widehat{Z}_1,\widehat{Z}_2,\cdots,\widehat{Z}_{2n-1}$. Let X be the median (中位数) of $\widehat{Z}_1,\widehat{Z}_2,\cdots,\widehat{Z}_{2n-1}$. Use Chebyshev's inequality to show that $\mathbf{Pr}[(1-\epsilon)Z\leq X\leq(1+\epsilon)Z]=1-O(1/n)$. (Remark: The bound can be drastically improved with Chernoff bound).

  Proof:

  设 Y_i 为一个随机变量，当 $(1-\epsilon)Z\leq\widehat{Z}_i\leq(1+\epsilon)Z$ 时取值为 1，否则取值为 0。

  由题意可知，$\mathbf{E}[Y_i]=\mathbf{Pr}[(1-\epsilon)Z\leq\widehat{Z}_i\leq(1+\epsilon)Z]\geq\frac{3}{4}$。

  设 $Y=\sum_{i=1}^{2n-1}Y_i$，则 $\mathbf{E}[Y]=(2n-1)\mathbf{E}[Y_i]\geq\frac{3}{2}n-\frac{3}{4}$。

  由 Chebyshev 不等式，

  $$\mathbf{Pr}\left[|Y-\mathbf{E}[Y]|\geq n/2\right]\leq\frac{\mathrm{Var}(Y)}{(n/2)^2}$$

  由于 Y_i 是独立的随机变量，有 $\mathrm{Var}(Y)=(2n-1)\mathrm{Var}(Y_i)$。

  又因为 Y_i 是一个伯努利随机变量，所以 $\mathrm{Var}(Y_i)=p(1-p)$，其中 $p=\mathbf{E}[Y_i]$。

  因此，

  $$\mathbf{Pr}\left[|Y-\mathbf{E}[Y]|\geq n/2\right]\leq\frac{(2n-1)p(1-p)}{(n/2)^2}=O(1/n)$$

  这意味着，以至少 $1-O(1/n)$ 的概率，Y 的值在其期望值的 $n/2$ 范围内。

  也就是说，

  $$\mathbf{Pr}\left[\frac{n}{4}<Y<2n-\frac{n}{4}\right]=1-O(1/n)$$

  这意味着，在至少 $n/4$ 个 $\widehat{Z}$ 中，$(1-\epsilon)Z$ 和 $(1+\epsilon)Z$ 之间有一个值。

  因此，在这些值中，中位数 X 必定在这个范围内。

  所以，

  $$\mathbf{Pr}[(1-\epsilon)Z\leq X\leq(1+\epsilon)Z]=1-O(1/n)$$

Problem 3 (Probability meets graph theory)

- [**Common neighbor**] Let $p\in(0,1)$ be a constant. Show that with a probability approaching to 1 (as n tends to infinity) the Erdős–Rényi random graph $\mathbf{G}(n,p)$ has the property that every pair of its vertices has a common neighbor. (Hint: You may use Markov's inequality.)

  Proof:

  对于任意两个顶点 u 和 v，它们没有公共邻居的概率为 $(1-p^2)^{n-2}$，因为剩下的 $n-2$ 个顶点都不与它们同时相邻。

由于有 $\binom{n}{2}$ 对顶点，根据线性期望，

$\mathbb{E}[X] = \binom{n}{2}(1-p^2)^{n-2}$

由于 p 是常数且大于 0，所以当 n 趋近于无穷大时，$(1-p^2)^{n-2}$ 趋近于 0。因此，$\mathbb{E}[X]$ 也趋近于 0 。

使用马尔科夫不等式来估计 $\mathbf{Pr}(X \geq 1)$。

对于任意正数 t，我们有：

$\mathbf{Pr}(X \geq 1) \leq \frac{\mathbb{E}[X]}{1} = \mathbb{E}[X]$

由于 $\mathbb{E}[X]$ 趋近于 0，所以 $\mathbf{Pr}(X \geq 1)$ 也趋近于 0。这意味着当 n 趋近于无穷大时，$\mathbf{G}(n,p)$ 中每对顶点都有一个公共邻居的概率趋近于 1。

- [**Isolated vertices**] Prove that $p = \log n/n$ is the threshold probability for the disappearance of isolated vertices. Formally, you are required to show that
  - with a probability approaching to 1 (as n tends to infinity) the Erdős–Rényi random graph $\mathbf{G} = \mathbf{G}(n,p)$ has the property that G has no isolated vertices when $p = \omega(\log n/n)$;

    Proof:

    设 X 为 Erdős–Rényi 随机图 $\mathbf{G} = \mathbf{G}(n,p)$ 中孤立顶点的数量。对于任意一个顶点 v，它是孤立的概率为 $(1-p)^{n-1}$，因为剩下的 $n-1$ 个顶点都不与它相邻。由于有 n 个顶点，根据线性期望，则，

    $$\mathbf{E}[X] = n(1-p)^{n-1}$$

    当 $p = \omega(\log n/n)$ 时，

    $$\lim_{n\to\infty} \frac{\log(1-p)}{\log n/n} = \lim_{n\to\infty} \frac{\log(1-\omega(\log n/n))}{\log n/n} = -\infty$$

    因此，

    $$\lim_{n\to\infty} (1-p)^{n-1} = 0$$

    所以，

    $$\lim_{n\to\infty} \mathbf{E}[X] = 0$$

    由 Markov 不等式，

    $$\mathbf{Pr}[X > 0] \leq \frac{\mathbf{E}[X]}{1} = \mathbf{E}[X]$$

    因此，

    $$\lim_{n\to\infty} \mathbf{Pr}[X > 0] = 0$$

    这意味着，当 $p = \omega(\log n/n)$ 时，以概率趋近于 1，Erdős–Rényi 随机图 $\mathbf{G} = \mathbf{G}(n,p)$ 没有孤立顶点。
  - with a probability approaching to 0 (as n tends to infinity) the Erdős–Rényi random graph $\mathbf{G} = \mathbf{G}(n,p)$ has the property that $\mathbf{G}$ has no isolated vertices when $p = o(\log n/n)$.

    Proof:

    设 X 表示在 $\mathbf{G}(n,p)$ 中孤立顶点的数量。

    $\mathbb{E}[X] = n(1-p)^{n-1}$

    当 $p = o(\log n/n)$ 时：

    $\mathbb{E}[X] = n(1-p)^{n-1} \leq ne^{-p(n-1)} = ne^{-pn} = o(1)$

    因此，当 $p = o(\log n/n)$ 时，$\mathbb{E}[X]$ 趋近于 0。

    由 Markov 不等式，对于任意正数 t，

$$\mathbf{Pr}(X \geq 1) \leq \frac{\mathbb{E}[X]}{1} = \mathbb{E}[X]$$

由于 $\mathbb{E}[X]$ 趋近于 0，所以 $\mathbf{Pr}(X \geq 1)$ 也趋近于 0。这意味着当 n 趋近于无穷大且 $p = o(\log n/n)$ 时，$\mathbf{G}(n, p)$ 中没有孤立顶点的概率趋近于 1。