Concrete Math: Homework 5

Due on June 9, 2022 at 14:00

Professor Xie Wei

SA21011018
Zhou Enshuai

2022 年 5 月 29 日

Problem 1

Show that $2\sqrt{n+1}-2 \le \sum_{k=1}^{n} \frac{1}{\sqrt{k}} \le 2\sqrt{n}-1$.

Solution

用积分放缩法证明上述不等式, 采用的积分函数是 $f(x)=\frac{1}{\sqrt{x}}$:

$$\int_1^{n+1} \frac{1}{\sqrt{x}}dx \le \sum_{k=1}^{n} \frac{1}{\sqrt{k}} \le 1 + \int_1^{n} \frac{1}{\sqrt{x}}dx$$

$$2\sqrt{x}\Big|_1^{n+1} \le \sum_{k=1}^{n} \frac{1}{\sqrt{k}} \le 1 + 2\sqrt{x}\Big|_1^{n}$$

$$2\sqrt{n+1}-2 \le \sum_{k=1}^{n} \frac{1}{\sqrt{k}} \le 2\sqrt{n}-1$$

Problem 2

Using the idea of generating function, solve the recurrences:

1. $f_0 = 1, f_1 = 2, f_n = 2f_{n-1} - f_{n-2} + (-2)^n \ for \ n \geq 2.$
2. $g_0 = 0, h_0 = 1, g_1 = h_1 = 2, g_n = 2h_{n-1} - g_{n-2}, h_n = g_{n-1} - h_{n-2} \ for \ n \geq 2.$

Solution

1.
注意到这里 $f_1 = 2$, 不满足初始递推关系, 修正如下:

$$f_n = 2f_{n-1} - f_{n-2} + (-2)^n[n \geq 0] + 2[n = 1]$$

带入生成函数可得:

$$\begin{aligned} F(z) = \sum_n f_n z^n &= \sum_n \left(2f_{n-1} - f_{n-2} + (-2)^n[n \geq 0] + 2[n = 1]\right) z^n \\ &= \sum_n 2f_{n-1}z^n - \sum_n f_{n-2}z^n + \sum_n (-2)^n[n \geq 0]z^n + \sum_n 2[n = 1]z^n \\ &= 2zF(z) - z^2F(z) + \frac{1}{1+2z} + 2z \end{aligned}$$

解得:

$$\begin{aligned} F(z) &= \frac{1}{(1+2z)(z-1)^2} + \frac{2z}{(z-1)^2} \\ &= \frac{4}{9} \cdot \frac{1}{1+2z} - \frac{16}{9} \cdot \frac{1}{1-z} + \frac{7}{3} \cdot \frac{1}{(1-z)^2} \end{aligned}$$

所以:

$$f_n = [z^n]F(z) = \frac{4}{9}(-2)^n - \frac{16}{9} + \frac{7}{3}(n+1), \ for \ n \geq 0.$$

2.
注意到这里 $h_0 = 1, h_1 = 2$, 不满足初始递推关系, 对所有的 n 修正如下:

$$\begin{cases} h_n = g_{n-1} - h_{n-2} + [n = 0] + 2[n = 1] \\ g_n = 2h_{n-1} - g_{n-2} \end{cases}$$

从而

$$\begin{cases} H(z) = \sum_n h_n z^n = zG(z) - z^2H(z) + 1 + 2z \\ G(z) = \sum_n g_n z^n = 2zH(z) - z^2G(z) \end{cases}$$

解得:

$$\begin{cases} H(z) = \frac{(1+z^2)(1+2z)}{z^4+1} = (2z^3 + z^2 + 2z + 1)\frac{1}{z^4+1} = (2z^3 + z^2 + 2z + 1)\sum_{n \geq 0}(-1)^n z^{4n} \\ G(z) = \frac{2z(1+2z)}{z^4+1} = (4z^2 + 2z)\frac{1}{z^4+1} = (4z^2 + 2z)\sum_{n \geq 0}(-1)^n z^{4n} \end{cases}$$

所以:

$$h_n = [z^n]H(z) = \begin{cases} (-1)^{\frac{n}{4}},\ n\%4 = 0 \\ 2(-1)^{\frac{n-1}{4}},\ n\%4 = 1 \\ (-1)^{\frac{n-2}{4}},\ n\%4 = 2 \\ 2(-1)^{\frac{n-3}{4}},\ n\%4 = 3 \end{cases} = \begin{cases} (-1)^{\lfloor \frac{n}{4} \rfloor},\ n\%4 = 0 \text{ or } n\%4 = 2 \\ 2 \cdot (-1)^{\lfloor \frac{n}{4} \rfloor},\ n\%4 = 1 \text{ or } n\%4 = 3 \end{cases}$$

$$g_n = [z^n]G(z) = \begin{cases} 0,\ n\%4 = 0 \text{ or } n\%4 = 3 \\ 2(-1)^{\frac{n-1}{4}},\ n\%4 = 1 \\ 4(-1)^{\frac{n-2}{4}},\ n\%4 = 2 \end{cases} = \begin{cases} 0,\ n\%4 = 0 \text{ or } n\%4 = 3 \\ 2 \cdot (-1)^{\lfloor \frac{n}{4} \rfloor},\ n\%4 = 1 \\ 4 \cdot (-1)^{\lfloor \frac{n}{4} \rfloor},\ n\%4 = 2 \end{cases}$$

Problem 3

A random variable X is said to have the Poisson distribution with mean λ if $\Pr(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$ for all $k \in \mathbb{N}$. Let X_1 and X_2 be independent random Poisson variables both with variance t. Calculate the distribution of $X_1 + X_2$.

Solution

先证明服从泊松分布的均值为 λ 的随机变量 X 的方差也是 λ, 设 $X \sim Pois(\lambda)$:

$$
\begin{aligned}
VarX &= EX^2 - (EX)^2 = EX^2 - \lambda^2 \\
&= \sum_{k\geq 0} \frac{e^{-\lambda}\lambda^k k^2}{k!} - \lambda^2 \\
&= e^{-\lambda}\lambda \sum_{k\geq 1} \frac{\lambda^{k-1}k}{(k-1)!} - \lambda^2 \\
&= e^{-\lambda}\lambda \sum_{k\geq 0} \frac{\lambda^k(k+1)}{k!} - \lambda^2 \\
&= e^{-\lambda}\lambda \sum_{k\geq 0} \frac{\frac{\mathrm{d}\lambda^{k+1}}{\mathrm{d}\lambda}}{k!} - \lambda^2 \\
&= e^{-\lambda}\lambda \frac{\mathrm{d}(\sum_{k\geq 0} \frac{\lambda^{k+1}}{k!})}{\mathrm{d}\lambda} - \lambda^2 \\
&= e^{-\lambda}\lambda \frac{\mathrm{d}(\lambda \sum_{k\geq 0} \frac{\lambda^k}{k!})}{\mathrm{d}\lambda} - \lambda^2 \\
&= e^{-\lambda}\lambda \frac{\mathrm{d}(\lambda e^{\lambda})}{\mathrm{d}\lambda} - \lambda^2 \\
&= e^{-\lambda}\lambda (e^{\lambda} + \lambda e^{\lambda}) - \lambda^2 \\
&= \lambda
\end{aligned}
$$

所以 X_1 和 X_2 独立同分布于参数为 t 的泊松分布, 所以 $X_1 + X_2$ 也服从泊松分布, 即 $X_1 \sim Pois(t)$ 和

Problem 3 continued on next page...

$X_2 \sim Pois(t)$, 设 $X = X_1 + X_2$, 则 X 的概率分布为:

$$
\begin{aligned}
\Pr(X = k) &= \Pr(X_1 + X_2 = k) \\
&= \sum_i \Pr(X_1 = i, X_2 = k - i) \\
&= \sum_i \left[\Pr(X_1 = i) \cdot \Pr(X_2 = k - i)\right] \\
&= \sum_{i=0}^{k} \left[\frac{e^{-t}t^i}{i!} \cdot \frac{e^{-t}t^{k-i}}{(k-i)!}\right] \\
&= \frac{e^{-2t}t^k}{k!} \sum_{i=0}^{k} \frac{k!}{i!(k-i)!} \\
&= \frac{e^{-2t}t^k}{k!} \sum_{i=0}^{k} \binom{k}{i} \\
&= \frac{e^{-2t}t^k}{k!} \cdot 2^k \\
&= \frac{e^{-2t}(2t)^k}{k!}
\end{aligned}
$$

即 $X_1 + X_2 \sim Pois(2t)$.
事实上利用指数生成函数的二项卷积公式可以得到: 若 X_1, X_2 独立, 且 $X_1 \sim Pois(t_1)$, $X_2 \sim Pois(t_2)$, 则 $X_1 + X_2 \sim Pois(t_1 + t_2)$.

Problem 4

If we toss a coin, it comes up heads with probability p, which is fixed but unknown. We toss the coin n times (different tosses are independent), and give an estimate $\hat{p}$ of p. Given small $\varepsilon, \delta > 0$, choose an n as small as possible such that $\Pr(|\hat{p} - p| \geq \varepsilon) \leq \delta$ is satisfied.

Solution

该实验是 n 次重复伯努利试验, 设第 i 次抛掷硬币结果随机变量为 X_i

$$X_i = \begin{cases} 1, & \text{if comes up head at the ith time} \\ 0, & \text{if comes up tail at the ith time} \end{cases}, \quad 0 \leq i \leq n$$

所以 $\Pr(X_i = 1) = p$, 则 $EX_i = p$, 且 $X_1, X_2, \cdots, X_n$ 相互独立.
设 $X = \sum_{i=1}^n X_i$, 则 $EX = \sum_{i=1}^n EX_i = np = \mu$,
本题使用了下面的 chernoff bound:

$$\Pr(|X - \mu| \geq \eta\mu) \leq 2e^{-\eta^2\mu/3}, \quad 0 \leq \eta \leq 1 \tag{1}$$

n 次重复实验对 p 的估计为随机变量 $\hat{p}$

$$\hat{p} = \frac{1}{n}\sum_{i=1}^n X_i = \frac{X}{n}$$

所以

$$\begin{aligned}\Pr(|\hat{p} - p| \geq \varepsilon) &= \Pr\left(\left|\frac{X}{n} - p\right| \geq \varepsilon\right) = \Pr\left(|X - np| \geq n\varepsilon\right) \\ &= \Pr\left(|X - \mu| \geq \frac{np\varepsilon}{p}\right) = \Pr\left(|X - \mu| \geq \mu\frac{\varepsilon}{p}\right)\end{aligned}$$

因为 p 是定值, 而 ε 足够小, 所以 $0 \leq \frac{\varepsilon}{p} \leq 1$, 所以取不等式 (1) 中的 $\eta = \frac{\varepsilon}{p}$ 可得:

$$\Pr(|X - \mu| \geq \frac{\varepsilon}{p}\mu) \leq 2e^{-(\frac{\varepsilon}{p})^2\mu/3} = 2e^{-n\varepsilon^2/3p}$$

所以若要使 $\Pr(|\hat{p} - p| \geq \varepsilon) = \Pr\left(|X - \mu| \geq \mu\frac{\varepsilon}{p}\right) \leq \delta$ 成立, 只需要让 $2e^{-n\varepsilon^2/3p} \leq \delta$ 成立即可:

$$\begin{aligned}2e^{-n\varepsilon^2/3p} \leq \delta &\Leftrightarrow -n\varepsilon^2/3p \leq \ln\frac{\delta}{2} \\ &\Leftrightarrow n\varepsilon^2/3p \geq \ln\frac{2}{\delta} \\ &\Leftrightarrow n \geq \frac{3p}{\varepsilon^2}\ln\frac{2}{\delta}\end{aligned}$$

所以一个尽可能小的满足条件的 n 应该是

$$n = \left\lceil \frac{3p}{\varepsilon^2}\ln\frac{2}{\delta} \right\rceil$$