Discrete Mathematics Homework from Handout 3

高橋 知也 / 23TB8303

Homework 18

■Homework 18. Let $k \geq 1$ be a positive integer and X a discrete random variable with a finite kth moment. Show that for any nonnegative integer $l \leq k$, the lth moment $\mathbb{E}\big(X^l\big)$ is finite. (Hint: if $0 \leq l \leq k$, then for any $i \in \mathbb{R}$, $|i|^l \leq \max\big(1, |i|^k\big)$.)

解答．$0 \leq l \leq k$ をみたす整数 l を取る．任意の $i \in \mathbb{R}$ について

$$|i|^l \leq \max(1, |i|^k) \leq 1 + |i|^k$$

が成り立つ．したがって

$$\mathbb{E}\big(|X|^l\big) = \sum_i |i|^l \Pr(X = i) \leq \sum_i (1 + |i|^k) \Pr(X = i) = 1 + \mathbb{E}\big(|X|^k\big) .$$

仮定より $\mathbb{E}\big(|X|^k\big)$ が有限なので，$1 + \mathbb{E}\big(|X|^k\big)$ は有限である．よって $\mathbb{E}\big(|X|^l\big)$ も有限である．

また $-|X|^l \leq X^l \leq |X|^l$ より

$$-\mathbb{E}\big(|X|^l\big) \leq \mathbb{E}\big(X^l\big) \leq \mathbb{E}\big(|X|^l\big)$$

が成り立つから，$\mathbb{E}\big(X^l\big)$ も有限である． □

Homework 19

■Homework 19. Let X be a nonnegative random variable with a finite positive expectation. Then, for any positive real $a > 0$,

$$\Pr\bigl(X \geq a\mathbb{E}(X)\bigr) \leq \frac{1}{a}.$$

解答. Markov の不等式より, 非負の確率変数 X と任意の $t > 0$ について

$$\Pr(X \geq t) \leq \frac{\mathbb{E}(X)}{t}$$

が成り立つ．ここで $t = a\mathbb{E}(X)$ とおくと，$\mathbb{E}(X) > 0$ かつ $a > 0$ より $t > 0$ であり，

$$\Pr\bigl(X \geq a\mathbb{E}(X)\bigr) \leq \frac{\mathbb{E}(X)}{a\mathbb{E}(X)} = \frac{1}{a}$$

を得る．よって主張が示された． □

Homework 20

■Homework 20. Give an example of a nonnegative random variable X such that there exists a positive a satisfying

$$\Pr(X \geq a\mathbb{E}(X)) = \frac{1}{a}.$$

(Hint: consider an indicator random variable.)

解答. $a > 1$ を 1 つ固定する. 事象 A を $\Pr(A) = \frac{1}{a}$ を満たすように取り, 確率空間 $(\Omega, \mathcal{F}, \Pr)$ の元 $\omega \in \Omega$ を「1 つの結果（標本点)」とする. このとき指示変数

$$X = \mathbf{1}_A = \begin{cases} 1 & (\omega \in A), \\ 0 & (\omega \notin A) \end{cases}$$

と定める. このとき X は非負の確率変数であり,

$$\mathbb{E}(X) = \Pr(A) = \frac{1}{a}$$

だから

$$a\mathbb{E}(X) = a \cdot \frac{1}{a} = 1$$

となる. よって

$$\Pr\bigl(X \geq a\mathbb{E}(X)\bigr) = \Pr(X \geq 1) = \Pr(X = 1) = \Pr(A) = \frac{1}{a}$$

を得る. したがって, この X が条件を満たす非負の確率変数の例である. □

Homework 21

■Homework 21. Show that the converse of Theorem 5 does not hold. (In other words, give an example of random variables X, Y with finite second moments such that $\mathrm{Cov}(X,Y) = 0$ but X and Y are not independent.)

解答. 離散型確率変数 X を

$$X = \begin{cases} -1 & \text{with probability } \frac{1}{4}, \\ 0 & \text{with probability } \frac{1}{2}, \\ 1 & \text{with probability } \frac{1}{4} \end{cases}$$

と定める．このとき X は 0 を中心に対称であり,

$$\mathbb{E}(X) = (-1) \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} = 0$$

である．ここで $Y = X^2$ とおくと,

$$Y = \begin{cases} 0 & \text{with probability } \frac{1}{2}, \\ 1 & \text{with probability } \frac{1}{2} \end{cases}$$

となる．共分散を計算すると,

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(X^3) - 0 \cdot \mathbb{E}(Y) = \mathbb{E}(X^3)$$

である．X^3 の期待値は

$$\mathbb{E}(X^3) = (-1)^3 \cdot \frac{1}{4} + 0^3 \cdot \frac{1}{2} + 1^3 \cdot \frac{1}{4} = -\frac{1}{4} + 0 + \frac{1}{4} = 0$$

なので

$$\mathrm{Cov}(X,Y) = 0$$

が成り立つ.

一方で $Y = X^2$ であるから,

$$\Pr(X = 1, Y = 0) = 0$$

であるが,

$$\Pr(X = 1)\Pr(Y = 0) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8} \neq 0$$

となる．したがって X と Y は独立ではない.

以上より,Theorem 5 の逆は成り立たない. □

Homework 22

■Homework 22. Let $p \in (0,1)$. Consider a sequence of n independent coin flips in which the coin will land on heads with probability p and tails with probability $1-p$. Define X to be the random variable that represents the number of heads in the sequence. Show that for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \Pr\bigl(X \geq (1+\varepsilon)\mathbb{E}(X)\bigr) = 0 \tag{1}$$

and

$$\lim_{n\to\infty} \Pr\bigl(X \leq (1-\varepsilon)\mathbb{E}(X)\bigr) = 0. \tag{2}$$

解答. X は n 回のコイン投げにおける表の回数を表すので,X は二項分布 $B(n,p)$ に従う．二項分布の基本的な性質より

$$\mathbb{E}(X) = np, \qquad \mathrm{Var}(X) = np(1-p) \tag{3}$$

が成り立つ．

チェビシェフ不等式より，任意の $a > 0$ について

$$\Pr\bigl(|X - \mathbb{E}(X)| \geq a\bigr) \leq \frac{\mathrm{Var}(X)}{a^2} \tag{4}$$

が成り立つ．ここで

$$a := \varepsilon\mathbb{E}(X)$$

とおくと，(4) と (3) より

$$\Pr\bigl(|X - \mathbb{E}(X)| \geq \varepsilon\mathbb{E}(X)\bigr) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2\mathbb{E}(X)^2} = \frac{np(1-p)}{\varepsilon^2 n^2 p^2} = \frac{1-p}{\varepsilon^2 np}. \tag{5}$$

式 (1) の証明. 任意の $\varepsilon > 0$ について

$$X \geq (1+\varepsilon)\mathbb{E}(X) \;\Rightarrow\; X - \mathbb{E}(X) \geq \varepsilon\mathbb{E}(X) \;\Rightarrow\; |X - \mathbb{E}(X)| \geq \varepsilon\mathbb{E}(X)$$

である．したがって

$$\Pr\bigl(X \geq (1+\varepsilon)\mathbb{E}(X)\bigr) \leq \Pr\bigl(|X - \mathbb{E}(X)| \geq \varepsilon\mathbb{E}(X)\bigr).$$

これと (5) より

$$0 \leq \Pr\bigl(X \geq (1+\varepsilon)\mathbb{E}(X)\bigr) \leq \frac{1-p}{\varepsilon^2 np}$$

となる．右辺は $n \to \infty$ で 0 に収束するので，はさみうちの原理から

$$\lim_{n\to\infty} \Pr\bigl(X \geq (1+\varepsilon)\mathbb{E}(X)\bigr) = 0,$$

すなわち (1) が従う．

式 (2) **の証明．** 任意の $\varepsilon > 0$ について

$$X \leq (1-\varepsilon)\mathbb{E}(X) \;\Rightarrow\; \mathbb{E}(X) - X \geq \varepsilon\mathbb{E}(X) \;\Rightarrow\; |X - \mathbb{E}(X)| \geq \varepsilon\mathbb{E}(X)$$

である．したがって

$$\Pr\bigl(X \leq (1-\varepsilon)\mathbb{E}(X)\bigr) \leq \Pr\bigl(|X - \mathbb{E}(X)| \geq \varepsilon\mathbb{E}(X)\bigr).$$

(5) を用いると

$$0 \leq \Pr\bigl(X \leq (1-\varepsilon)\mathbb{E}(X)\bigr) \leq \frac{1-p}{\varepsilon^2 np}$$

となり，右辺は $n \to \infty$ で 0 に収束するので，はさみうちの原理より

$$\lim_{n\to\infty} \Pr\bigl(X \leq (1-\varepsilon)\mathbb{E}(X)\bigr) = 0,$$

すなわち (2) が成り立つ．

以上より，任意の $\varepsilon > 0$ について (1) と (2) が示された． □

Homework 23

■Homework 23. Complete the proof of Theorem 12 by proving the lower tail bound

$$\Pr\bigl(X \le (1-\delta)\mathbb{E}(X)\bigr) \le \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mathbb{E}(X)} \qquad (0<\delta<1). \tag{6}$$

解答．Theorem 12 の証明と同様に，$X=\sum_{i=1}^n X_i$ を互いに独立な指示変数の和とする．各 i について，ある $p_i \in [0,1]$ が存在して

$$X_i = \begin{cases} 1 & \text{with probability } p_i, \\ 0 & \text{with probability } 1-p_i \end{cases}$$

と書ける．このとき，各 X_i のモーメント母関数 $M_{X_i}(t)$ について

$$M_{X_i}(t) = \mathbb{E}(e^{tX_i}) = p_i e^t + (1-p_i) = 1 + p_i(e^t-1) \le e^{p_i(e^t-1)}$$

が成り立つ．ここでは任意の実数 x に対して $1+x \le e^x$ という不等式を用いた．これは，e^x に対してマクローリンの定理を適用すると

$$e^x = 1 + x + \frac{x^2}{2}e^{\theta x} \qquad (0<\theta<1)$$

が成り立つので，

$$e^x - (1+x) = \frac{x^2}{2}e^{\theta x} \ge 0$$

となる．したがって，$1+x \le e^x$ が従う．

よって Lemma 10 より X のモーメント母関数 $M_X(t)$ は

$$M_X(t) = \prod_{i=1}^n M_{X_i}(t) \le \prod_{i=1}^n e^{p_i(e^t-1)} = e^{(e^t-1)\sum_{i=1}^n p_i} = e^{(e^t-1)\mathbb{E}(X)} \tag{7}$$

を満たす（$t \in \mathbb{R}$ の任意の値でよい）．

ここで $0<\delta<1$ を固定し，

$$a := (1-\delta)\mathbb{E}(X)$$

とおく．$t<0$ のとき，$X \le a \iff e^{tX} \ge e^{ta}$ であり，Markov の不等式から

$$\Pr(X \le a) = \Pr\bigl(e^{tX} \ge e^{ta}\bigr) \le \frac{\mathbb{E}(e^{tX})}{e^{ta}} = \frac{M_X(t)}{e^{ta}} \tag{8}$$

を得る．(7) と (8) をあわせると

$$\begin{aligned}\Pr\bigl(X \le (1-\delta)\mathbb{E}(X)\bigr) &\le \frac{e^{(e^t-1)\mathbb{E}(X)}}{e^{t(1-\delta)\mathbb{E}(X)}} \\ &= e^{(e^t-1-t(1-\delta))\mathbb{E}(X)} \end{aligned} \tag{9}$$

が得られる．

$t<0$ について,$0<\delta<1$ のとき $\ln(1-\delta)<0$ なので,

$$t := \ln(1-\delta)$$

とすると $e^t = 1-\delta$ であり,

$$e^t - 1 - t(1-\delta) = (1-\delta) - 1 - (1-\delta)\ln(1-\delta) = -\delta - (1-\delta)\ln(1-\delta)$$

となる．これを (9) に代入すると

$$\begin{aligned}
\Pr\bigl(X \le (1-\delta)\mathbb{E}(X)\bigr) &\le e^{\bigl(-\delta-(1-\delta)\ln(1-\delta)\bigr)\mathbb{E}(X)} \\
&= \Bigl(e^{-\delta}(1-\delta)^{-(1-\delta)}\Bigr)^{\mathbb{E}(X)} \\
&= \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mathbb{E}(X)}.
\end{aligned}$$

これは (6) の右辺である．したがって，式 (6) が示された． □

Homework 24

■Homework 24. Assume that $L \geq 1$ in the randomized load balancing algorithm given above. By using Chebyshev 's inequality, instead of the Chernoff bound, derive an upper bound on

$$\Pr\Bigl(\bigcup_{j=0}^{m-1}\Bigl\{X_j \geq (1+\varepsilon)\frac{L}{m}\Bigr\}\Bigr). \tag{10}$$

Set $m = 10$, $L = 25000$, and $\varepsilon = 0.1$, and compare the resulting numerical bound with the one via the Chernoff bound in this section.

解答. ランダム負荷分散アルゴリズムでは

$$X_j = \sum_{i=0}^{n-1} l_i X_{i,j}, \qquad 0 \leq l_i \leq 1, \qquad L = \sum_{i=0}^{n-1} l_i$$

であり，各 $X_{i,j}$ は

$$X_{i,j} = \begin{cases} 1 & \text{with probability } 1/m, \\ 0 & \text{with probability } 1 - 1/m \end{cases}$$

かつ i ごとに独立とする．このとき

$$\mathbb{E}(X_j) = \sum_{i=0}^{n-1} l_i \mathbb{E}(X_{i,j}) = \frac{1}{m}\sum_{i=0}^{n-1} l_i = \frac{L}{m}$$

である．また，独立性より

$$\begin{aligned} \mathrm{Var}(X_j) &= \sum_{i=0}^{n-1} \mathrm{Var}(l_i X_{i,j}) = \sum_{i=0}^{n-1} l_i^2\,\mathrm{Var}(X_{i,j}) \\ &= \Bigl(\frac{1}{m} - \frac{1}{m^2}\Bigr)\sum_{i=0}^{n-1} l_i^2 = \frac{m-1}{m^2}\sum_{i=0}^{n-1} l_i^2 \leq \frac{m-1}{m^2}\sum_{i=0}^{n-1} l_i = \frac{m-1}{m^2}L, \end{aligned} \tag{11}$$

ここで $0 \leq l_i \leq 1$ より $l_i^2 \leq l_i$ を用いた．

いま $\mu := \mathbb{E}(X_j) = L/m$ とおく．任意の $\varepsilon > 0$ に対して

$$X_j \geq (1+\varepsilon)\mu \;\Rightarrow\; X_j - \mu \geq \varepsilon\mu \;\Rightarrow\; |X_j - \mu| \geq \varepsilon\mu$$

なので，

$$\Pr\bigl(X_j \geq (1+\varepsilon)\mu\bigr) \leq \Pr\bigl(|X_j - \mu| \geq \varepsilon\mu\bigr).$$

チェビシェフの不等式より

$$\Pr\bigl(|X_j - \mu| \geq \varepsilon\mu\bigr) \leq \frac{\mathrm{Var}(X_j)}{\varepsilon^2\mu^2} \leq \frac{\frac{m-1}{m^2}L}{\varepsilon^2(L^2/m^2)} = \frac{m-1}{\varepsilon^2 L}$$

であるから，

$$\Pr\bigl(X_j \geq (1+\varepsilon)\tfrac{L}{m}\bigr) \leq \frac{m-1}{\varepsilon^2 L} \qquad (0 \leq j \leq m-1). \tag{12}$$

union bound を用いる：

$$\begin{aligned}\Pr\Biggl(\bigcup_{j=0}^{m-1}\Bigl\{X_j \geq (1+\varepsilon)\tfrac{L}{m}\Bigr\}\Biggr) &\leq \sum_{j=0}^{m-1}\Pr\bigl(X_j \geq (1+\varepsilon)\tfrac{L}{m}\bigr) \\ &\leq m \cdot \frac{m-1}{\varepsilon^2 L} = \frac{m(m-1)}{\varepsilon^2 L}. \end{aligned}\tag{13}$$

したがって (10) に対するチェビシェフ不等式による上界は

$$\Pr\Biggl(\bigcup_{j=0}^{m-1}\Bigl\{X_j \geq (1+\varepsilon)\tfrac{L}{m}\Bigr\}\Biggr) \leq \frac{m(m-1)}{\varepsilon^2 L}$$

である．

つぎに $m = 10$, $L = 25000$, $\varepsilon = 0.1$ を代入すると

$$\frac{m(m-1)}{\varepsilon^2 L} = \frac{10 \cdot 9}{(0.1)^2 \cdot 25000} = 0.36$$

を得る．

一方，本文中の Chernoff 限界からは,「少なくとも 1 台のサーバが平均より 10% 以上重い負荷を受ける確率」の上界として

$$5.6 \times 10^{-5}$$

が得られていた. したがって, チェビシェフの不等式による評価 (13) は 0.36 とかなり粗く,Chernoff 限界の方が精度良く評価できていることがわかる.

考察として、Chernoff 限界の方は定理の適用に「$X_{i,j}$ が独立であること」と「重み l_i が 0 以上 1 以下であること」を必要とするため問題設定を活かしているのに対し、チェビシェフの不等式は定理の適用にそれらの条件を仮定しないため、問題設定を活用できず定理内での評価が甘くなり、結果としてチェビシェフの不等式による評価の方が粗くなったと考えられる。

$\square$