2021 年度　数理科学特別講義 C, G
大阪府立大学
8 月 26 日 第 3 コマ

今野良彦

日本女子大学 理学部 数物科学科

2021 年 8 月 26 日

内容

(1) チェビシェフの不等式

(2) ヘフディンの不等式

記号

$(\Omega, \mathcal{F}, \mathbb{P})$：確率空間

X：$\mathcal{F}$ 可測な実数値確率変数

$^\forall B \in \mathcal{B}(\mathbb{R})$ に対して，$X^{-1}(B) \in \mathcal{F}$ であるので

$$P(B) := \mathbb{P}(\{\omega \in \Omega ; X(\omega) \in B\}) = \mathbb{P}(X \in B)$$

これを $P = \mathbb{P} \circ X^{-1}$ と記す．

$(\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$：$X$ によって誘導された確率空間

定理 8.1 (チェビシェフの不等式)

$$\phi : \mathbb{R} \to [0, \infty)$$

は増加可測関数。このとき, $^\forall a \in \mathbb{R}$ に対して

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[\phi(X)]}{\phi(a)}.$$

ただし, $\mathbb{E}[\phi(X)] < \infty$ である。

証明:

$$\mathbb{E}[\phi(X)] = \int_a^\infty \phi(x)\, dP(x) + \int_{-\infty}^a \phi(x)\, dP(x)$$

$$\geq \phi(a) \int_a^\infty dP(x) = \phi(a)\, \mathbb{P}(X \geq a)$$

□

例 8.1 $X_1, X_2, \ldots, X_n$ は X の独立複製.

$\mu = \mathbb{E}[X]$, $\sigma^2 = \mathbb{V}[X] < \infty$ と仮定. 定理 8.1 において

$$\phi(x) = \begin{cases} x^2 & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

とおけば,

$$\mathbb{P}(|\bar{X}_n - \mu| \geq a) \leq \frac{\mathbb{E}[|\bar{X}_n - \mu|^2]}{a^2} = \frac{\sigma^2}{n a^2}$$

よって,

$$\bar{X}_n \xrightarrow{P} 0 \quad (n \to \infty)$$

特に $a = \sigma\sqrt{\frac{t}{n}}$ $(t > 0)$ とおけば

$$\mathbb{P}\left(|\bar{X}_n - \mu| \geq \sigma\sqrt{\frac{t}{n}}\right) \leq \frac{1}{t} \rightsquigarrow \bar{X}_n = \mu + O_p\left(\frac{1}{\sqrt{n}}\right).$$

例 8.2　$N \in \mathbb{N}$ とし，$j = 1, 2, \dots, N$ に対し，

$g_j : \mathbb{R} \to \mathbb{R}$ は可測，

で，ある $\sigma^2 < \infty$ が存在して，

$$\mathbb{V}[g_j(X)] < \sigma^2 \quad (j = 1, \dots, N)$$

とする．$\forall a > 0$ に対して，

$$\mathbb{P}\left(\max_{1 \leq j \leq N} |(P_n - P) g_j| \geq a \right) \leq \frac{N}{n} \frac{\sigma^2}{a^2}$$

となる。

$$P_n g_j = \frac{1}{n} \sum_{i=1}^{n} g_j(X_i)$$

$$P g_j = \int_{\mathbb{R}} g_j(x) \, dP(x) .$$

ここで，

$$a = \gamma\sqrt{\frac{Nt}{n}} \quad (t > 0)$$

とおけば，

$$\mathbb{P}\left(\max_{1\le j\le N} |(P_n - P)g_j| \ge \gamma\sqrt{\frac{Nt}{n}}\right) \le \frac{1}{t}$$

したがって

$$\max_{1\le j\le N} |(P_n - P)g_j| = O_p\left(\sqrt{\frac{N}{n}}\right).$$

$\boxed{/\!/}$

No. 6

<u>ヘフディン条件</u> $X_1, X_2, \ldots, X_n$ は $(\Omega, \mathcal{F}, \mathbb{P})$ 上で定義された独立な確率変数列（同一でなくてよい）とし、ある $c_j > 0$ $(j=1, \ldots, n)$ が存在して、

$$|X_j| < c_j \quad かつ \quad \mathbb{E}[X_j] = 0.$$

補題 8.1 ヘフディング条件を仮定し、

No. 7

$$b^2 = \sum_{j=1}^{n} c_j^2$$

とおく。このとき、任意の $\lambda > 0$ に対して、

$$\mathbb{E}\left[\exp\left\{\lambda \sum_{j=1}^{n} X_j\right\}\right] \leq \exp\left\{\frac{\lambda^2 b^2}{2}\right\}$$

証明 関数 $\mathbb{R} \ni x \longmapsto e^{\lambda x} \in (0, \infty)$ の凸性より

$$\mathbb{E}[\exp(\lambda X_j)] \leq \frac{1}{2}\exp(-\lambda c_j) + \frac{1}{2}\exp(\lambda c_j)$$

$(j = 1, 2, \ldots n)$。

さらに

$$e^{-t} + e^{t} \leq 2\exp\left(\frac{t^2}{2}\right) \qquad (t \geq 0)$$

を利用すれば

$$\mathbb{E}\left[\exp\left(\lambda\sum_{j=1}^{n} X_j\right)\right] = \prod_{j=1}^{n} \mathbb{E}\left[\exp(\lambda X_j)\right]$$

$$\leq \frac{1}{2}\prod_{j=1}^{n}\left[\exp\{-\lambda c_j\} + \exp\{\lambda c_j\}\right]$$

$$\leq \prod_{j=1}^{n}\exp\left(\frac{\lambda^2 c_j^2}{2}\right)$$

□

定理 8.2. ヘフディン条件を仮定し,

$$b^2 := \sum_{j=1}^{n} c_j^2.$$

$\forall a > 0$ に対し,

$$\mathbb{P}\left(\sum_{j=1}^{n} X_j \geq a\right) \leq \exp\left\{-\frac{a^2}{2b^2}\right\}.$$

さらに, $\forall t > 0$ に対し,

$$\mathbb{P}\left(\sum_{j=1}^{n} X_j \geq b\sqrt{2t}\right) \leq e^{-t}.$$

No.10

証明　$\lambda > 0$ とし，$\phi(x) = \exp\{\lambda x\}$ とし

チェビシェフの不等式を用いれば，$\forall a > 0$ に対し，

$$\mathbb{P}\left(\sum_{j=1}^{n} X_j \geq a\right) \leq \exp\left(\frac{\lambda^2 b^2}{2} - \lambda a\right).$$

ここで，

$$\lambda = \frac{a}{b^2}$$

とおけば，

$$\mathbb{P}\left(\sum_{j=1}^{n} X_j \geq a\right) \leq \exp\left(-\frac{a^2}{2b^2}\right).$$

最後に，

$a = b\sqrt{2t}$　$(t > 0)$　とおけばよい。　□

系 8.3

- $X_1, \dots, X_n$ は X の独立複製

- $g_j : \mathbb{R} \to \mathbb{R}$ は可測 $(j = 1, \dots, N,\ N \in \mathbb{N})$.

ある $K > 0$ が存在して,

$$\mathbb{E}[g_j(X)] = 0, \quad \sup_{x \in \mathbb{R}} |g_j(x)| \le K$$

このとき, $t > 0$ に対して,

$$\mathbb{P}\left(|P_n g_j| \le K\sqrt{\frac{2t}{n}} \right) \le 2e^{-t}$$

さらに,

$$\mathbb{P}\left(\max_{1 \le j \le N} |P_n g_j| \ge K\sqrt{\frac{2(t + \log N)}{n}} \right) \le 2e^{-t}.$$

証明. 各 $j=1,\ldots,N$ に対して,

$\{g_j(X_1), g_j(X_2), \ldots, g_j(X_n)\}$ はヘフディング条件をみたすので

定理8.2より, $a>0$ に対し

$$\mathbb{P}\left(|P_n g_j| \geq a\right) \leq 2\exp\left\{-\frac{na^2}{2K^2}\right\}. \quad \text{(a)}$$

上の式で,

$$t = \frac{na^2}{2K^2} \quad (t>0) \iff a = K\sqrt{\frac{2t}{n}}$$

とおけば

$$\mathbb{P}\left(|P_n g_j| \geq K\sqrt{\frac{2t}{n}}\right) \leq 2e^{-t}.$$

(a) より, $a>0$ に対し,

$$\mathbb{P}\left(\max_{1\le j\le N} |P_n g_j| \ge a\right) \le \sum_{j=1}^{N} \mathbb{P}\left(|P_n g_j| \ge a\right)$$

$$\le 2N\exp\left\{-\frac{na^2}{2k^2}\right\} = 2\exp\left\{-\frac{na^2}{2k^2} + \log N\right\}.$$

ここで,

$$t = \frac{1}{2}\left(\frac{na^2}{k^2} - 2\log N\right) \iff a = k\sqrt{\frac{2(t+\log N)}{n}}$$

とおけば:

$$\mathbb{P}\left(\max_{1\le j\le N} |P_n g_j| \ge k\sqrt{\frac{2(t+\log N)}{n}}\right) \le 2e^{-t}.$$

補題 8.2. 系 8.3と同じ設定のもと, No.14

$$\mathbb{E}\left[\max_{1\leq j\leq N}|P_n g_j|\right]\leq\frac{K}{n}\sqrt{2\log(2N)}.$$

証明 $t\in\mathbb{R}$に対し,

$$e^{|t|}\leq e^{t}+e^{-t}$$

に注意する. $\forall\lambda>0$ と $j=1,\dots,N$ に対し

$$\mathbb{E}\left[\exp(\lambda n|P_n g_j|)\right]$$

$$\leq\mathbb{E}\left[\exp(\lambda n P_n g_j)\right]+\mathbb{E}\left[\exp(-\lambda n P_n g_j)\right]$$

$$\leq 2\exp\left(\frac{\lambda K^2}{2}\right).\quad(\text{補題 8.1 より})\quad(a)$$

イエンセンの不等式と(a)より

$$\mathbb{E}\left[\max_{1\le j\le N} n|P_n g_j|\right]$$

$$=\frac{1}{\lambda}\mathbb{E}\left[\log\left\{\exp\left(\lambda\max_{1\le j\le N} n|P_n g_j|\right)\right\}\right]$$

$$\le\frac{1}{\lambda}\log\left\{\mathbb{E}\left[\exp\left(\lambda\max_{1\le j\le N} n|P_n g_j|\right)\right]\right\}$$

(イエンセンの不等式)

$$\le\frac{1}{\lambda}\log\left\{\sum_{j=1}^{N}\mathbb{E}\left[\exp\left(\lambda n|P_n g_j|\right)\right]\right\}$$

$$\le\frac{1}{\lambda}\log\left\{2N\exp\left(\frac{\lambda K^2}{2}\right)\right\}$$ (a)より

$$= \frac{\log(2N)}{\lambda} + \frac{\lambda k^2}{2}.$$

最右辺は $\lambda = \sqrt{\frac{2\log(2N)}{k}}$ のとき最小化され、

最小値は $k\sqrt{2\log(2N)}$ となる。

よって、

$$\mathbb{E}\left[\max_{1\le j\le N} n|P_n g_j|\right] \le k\sqrt{2\log(2N)}.$$

より

$$\mathbb{E}\left[\max_{1\le j\le N} |P_n g_j|\right] \le \frac{k}{n}\sqrt{2\log(2N)}.$$

□