第5章　確率集中不等式

この章では, 独立な確率変数列の期待値からの偏差を制御する道具を導出する. 確率不等式は

$$\mathsf{Pr}\big(Z - \mathsf{E}[Z] \geq t\big) \leq g(t) \qquad (t \in \mathbb{R})$$

という形式で与えられる. ただし, Z はある確率変数で, $g : \mathbb{R} \longrightarrow \mathbb{R}$ は関数である. 特に, $X_1, X_2, \ldots, X_n\,(n \in \mathbb{N})$ は i.i.d. 確率変数列で

$$Z := Z_n = f_n(X_1, X_2, \ldots, X_n), \qquad f_n : \mathbb{R}^n \longrightarrow \mathbb{R}$$

としたとき

$$\mathsf{Pr}\big(Z - \mathsf{E}[Z] \geq t\big) \leq g(n, t), \qquad (t \in \mathbb{R})$$

と書ける. ここで, $g(\,\cdot\,, \cdot\,)$ は 2 つの変数それぞれに関して非減少であることが期待される.

5.1　Chernoff 限界

5.1.1　基本原理

系 5.1 (Chebyshev の不等式)**.** X を $\mathsf{E}[X^2] < \infty$ なる確率変数とする. このとき

$$\mathsf{Pr}\big(|X - \mathsf{E}[X]| \geq t)\big) \leq \frac{\mathsf{Var}[X]}{t^2}, \qquad (t > 0)$$

が成立する.

Proof. Markov の不等式 (命題 1.33) を用いると

$$\begin{aligned}\mathsf{Pr}\big(|X - \mathsf{E}[X]| \geq t)\big) = \mathsf{Pr}\big(|X - \mathsf{E}[X]|^2 \geq t^2)\big) &\leq \frac{\mathsf{E}[(X - \mathsf{E}[X])^2]}{t^2} \\ &= \frac{\mathsf{Var}[X]}{t^2}\end{aligned}$$

からわかる. □

Chebyshev の不等式では, 変換

$$x \mapsto x^2$$

を用いた. この変換を

$$x \mapsto e^{\lambda x}$$

に変えたものを考える. ただし, $\lambda > 0$ は定数である.

$\lambda_0 > 0$ とし, Z を確率変数とする. このとき, 関数

$$[0,\, \lambda_0) \ni \lambda \mapsto \Phi_Z(\lambda) := \log \mathsf{E}\bigl[\exp(\lambda Z)\bigr]$$

を **確率変数 Z の Cramér-Chernoff 変換**[1] という. ただし, $\forall\lambda \in [0,\, \lambda_0)$ に対して, $\mathsf{E}\bigl[e^{\lambda Z}\bigr] < \infty$ と仮定する. 次に, Φ_Z の Fenchel-Legendre **変換**を

$$\Phi_Z^{\vee}(t) := \sup_{\lambda \geq 0}\bigl\{\lambda t - \Phi_Z(\lambda)\bigr\} \qquad (t \geq 0)$$

で定める[2].

系 5.2. $\lambda \geq 0$ とする. 原点を含む近傍で $\mathsf{E}[e^{\lambda Z}] < \infty$ なる確率変数 Z に対して

$$\mathsf{Pr}\bigl(Z \geq t\bigr) \leq \exp\Bigl(-\Phi_Z^{\vee}(t)\Bigr) \qquad (t > 0)$$

が成り立つ.

Proof. 原点を含むある近傍に含まれる $\lambda > 0$ を取る. Markov の不等式 (命題 1.33) を用いると

$$\begin{aligned}
\mathsf{Pr}\bigl(Z \geq t\bigr) &= \mathsf{Pr}\Bigl(\exp(\lambda Z) \geq \exp(\lambda t)\Bigr) \\
&\leq \frac{1}{\exp(\lambda t)}\mathsf{E}\bigl[\exp(\lambda Z)\bigr] \\
&= \exp\Bigl\{-\bigl\{\lambda t - \log \mathsf{E}\bigl[\exp(\lambda Z)\bigr]\bigr\}\Bigr\} \\
&= \exp\Bigl\{-\bigl\{\lambda t - \Phi_Z(\lambda)\bigr\}\Bigr\} \qquad (5.1)
\end{aligned}$$

となる. さらに, $\lambda = 0$ のときは

$$\mathsf{Pr}(Z \geq t) \leq 1 = \exp\bigl[-\{0 \times t - \Phi_Z(0)\}\bigr]$$

[1] $\lambda = 0$ のとき, $\mathsf{E}[e^{\lambda Z}] = 1$ 　となるので, $\Phi(0) = 0$ である.

[2] $\Phi_Z(\lambda) \geq 0$ で $\lambda = 0$ のとき, $\Phi_Z(0) = 0$ となる. よって, $t = 0$ のとき, $\Phi_Z^{\vee}(0) = \sup_{\lambda \geq 0}\bigl\{-\Phi_Z(\lambda)\bigr\} = \Phi_Z(0) = 0$ となる.

となるので, (5.1) は $\lambda \geq 0$ で成立する. $\lambda \geq 0$ は任意だったので

$$\begin{aligned}\mathsf{Pr}\bigl(Z \geq t\bigr) &\leq \exp\Bigl\{-\sup_{\lambda\geq 0}\bigl\{\lambda t - \Phi_Z(\lambda)\bigr\}\Bigr\} \\ &= \exp\Bigl\{-\Phi_Z^{\vee}(t)\Bigr\}\end{aligned}$$

を得る. □

5.1.2 例

例 5.3 (正規分布)**.** $Z \sim \mathsf{N}(0, \sigma^2)\,(\sigma > 0)$ とする. m を $\mathbb{R}$ 上の Lebesgue 測度としたとき, $\lambda \geq 0$ に対して

$$\begin{aligned}\mathsf{E}[e^{\lambda Z}] &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Bigl\{\lambda z - \frac{z^2}{2\sigma^2}\Bigr\} \mathsf{dm}(z) \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Bigl\{-\frac{1}{2\sigma^2}\bigl(z - \sigma^2\lambda\bigr)^2 + \frac{\sigma^2\lambda^2}{2}\Bigr\} \mathsf{dm}(z) \\ &= \exp\Bigl(\frac{\sigma^2\lambda^2}{2}\Bigr) \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Bigl\{-\frac{1}{2\sigma^2}\bigl(z - \sigma^2\lambda\bigr)^2\Bigr\} \mathsf{dm}(z)}_{=1} \\ &= \exp\Bigl(\frac{\sigma^2\lambda^2}{2}\Bigr)\end{aligned}$$

となる. よって

$$\Phi_Z(\lambda) = \exp\Bigl(\frac{\sigma^2\lambda^2}{2}\Bigr)$$

である. これより

$$\begin{aligned}\Phi_Z^{\vee}(t) &= \sup_{\lambda\geq 0}\Bigl\{\lambda t - \frac{\sigma^2\lambda^2}{2}\Bigr\} \\ &= \sup_{\lambda\geq 0}\Bigl\{-\frac{\sigma^2}{2}\Bigl(\lambda - \frac{t}{\sigma^2}\Bigr)^2 - \frac{t^2}{2\sigma^2}\Bigr\} \\ &= \exp\Bigl(-\frac{t^2}{2\sigma^2}\Bigr)\end{aligned}$$

となる. よって, 系 4.5 を用いると

$$\mathsf{Pr}\bigl(Z \geq t\bigr) \leq \exp\Bigl(-\frac{t^2}{2\sigma^2}\Bigr) \qquad (t > 0)$$

を得る. □

例 5.4 (Poisson 分布)**.** $Y \sim \mathsf{Poi}(\nu)\,(\nu > 0)$ とする. $Z := Y - \nu$ とおくと, Z の積率母関数は

$$\begin{aligned}\mathsf{E}\big[e^{\lambda Z}\big] &= e^{-\lambda\mu}\mathsf{E}\big[e^{\lambda Y}\big] = e^{-\lambda\mu}\sum_{k=0}^{\infty} e^{\lambda k}\frac{e^{-\nu}}{k!}\nu^k = e^{-\lambda\mu}e^{-\nu}\sum_{k=0}^{\infty}\frac{(\nu e^{\lambda})^k}{k!}]\\ &= e^{-\lambda\mu}e^{-\nu}e^{\nu e^{\lambda}}\end{aligned}$$

となる. これより

$$\Phi_Z(\lambda) = \log\big\{e^{-\lambda\mu}e^{-\nu}e^{\nu e^{\lambda}}\big\} = \nu\big\{e^{\lambda} - \lambda - 1\big\}$$

となる. ここで

$$f_t(\lambda) := \lambda t - \nu\big\{e^{\lambda} - \lambda - 1\big\}$$

とおくと

$$\dot{f}_t\,(\lambda = t - \nu(e^{\lambda} - 1)$$

となるので, f_t は $\lambda = \log\left(1 + \dfrac{t}{\nu}\right)$ で最大値を取る. このことから

$$\begin{aligned}\Phi_Z^{\vee}(t) &= t\log\left(1 + \frac{t}{\nu}\right) - \nu\left(\left(1 + \frac{t}{\nu}\right) - \log\left(1 + \frac{t}{\nu}\right) - 1\right)\\ &= (t + \nu)\log\left(1 + \frac{t}{\nu}\right) - t\\ &= \nu\left\{\left(1 + \frac{t}{\nu}\right)\log\left(1 + \frac{t}{\nu}\right) - \frac{t}{\nu}\right\}\\ &= \nu h\left(\frac{t}{\nu}\right)\end{aligned}$$

となる. ただし, $h(x) = (1 + x)\log(1 + x) - x$ である. □

==============================

これが劣ガンマであることを加筆する?

==============================

5.1.3 劣 Gauss と劣 Gamma 確率変数

定義 5.5. (1) 確率変数 X は定数 $\nu > 0$ の劣 **Gauss** であるとは

$$\Phi_X(\lambda) \le \frac{\lambda^2\nu}{2} \qquad (\lambda \in \mathbb{R})$$

であることをいう. これを $X \sim \mathsf{SubG}(\nu)$ と記す.
(2) 確率変数 X は定数 $\nu > 0$ と $c > 0$ の右からの劣 **Gamma** であるとは

$$\Phi_X(\lambda) \leq \frac{\lambda^2 \nu}{2(1-2c\lambda)} \qquad \left(0 < \lambda < \frac{1}{c}\right)$$

であることをいう. これを $X \sim \mathsf{\Gamma}_+(\nu, c)$ と記す. さらに, $-X \in \mathsf{\Gamma}_+(\nu, c)$ のとき, $X \in \mathsf{\Gamma}_-(\nu, c)$ と記す. 最後に, $\mathsf{\Gamma}(\nu, c) = \mathsf{\Gamma}_+(\nu, c) \bigcap \mathsf{\Gamma}_-(\nu, c)$ と定める.

補題 5.6. $X \sim \mathsf{subG}(\sigma^2)$ とする. このとき, $\forall t > 0$ に対して

$$\mathsf{Pr}(X \geq t) \;\leq\; \exp\left(-\frac{t^2}{2\sigma^2}\right); \qquad \mathsf{Pr}(X \leq -t) \leq \exp\left(-\frac{t^2}{2\sigma^2}\right)$$

となる.

Proof. Chernoff 限界と呼ばれるテクニックを用いる. $\forall s > 0$ に対して, Markov の不等式 (命題 1.33) を用いる. $X \sim \mathsf{subG}(\sigma^2)$ であることより

$$\mathsf{Pr}(X \geq t) = \mathsf{Pr}\left(e^{sX} \geq e^{st}\right) \leq \frac{\mathsf{E}[e^{sX}]}{e^{st}} \leq \exp\left(\frac{\sigma^2}{2}s^2 - st\right) \tag{5.2}$$

となる. ここで $\phi(s) := (\sigma^2/2)s^2 - st\,(s > 0)$ とおくと

$$\phi(s) \geq \inf_{s>0} \phi(s) = -\frac{t^2}{2\sigma^2}$$

となるので, 不等式 (5.2) は任意の $s > 0$ で成立しているので

$$\mathsf{Pr}(X \geq t) \leq \exp\left(-\frac{t^2}{2\sigma^2}\right)$$

を得る. □

補題 5.7. 確率変数 X は

$$\mathsf{Pr}\left(|X| \geq t\right) \leq 2\exp\left(-\frac{t^2}{2\sigma^2}\right) \tag{5.3}$$

をみたすとする. このとき, $\forall k \geq 1\,(k \in \mathbb{N})$ に対して

$$\mathsf{E}\left[|X|^k\right] \leq (2\sigma^2)^{k/2} k \Gamma\left(\frac{k}{2}\right)$$

となる. 特に

$$\left(\mathsf{E}\left[|X|^k\right]\right)^{1/k} \leq \sigma e^{1/e}\sqrt{k}, \qquad (k \geq 2)$$

と

$$\mathsf{E}\left[|X|\right] \leq \sigma\sqrt{2\pi}$$

である.

Proof. 命題 1.32 と補題 5.6 より, $t > 0$ に対して

$$\begin{aligned}
\mathsf{E}\left[|X|^k\right] &= \int_0^\infty \mathsf{Pr}(|X| > t^{1/k})\,\mathrm{d}t \qquad (\because \text{命題 } 1.32)\\
&\le 2\int_0^\infty \exp\left(-\frac{t^{2/k}}{2\sigma^2}\right)\mathrm{d}t \qquad (\because (5.3))\\
&= (2\sigma^2)^{k/2}k\int_0^\infty e^{-u}u^{(k/2)-1}\,\mathrm{d}u \qquad \left(\because u = \frac{t^{2/k}}{2\sigma^2} \text{ と変換}\right)\\
&= (2\sigma^2)^{k/2}k\Gamma\left(\frac{k}{2}\right)
\end{aligned}$$

が成り立つことがわかる. 2 番目の主張は, $\forall k \ge 2$ に対して

$$\Gamma\left(\frac{k}{2}\right) \le \left(\frac{k}{2}\right)^{k/2}; \qquad k^{1/k} \le e^{1/e}$$

を用いると

$$\left\{(2\sigma^2)^{k/2}k\Gamma\left(\frac{k}{2}\right)\right\}^{1/k} \le k^{1/k}\sqrt{\frac{2\sigma^2 k}{2}} \le e^{1/e}\sigma\sqrt{k}$$

となる. さらに, $k = 1$ に対しては

$$\sqrt{2}\Gamma\left(\frac{1}{2}\right) = \sqrt{2\pi}$$

よりわかる. □

補題 5.8. X を確率空間 $(\Omega, \mathcal{F}, \mathsf{Pr})$ 上の確率変数で以下をみたすものとする.
(1) $\mathsf{E}[X] = 0$.
(2) $\forall t > 0$ に対して

$$\mathsf{Pr}(X \ge t) \le \exp\left(-\frac{t^2}{2\sigma^2}\right); \qquad \mathsf{Pr}(X \le -t) \le \exp\left(-\frac{t^2}{2\sigma^2}\right).$$

ただし, $\sigma > 0$ である. このとき, $\forall s > 0$ に対して

$$\mathsf{E}\left[e^{sX}\right] \le e^{4\sigma^2 s^2}$$

となる.

Proof. 指数関数に対する Taylor 展開を用いる.

$$
\begin{aligned}
\mathsf{E}[e^{sX}] &\le 1 + \sum_{k=1}^{\infty} \frac{s^k \mathsf{E}[|X|^k]}{k!} \\
&\le 1 + \sum_{k=1}^{\infty} \frac{(2\sigma^2 s^2)^{k/2} k\Gamma(k/2)}{k!} \qquad (\because \text{補題 } 5.7) \\
&= 1 + \sum_{k=1}^{\infty} \frac{(2\sigma^2 s^2)^{k} (2k)\Gamma(k)}{(2k)!} \\
&\qquad + \sum_{k=1}^{\infty} \frac{(2\sigma^2 s^2)^{k+1/2} (2k+1)\Gamma(k+1/2)}{(2k+1)!} \\
&\le 1 + \left(2 + \sqrt{2\sigma^2 s^2}\right) \sum_{k=1}^{\infty} \frac{(2\sigma^2 s^2)^{k} k!}{(2k)!} \\
&\qquad \left(\because\ 2(k!)^2 \le (2k)!\right) \\
&= 1 + \left(1 + \sqrt{\frac{\sigma^2 s^2}{2}}\right)\left(e^{2\sigma^2 s^2} - 1\right) \\
&= e^{2\sigma^2 s^2} + \sqrt{\frac{\sigma^2 s^2}{2}}\left\{e^{2\sigma^2 s^2} - 1\right\} \\
&\le e^{4\sigma^2 s^2}
\end{aligned}
$$

からわかる. □

注意 5.9. 補題 5.8 の証明の最後の不等式は以下からわかる. $x > 0$ に対して, $x = 2\sigma^2 s^2$ とおいて考える.

$$
\begin{aligned}
e^{2x} \ge e^x + \sqrt{\frac{x}{4}}(e^x - 1) \ &\Leftrightarrow\ e^x \ge 1 + \sqrt{\frac{x}{4}}(1 - e^{-x}) \qquad (\because \text{両辺 } e^{-x} \text{ をかけた}) \\
&\Leftrightarrow\ e^x + e^{-x}\sqrt{\frac{x}{4}} \ge 1 + \sqrt{\frac{x}{4}}
\end{aligned}
$$

となる. 最後の不等式は以下からわかる. $x > 0$ であった. 左辺から右辺を引くと

$$
e^{2x} - e^x - \sqrt{\frac{x}{4}}(e^x - 1) = (e^x - 1)\left(e^x - \sqrt{\frac{x}{4}}\right)
$$

となる. ここで $x > 0$ だから

$$
e^x - 1 > 0
$$

である. また

$$
e^{2x} - \frac{x}{4}
$$

を考えると，その導関数は

$$2e^{2x} - \frac{1}{4} > 0$$

であるから，これは単調増加である．したがって

$$e^{2x} - \frac{x}{4} > 1$$

となり

$$e^{2x} > \frac{x}{4}$$

である．両辺は正であるから平方根をとって

$$e^x > \sqrt{\frac{x}{4}}$$

を得る．よって

$$(e^x - 1)\left(e^x - \sqrt{\frac{x}{4}}\right) > 0$$

であり

$$e^{2x} \geq e^x + \sqrt{\frac{x}{4}}(e^x - 1)$$

が従う． □

命題 5.10. $X \in \mathsf{\Gamma}(\nu, c)\,(\nu > 0,\, c > 0)$ とする．このとき, 任意の $t > 0$ に対して

$$\mathsf{Pr}\big(X > \sqrt{2\nu t} + ct\big) \leq e^{-t}, \qquad \mathsf{Pr}\big(-X > \sqrt{2\nu t} + ct\big) \leq e^{-t}$$

が成り立つ.

Proof. $\Phi_X(\lambda) \leq \dfrac{\nu\lambda^2}{2(1-c\lambda)}$ から, $0 < \lambda < c^{-1}$ に対して

$$t\lambda - \Phi_X(\lambda) \geq t\lambda - \frac{\nu\lambda^2}{2(1-c\lambda)}$$

となることに注意する．すると

$$\begin{aligned}
\Phi_X^{\vee}(t) &= \sup_{\lambda \geq 0}\big\{t\lambda - \Phi_X(\lambda)\big\} \\
&\geq \sup_{\lambda \in (0,\,1/c)}\big\{t\lambda - \Phi_X(\lambda)\big\} \\
&\geq \sup_{\lambda \in (0,\,1/c)}\left(t\lambda - \frac{\nu\lambda^2}{2(1-c\lambda)}\right) = \frac{\nu}{c^2} g\left(\frac{ct}{\nu}\right)
\end{aligned}$$

となる[3]. ただし

$$g(u) = 1 + u - \sqrt{1+2u} \qquad (u \geq 0)$$

である. □

補題 5.11. $0 < b \leq \infty$ とする. $\Phi : [0, b) \longrightarrow [0, \infty)$ は凸関数で $[0, b)$ 上で微分可能で $\Phi(0) = \dot{\Phi}(0) = 0$ をみたすものとする. 任意の $t \geq 0$ に対して

$$\Phi^\vee(t) = \sup_{\lambda \in [0, b)} \left\{ \lambda t - \Phi(\lambda) \right\}$$

とおく. このとき, $\Phi^\vee$ は $(0, \infty)$ 上の正値, 狭義増加かつ凸関数で, $\Phi^\vee(0) = 0$ と

$$\{\Phi^\vee\}^{-1}(y) = \inf_{\lambda \in (0, b)} \left[\frac{y + \Phi(\lambda)}{\lambda} \right] \tag{5.4}$$

となる.

Proof. ① $\Phi^\vee(0) = 0$ の証明: まず, 仮定から Φ は非減少関数で, $[0, b)$ 上で非負である. すると $\Phi^\vee(0)$ は 0 を含む非正値関数の最大なので, $\Phi^\vee(0) = 0$ がわかる.
② $\Phi^\vee$ の凸性の証明: sup の性質から, $\forall t_1, t_2 > 0$ と $0 < \forall a < 1$ に対して

$$\begin{aligned}
&\Phi^\vee(at_1 + (1-a)t_2) \\
&\quad = \sup_{\lambda \in (0, b)} \left\{ \lambda \left\{ at_1 + (1-a)t_2 \right\} - \Phi(\lambda) \right\} \\
&\quad \leq \sup_{\lambda \in (0, b)} \left\{ a\lambda t_1 + (1-a)\lambda t_2 - a\Phi(\lambda) - (1-a)\Phi(\lambda) \right\} \\
&\quad \leq a \sup_{\lambda \in (0, b)} \left\{ \lambda t_1 - \Phi(\lambda) \right\} + (1-a) \sup_{\lambda \in (0, b)} \left\{ \lambda t_2 - \Phi(\lambda) \right\} \\
&\quad \leq a\Phi^\vee(t_1) + (1-a)\Phi^\vee(t_2)
\end{aligned}$$

[3]

$$f_t(\lambda) = t\lambda - \frac{\nu\lambda^2}{2(1-c\lambda)}$$

とおく. すると

$$\dot{f}_t(\lambda) = t - \frac{\lambda\nu}{1-c\lambda} - \frac{c\nu\lambda^2}{2(1-c\lambda)^2} = 0$$

を解けばよい.

からわかる.
③ $\Phi^{\vee}$ の狭義単調増加性の証明: $t_1 < t_2$ に対して, $\lambda t_1 - \Phi(\lambda) \le \lambda t_2 - \Phi(\lambda)\ (\forall\lambda \in (0,\, b))$ なので,

$$\Phi^{\vee}(t_1) = \sup_{\lambda\in(0,\,b)}\big\{\lambda t_1 - \Phi(\lambda)\big\} < \sup_{\lambda\in(0,\,b)}\big\{\lambda t_2 - \Phi(\lambda)\big\} = \Phi^{\vee}(t_2)$$

からわかる.
④ (5.4) の証明:

$$u := \inf_{\lambda\in(0,\,b)}\left[\frac{y+\Phi(\lambda)}{\lambda}\right]$$

とおく. 任意の $t \ge 0$ に対して

$$\begin{aligned} u \ge t \ &\Leftrightarrow\ \frac{y+\Phi(\lambda)}{\lambda} \ge t \quad (\forall\lambda \in (0,\, b)) \\ &\Leftrightarrow\ y \ge \lambda t - \Phi(\lambda) \quad (\forall\lambda \in (0,\, b)) \\ &\Leftrightarrow\ y \ge \sup_{\lambda\in(0,\,b)}\left\{\lambda t - \Phi(\lambda)\right\} = \Phi^{\vee}(t) \end{aligned}$$

を得る. 一方, $\Phi^{\vee}$ の右逆関数 $\{\Phi^{\vee}\}^{\leftarrow}$ の定義と $u \ge t$ に注意すると

$$\{\Phi^{\vee}\}^{\leftarrow}(y) = \sup\big\{t \ge 0;\ \Phi^{\vee}(t) \le y\big\} = u$$

となる. さらに, $\Phi^{\vee}$ は狭義単調増加なので, 逆関数を持つので, 逆関数は右逆関数に一致するので

$$\{\Phi^{\vee}\}^{-1}(y) = u$$

がわかる. □

命題 5.12. $b > 0$ とする. $Z_1,\, Z_2,\, \ldots,\, Z_N$ は確率変数列で, $\forall\lambda \in (0,\, b)$ に対して

$$\Phi_{Z_j}(\lambda) \le \Phi(\lambda) \qquad \big(j \in \{1,\, 2,\, \ldots,\, N\}\big)$$

をみたすとする. ただし, Φ は微分可能な凸関数で $\Phi(0) = \dot{\Phi}(0) = 0$ をみたすのものある. このとき

$$\mathsf{E}\left[\max_{j=1,\,2,\,\ldots,\,N} Z_j\right] \le \{\Phi^{\vee}\}^{-1}(\log N)$$

が成り立つ.

Proof. 任意の $\lambda \in (0,\, b)$ に対して

$$\exp\left(\lambda \mathsf{E}\left[\max_{j=1,\,2,\,\ldots,\,N} Z_j\right]\right) \le \sum_{j=1}^{n} \mathsf{E}\bigl[\exp(\lambda Z_j)\bigr] \le N\exp\bigl(\Phi(\lambda)\bigr)$$

となる．この不等式の辺々に対数をとると

$$\lambda \mathsf{E}\left[\max_{j=1,\,2,\,\ldots,\,N} Z_j\right] \le \log N + \Phi(\lambda)$$

を得る．$\lambda \in (0,\, b)$ は任意だったことと (5.4) に注意すると

$$\mathsf{E}\left[\max_{j=1,\,2,\,\ldots,\,N} Z_j\right] \le \inf_{\lambda\in(0,\,b)}\left[\frac{\log N + \Phi(\lambda)}{\lambda}\right] = \bigl\{\Phi^{\vee}\bigr\}^{-1}(\log N)$$

を得る． □

5.1.4 Hoeffding の不等式

補題 5.13. (Hoeffding の補題) X を確率変数で $\mathsf{E}[X]=0$ とし，ほとんど確実に $X \in [a,\, b]$ とする．ただし，$a<b$ である．このとき，$\forall s \in \mathbb{R}$ に対して

$$\mathsf{E}[e^{sX}] \le \exp\left(\frac{s^2}{8}(b-a)^2\right)$$

が成立する．特に，$X \sim \mathsf{subG}\left(\dfrac{(b-a)^2}{4}\right)$ である．

Proof. $\Phi_X(s) := \log \mathsf{E}\bigl[e^{sX}\bigr]$ $(s\in\mathbb{R})$ とする．このとき

$$\begin{aligned}
\overset{\bullet}{\Phi}_X(s) &:= \frac{\mathrm{d}}{\mathrm{d}s}\Phi_X(s) = \frac{\mathsf{E}[Xe^{sX}]}{\mathsf{E}[e^{sX}]},\\
\overset{\bullet\bullet}{\Phi}_X(s) &:= \frac{\mathrm{d}^2}{\mathrm{d}s^2}\Phi_X(s) = \frac{\mathsf{E}[X^2e^{sX}]}{\mathsf{E}[e^{sX}]} - \left\{\frac{\mathsf{E}[Xe^{sX}]}{\mathsf{E}[e^{sX}]}\right\}^2
\end{aligned} \tag{5.5}$$

となる．しかし，ほとんど確実に $X\in[a,\, b]$ なので

$$\mathsf{Var}[X] \;=\; \mathsf{Var}\left[X - \frac{a+b}{2}\right] \le \mathsf{E}\left[\left(X-\frac{a+b}{2}\right)^2\right] \le \frac{(b-a)^2}{4}$$

となる．このこととキュムラント関数の性質 [4] から

$$\overset{\bullet\bullet}{\Phi}_X(s) \le \frac{(b-a)^2}{4}$$

[4] X の p.d.f./p.m.f. を p としたとき，$\mathsf{q}_s(x) = e^{sx-\kappa(s)}\mathsf{p}(x)$ は指数型分布族となる．ただし，$\kappa(s) = \int_a^b e^{-sx}\mathsf{p}(x)\,\mathrm{d}x$(連続型の場合) である．この分布に従う確率変数を $\widetilde{X}$ と記す．すると (5.5) と指数型分布族の性質から

$$\overset{\bullet\bullet}{\Phi}_X(s) = \overset{\bullet\bullet}{\kappa}(x) = \mathsf{Var}[\widetilde{X}]$$

である．ここで, $\Phi_X(0) = \log 1 = 0$ および $\dot{\Phi}_X(0) = \mathsf{E}[X] = 0$ に注意すれば

$$\begin{aligned}
\Phi_X(s) &= \Phi_X(s) - \Phi_X(0) \\
&= \int_0^s \dot{\Phi}_X(u)\,\mathrm{d}\mathsf{m}(u) \\
&= \int_0^s \left\{ \dot{\Phi}_X(u) - \dot{\Phi}_X(0) \right\} \mathrm{d}\mathsf{m}(u) \\
&= \int_0^s \left\{ \int_0^u \ddot{\Phi}_X(r)\,\mathrm{d}r \right\} \mathrm{d}\mathsf{m}(u) = \int_0^s \int_0^u \ddot{\Phi}_X(r)\,\mathrm{d}\mathsf{m}(r)\,\mathrm{d}\mathsf{m}(u) \\
&\le \frac{(b-a)^2}{4} \int_0^s \int_0^u \mathrm{d}r\,\mathrm{d}\mathsf{m}(u) = \frac{(b-a)^2}{4} \int_0^s u\,\mathrm{d}\mathsf{m}(u) \\
&= \frac{(b-a)^2}{8} s^2
\end{aligned}$$

を得る．ただし, m は $\mathbb{R}$ 上の Lebesgue 測度である． □

命題 5.14. (Hoeffding の不等式) $X_1, X_2, \ldots, X_n$ は独立な確率変数列で, ほとんと確実に $X_j \in [a_j, b_j]\,(a_j < b_j;\, j = 1, 2, \ldots, n)$ とする．

$$\overline{X} = \frac{1}{n} \sum_{j=1}^{n} X_j$$

としたとき, $\forall t > 0$ に対して

$$\begin{aligned}
\mathsf{Pr}\big(\overline{X} - \mathsf{E}[\overline{X}] \ge t\big) &\le \exp\left(-\frac{2n^2t^2}{\sum_{j=1}^{n}(b_j - a_j)^2} \right), \\
\mathsf{Pr}\big(\overline{X} - \mathsf{E}[\overline{X}] \le -t\big) &\le \exp\left(-\frac{2n^2t^2}{\sum_{j=1}^{n}(b_j - a_j)^2} \right)
\end{aligned}$$

を得る．

Proof. $Y_j = X_j - \mathsf{E}[X_j]\,(j = 1, 2, \ldots, n)$ とおく．すると $Y_j \in \big[a_j - \mathsf{E}[X_j],\, b_j - \mathsf{E}[X_j]\big]$ かつ $\mathsf{E}[Y_j] = 0$ となる．ここで $Y_j\,(j = 1, 2, \ldots, n)$ に対して補題 5.13 を適用すると

$$\mathsf{E}\big[e^{sY_j}\big] \le \exp\left(\frac{s^2}{8}(b_j - a_j)^2 \right) \tag{5.6}$$

となる．一方, $\mathsf{Pr}(a \le \widetilde{X} \le b) = 1$ であることと関数 $[a, b] \ni x \mapsto \left(x - \frac{a+b}{2} \right)^2 \in \mathbb{R}$ は端点のいずれかで最大となる．その値はいずれも $(b-a)^2/4$ である．よって

$$\mathsf{Var}[\widetilde{X}] \le \mathsf{Var}\left[\widetilde{X} - \frac{a+b}{2} \right] \le \frac{(b-a)^2}{4}$$

からわかる．したがって, $\ddot{\Phi}_X(s) \le \dfrac{(b-a)^2}{4}$ がわかる．

を得る．よって

$$
\begin{aligned}
\mathsf{Pr}\bigl(\overline{X}-\mathsf{E}[\overline{X}]\geq t\bigr) &= \mathsf{Pr}\left(\sum_{j=1}^{n} Y_j \geq nt\right) \\
&= \mathsf{Pr}\left(\exp\left(s\sum_{j=1}^{n} Y_j\right) \geq e^{nst}\right) \\
&\leq \mathsf{E}\left[\exp\left(s\sum_{j=1}^{n} Y_j\right)\right]e^{-nst} \\
&\qquad (\because \text{Markov の不等式 (命題 1.33)}) \\
&= \prod_{j=1}^{n} \mathsf{E}[\exp(sY_j)]e^{-nst} \\
&\leq \prod_{j=1}^{n} \exp\left(\frac{s^2}{8}(b_j-a_j)^2\right)e^{-nst} \qquad (\because (5.6)) \\
&= \exp\left(\frac{s^2}{8}\sum_{j=1}^{n}(b_j-a_j)^2 - nst\right) \\
&= \exp\left\{\frac{\sum_{j=1}^{n}(b_j-a_j)^2}{8}\left(s-\frac{4nt}{\sum_{j=1}^{n}(b_j-a_j)^2}\right)^2 \right. \\
&\qquad \left. - \frac{2n^2t^2}{\sum_{j=1}^{n}(b_j-a_j)^2}\right\}
\end{aligned}
$$

を得る．ここで

$$
s = \frac{4nt}{\sum_{j=1}^{n}(b_j-a_j)^2}
$$

とおくと

$$
\mathsf{Pr}\bigl(\overline{X}-\mathsf{E}[\overline{X}]\geq t\bigr) \leq \exp\left(-\frac{2n^2t^2}{\sum_{j=1}^{n}(b_j-a_j)^2}\right)
$$

を得る． □

注意 5.15 (Chernoff-Okamoto の不等式)**.** $X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{Ber}(p)$ $(0 < p < 1)$ とする．$\overline{X} = n^{-1}\sum_{j=1}^{n} X_j$ とおくと $n\overline{X} \sim \mathsf{Bino}(n, p)$ となる．命題 5.14 から, $t > 0$ に対して

$$
\mathsf{Pr}\bigl(\overline{X} \geq p+t\bigr) \leq \exp\bigl(-2nt^2\bigr)
$$

を得る．

たとえば, $p = \dfrac{1}{2}$ とすると

$$
\mathsf{Pr}(n\overline{X}=n) = \mathsf{Pr}(n\overline{X}\geq n) = \mathsf{Pr}\left(\overline{X}-\frac{1}{2}\geq\frac{1}{2}\right) \leq \exp\left(-\frac{n}{2}\right)
$$

n	2	10	20
$\mathsf{Pr}(X = n)$	0.25	0.00098	9.54×10^{-7}
$\exp(-n/2)$	0.37	0.00674	4.54×10^{-5}

となる．また

$$\mathsf{Pr}(n\overline{X} \geq \frac{n3}{4}) = \mathsf{Pr}\left(\overline{X} - \frac{1}{2} \geq \frac{1}{4}\right) \leq \exp\left(-\frac{n}{8}\right)$$

である． □

5.1.5 Bennett の不等式

命題 5.16 (Bennett の不等式)**.** $b > 0$ を定数とし，$\phi(u) = e^u - u - 1\,(n \in \mathbb{R})$ とする．$X_1, X_2, \ldots, X_n$ は独立な確率変数列で，$X_j \leq b\,(j = 1, 2, \ldots, n)$ とする．

$$v = \sum_{j=1}^{n} \mathsf{E}\big[X_j^2\big] \quad \text{と} \quad S = \sum_{j=1}^{n} \big\{X_j - \mathsf{E}[X_j]\big\}$$

とおく．このとき

$$\Phi_S(\lambda) = \log \mathsf{E}\big[e^{\lambda S}\big] \leq \frac{v}{b^2}\phi(b\lambda) \qquad (\lambda \geq 0) \tag{5.7}$$

が成り立つ．さらに，$\forall t \geq 0$ に対して

$$\mathsf{Pr}\big(S \geq t) \leq \exp\left\{-\frac{v}{b^2}h\left(\frac{tb}{v}\right)\right\} \tag{5.8}$$

が成り立つ．ただし，$h(x) = (1+x)\log(1+x) - x\,(x > 0)$ である．

Proof. ① (5.7) の証明: X_j/b を X_j とすることで $b = 1$ とできる．関数 $u \mapsto \dfrac{\phi(u)}{u^2}$ は $\mathbb{R}$ 上の減少関数である．$X_j \leq 1\,(j = 1, 2, \ldots, n)$ なので $\lambda X_j \leq \lambda$ であることに注意すると

$$\frac{e^{\lambda X_j} - (\lambda X_j) - 1}{(\lambda X_j)^2} \leq \frac{e^\lambda - \lambda - 1}{\lambda^2} \Rightarrow e^{\lambda X_j} - \lambda X_j - 1 \leq X_j^2(e^\lambda - \lambda - 1)$$

を得る．この不等式の両辺の期待値を取ると

$$\begin{aligned} &\mathsf{E}\big[e^{\lambda X_j}\big] - \lambda\mathsf{E}[X_j] - 1 \leq \mathsf{E}\big[X_j^2\big]\phi(\lambda) \\ &\qquad \Rightarrow \mathsf{E}\big[e^{\lambda X_j}\big] \leq 1 + \lambda\mathsf{E}[X_j] + \mathsf{E}\big[X_j^2\big]\phi(\lambda) \end{aligned}$$

がわかる. さらに, 上の不等式の両辺に対数をとると

$$\Phi_{X_j}(\lambda) = \log \mathsf{E}[e^{\lambda X_j}] \le \log\Big(1 + \lambda \mathsf{E}[X_j] + \mathsf{E}[X_j^2]\phi(\lambda)\Big)$$

を得る. このとき

$$\begin{aligned}
\Phi_S(\lambda) &= \log \mathsf{E}\big[e^{\lambda S}\big] = \log\left(\prod_{j=1}^n \mathsf{E}\big[\exp\big\{\lambda\big(X_j - \mathsf{E}[X_j]\big)\big\}\big]\right) \\
&= \sum_{j=1}^n \log \mathsf{E}\big[\exp\big\{\lambda\big(X_j - \mathsf{E}[X_j]\big)\big\}\big] \\
&= \sum_{j=1}^n \log\Big(\mathsf{E}\big[\exp\{\lambda X_j\}\big]\exp\big\{-\lambda \mathsf{E}[X_j]\big\}\Big) \\
&= \sum_{j=1}^n \Big(\log \mathsf{E}\big[\exp\{\lambda X_j\}\big] - \lambda \mathsf{E}[X_j]\Big) \\
&= \sum_{j=1}^n \Big(\Phi_{X_j}(\lambda) - \lambda \mathsf{E}[X_j]\Big) \\
&\le \sum_{j=1}^n \Big\{\log\Big(1 + \lambda \mathsf{E}[X_j] + \mathsf{E}[X_j^2]\phi(\lambda)\Big) - \lambda \mathsf{E}[X_j]\Big\} \\
&\le \sum_{j=1}^n \mathsf{E}[X_j^2]\phi(\lambda) \qquad \big(\because \log(1+y) \le y \; (y > -1)\big) \\
&= v\phi(\lambda)
\end{aligned}$$

となる. あとは, $X_j \to \dfrac{X_j}{b}$; $\lambda \to b\lambda$ と置き換えればよい.

<u>② (5.8) の証明</u>: $\lambda \ge 0$ に対して

$$\begin{aligned}
\mathsf{Pr}\big(S \ge t\big) = \mathsf{Pr}\big(e^{\lambda S} \ge e^{\lambda t}\big) \le e^{-\lambda t}\mathsf{E}\big[e^{\lambda S}\big] &= \exp\Big\{-\Big(\lambda t - \Phi_S(\lambda)\Big)\Big\} \\
&\le \exp\Big\{-\Big(\lambda t - \frac{v}{b^2}\phi(b\lambda)\Big)\Big\}
\end{aligned}$$

となる. 最後の不等号は (5.7) からわかる. $\lambda\,(\ge 0)$ は任意だったので

$$\mathsf{Pr}\big(S \ge t\big) \le \exp\Big\{-\sup_{\lambda \ge 0}\Big(\lambda t - \frac{v}{b^2}\phi(b\lambda)\Big)\Big\}$$

となる. いま

$$g(t) := \lambda t - \frac{v}{b^2}\phi(b\lambda)$$

とおく. すると

$$\dot{g}(t) = t - \frac{v}{b^2}\bigl(be^{b\lambda} - b\bigr) = 0 \Leftrightarrow \lambda = \frac{1}{b}\log\left(1 + \frac{bt}{v}\right)$$

となるので

$$\begin{aligned}
&\sup_{\lambda \geq 0}\left(\lambda t - \frac{v}{b^2}\phi(b\lambda)\right) \\
&\quad = \frac{t}{b}\log\left(1 + \frac{bt}{v}\right) - \frac{v}{b^2}\phi\left\{\log\left(1 + \frac{bt}{v}\right)\right\} \\
&\quad = \frac{t}{b}\log\left(1 + \frac{bt}{v}\right) - \frac{v}{b^2}\left\{\left(1 + \frac{bt}{v}\right) - \log\left(1 + \frac{bt}{v}\right) - 1\right\} \\
&\quad = \frac{t}{b}\log\left(1 + \frac{bt}{v}\right) \\
&\quad = \frac{v}{b^2}\left\{\left(1 + \frac{bt}{v}\right)\log\left(1 + \frac{bt}{v}\right) - \frac{bt}{v}\right\} \\
&\quad = \frac{v}{b^2}h\left(\frac{bt}{v}\right)
\end{aligned}$$

となるので

$$\mathsf{Pr}\bigl(S \geq t) \leq \exp\left\{-\frac{v}{b^2}h\left(\frac{bt}{v}\right)\right\}$$

がわかる. □

5.2 分散不等式

5.2.1 Efron-Stein の不等式

$X_1, X_2, \ldots, X_n$ を独立な確率変数列とし, $Z := X_1 + X_2 + \cdots + X_n$ とする. すると, 独立性から $\mathsf{E}[Z^2] < \infty$ のとき

$$\mathsf{Var}[Z] = \sum_{j=1}^{n}\mathsf{Var}[X_j] \tag{5.9}$$

が成立する.

いま, $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ を可測関数とし, $Z = h(X_1, X_2, \ldots, X_n)$ とする. Efron-Stein の不等式は, (5.9) に対応する $\mathsf{Var}[Z]$ の上限を与えるものである.

定理 5.17 (Efron-Stein の不等式)**.** $X_1, X_2, \ldots, X_n$ を独立な確率変数列とし, $Z = h(X_1, X_2, \ldots, X_n)$ とする. ただし, $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ を可測関数である. さらに, $\mathsf{E}[Z^2] < \infty$ を仮定する. このとき

$$\mathsf{Var}[Z] \le \sum_{j=1}^{n} \mathsf{E}\big[\mathsf{Var}^{(j)}[Z]\big]$$

が成り立つ. ただし, $j = 1, 2, \ldots, n$ に対して

$$\begin{aligned}\mathsf{E}^{(j)}[\,\cdot\,] &= \mathsf{E}[\,\cdot \,|\, X_1, X_2, \ldots, X_{j-1}, X_{j+1}, \ldots, X_n],\\ \mathsf{Var}^{(j)}[Z] &= \mathsf{E}^{(j)}\big[\big(Z - \mathsf{E}^{(j)}[Z]\big)^2\big]\end{aligned}$$

である.

Proof. $\mathsf{E}_0[\,\cdot\,] = \mathsf{E}[\,\cdot\,]$ とし, $j = 1, 2, \ldots, n$ に対して

$$\mathsf{E}_j[\,\cdot\,] = \mathsf{E}\big[\,\cdot\,\big|\, X_1, X_2, \ldots, X_j\big]$$

とし

$$\Delta_j = \mathsf{E}_j[Z] - \mathsf{E}_{j-1}[Z]$$

とおく. すると

$$\begin{aligned}Z - \mathsf{E}[Z] &= \big(\underbrace{\mathsf{E}_n[Z]}_{=Z} - \mathsf{E}_{j-1}[Z]\big) + \big(\mathsf{E}_{j-1}[Z] - \mathsf{E}_{j-2}[Z]\big) + \cdots\\ &\qquad + \big(\mathsf{E}_1[Z] - \underbrace{\mathsf{E}_0[Z]}_{=\mathsf{E}[Z]}\big)\\ &= \sum_{j=1}^{n} \Delta_j\end{aligned}$$

と書ける. これより

$$\mathsf{Var}[Z] = \mathsf{E}\left[\left(\sum_{j=1}^{n} \Delta_j\right)^2\right] = \sum_{j=1}^{n} \mathsf{E}\big[\Delta_j^2\big] + 2\sum_{\ell > j} \mathsf{E}\big[\Delta_j \Delta_\ell\big] \tag{5.10}$$

となる. $\ell > j$ のとき

$$\begin{aligned}\mathsf{E}_j[\Delta_\ell] &= \mathsf{E}\big[\mathsf{E}_\ell[Z] - \mathsf{E}_{\ell-1}[Z]\big|\, X_1, X_2, \ldots, X_j\big]\\ &= \mathsf{E}\big[\mathsf{E}_\ell[Z]\big|\, X_1, X_2, \ldots, X_j\big] - \mathsf{E}\big[\mathsf{E}_{\ell-1}[Z]\big|\, X_1, X_2, \ldots, X_j\big]\\ &= \mathsf{E}\Big[\mathsf{E}\big[Z\big|\, X_1, X_2, \ldots, X_\ell\big]\Big|\, X_1, X_2, \ldots, X_j\Big]\\ &\qquad - \mathsf{E}\Big[\mathsf{E}\big[Z\big|\, X_1, X_2, \ldots, X_{\ell-1}\big]\Big|\, X_1, X_2, \ldots, X_j\Big]\\ &= \mathsf{E}\big[Z\big|\, X_1, X_2, \ldots, X_j\big] - \mathsf{E}\big[Z\big|\, X_1, X_2, \ldots, X_j\big]\\ &\qquad (\because \text{ towering property(定理 1.51(1))}\\ &= 0\end{aligned}$$

となる. 上の式から, $\ell > j$ のとき

$$\mathsf{E}_j\big[\Delta_\ell \Delta_j\big] = \Delta_j \mathsf{E}_j[\Delta_\ell] = 0$$

がわかる. 再度, towering property(定理 1.51(1)) を用いると

$$\mathsf{E}\big[\Delta_j \Delta_\ell\big] = \mathsf{E}\Big[\mathsf{E}\big[\Delta_j \Delta_\ell \big| X_1, X_2, \ldots, X_j\big]\Big] = \mathsf{E}\Big[\mathsf{E}_j\big[\Delta_j \Delta_\ell\big]\Big] = 0 \quad (5.11)$$

がわかる. (5.11) を (5.10) に代入すると

$$\mathsf{Var}[Z] = \sum_{j=1}^{n} \mathsf{E}\big[\Delta_j^2\big]$$

を得る.

次に

$$\Delta_j = \mathsf{E}_j[Z] - \mathsf{E}_{j-1}[Z] = \mathsf{E}_j[Z] - \mathsf{E}_j\big[\mathsf{E}^{(j)}[Z]\big] = \mathsf{E}_j\Big[Z - \mathsf{E}^{(j)}[Z]\Big]$$

と書きかえて, Jensen の不等式を用いると

$$\Delta_j^2 = \left\{\mathsf{E}_j\Big[Z - \mathsf{E}^{(j)}[Z]\Big]\right\}^2 \le \mathsf{E}_j\Big[\big\{Z - \mathsf{E}^{(j)}[Z]\big\}^2\Big]$$

を得る. したがって

$$\begin{aligned}
\mathsf{E}\big[\Delta_j^2\big] &\le \mathsf{E}\Big[\mathsf{E}_j\big[\big\{Z - \mathsf{E}^{(j)}[Z]\big\}^2\big]\Big] \\
&= \mathsf{E}\Big[\big\{Z - \mathsf{E}^{(j)}[Z]\big\}^2\Big] && \big(\because \text{towering property(定理 1.51(1))}\big) \\
&= \mathsf{E}\Big[\mathsf{E}^{(j)}\Big[\big\{Z - \mathsf{E}^{(j)}[Z]\big\}^2\Big]\Big] && \big(\because \text{towering property(定理 1.51(1))}\big) \\
&= \mathsf{E}\Big[\mathsf{Var}^{(j)}[Z]\Big]
\end{aligned}$$

がわかる. よって, 定理は証明された. □

5.3 修正 log-Sobolev 不等式

$\big(\mathbb{U}, \mathcal{B}(\mathbb{U}), \mu\big)$ を確率空間とする. 非負値可測関数 $h : \mathbb{U} \longrightarrow [0, \infty)$ に対して, **エントロピー関数** Ent_μ を

$$\mathsf{Ent}_\mu(h) = \int \big(h \log h\big)\,\mathrm{d}\mu - \int h\,\mathrm{d}\mu \times \log\!\left(\int h\,\mathrm{d}\mu\right) \quad (5.12)$$

で定める．ただし, $0\log 0 = \lim_{x\to 0;\, x>0} x\log x = 0$ と約束する．$x\log x \geq -1/e\,(x\geq 0)$ なので, エントロピー関数は $h\in L_1(\mu)$ に対して定義されることがわかる．

いま, $m\in\mathbb{N}$ とし

$$W_2^1(\gamma_m) = \left\{h : \mathbb{R}^m \longrightarrow \mathbb{R};\ \int \left|h\right|^2 + |\nabla h|_{2,m}^2\big)\,\mathrm{d}\gamma_m\right\} \tag{5.13}$$

で定まる．ただし, γ_m は平均が $\mathbf{0}_m$, 分散業分散行列が $\mathbf{I}_m$ の m 次元多変量正規分布 γ_m は $\mathsf{N}_m(\mathbf{0}_m, \mathbf{I}_m)$ の確率測度である．

補題 5.18. $\big(\mathbb{U}, \mathcal{B}(\mathbb{U}), \mu\big)$ を確率空間とする．$h : \mathbb{U} \longrightarrow [0,\infty)$ を $h\log h \in L_1(\mu)$ なる非負値可測関数とする．ただし, $0\log 0 = 0$ と約束する．このとき, 以下が成立する．

(1) 可測関数 $g : \mathbb{U} \longrightarrow, \mathbb{R}$ に対して

$$\mathsf{Ent}_\mu(h) = \sup\left\{\int hg\,\mathrm{d}\mu;\ \int e^g\,\mathrm{d}\mu \leq 1\right\} \tag{5.14}$$

が成り立つ．

(2) 可測関数 $g : \mathbb{U} \longrightarrow, \mathbb{R}$ に対して

$$\mathsf{Ent}_\mu(h) = \inf_{g>0}\left\{\int \big\{h(\log h - \log g) - (h-g)\big\}\,\mathrm{d}\mu\right\} \tag{5.15}$$

が成り立つ．

Proof. 一般性を失わず

$$\int h\,\mathrm{d}\mu = 1$$

として, 証明すればよいことに注意する．なぜならば, 命題 1.31 から, $\int h\,\mathrm{d}\mu = 0$ のとき, $\mu(h=0)=1$ となる．$0\log 0 = 0$ に注意すると

$$\begin{aligned}\mathsf{Ent}_\mu(h) &= \int (h\log h)\,\mathrm{d}\mu - \int h\,\mathrm{d}\mu \times \left(\int h\,\mathrm{d}\mu\right)\\ &= \int 0\log 0\,\mathrm{d}\mu - 0\log\underbrace{\left(\int h\,\mathrm{d}\mu\right)}_{=0}\\ &= 0\end{aligned}$$

となる．一方

$$\sup\left\{\int \underbrace{h}_{=0} g\,\mathrm{d}\mu;\ \int e^g\,\mathrm{d}\mu \leq 1\right\} = 0$$

と

$$\inf_{g>0}\left\{\int\{\underbrace{\{h\log h-h\log g\}}_{=0}-(h-g)\}\,\mathrm{d}\mu\right\}=\inf_{g>0}\int g\,\mathrm{d}\mu=0$$

となるので, 結果は自明となる. よって, $\int h\,\mathrm{d}\mu>0$ として証明すればよい. さらに, エントロピー関数は $c>0$ に対して, $\mathsf{Ent}_\mu(ch)=c\mathsf{Ent}_\mu(h)$ となる. 実際

$$\begin{aligned}\mathsf{Ent}_\mu(ch)&=\int(ch)\log(ch)\,\mathrm{d}\mu-\int ch\,\mathrm{d}\mu\times\log\left(\int ch\,\mathrm{d}\mu\right)\\&=c\int h\big(\log h+\log c\big)\,\mathrm{d}\mu-c\int h\mathrm{d}\mu\times\left\{\log\left(\int h\,\mathrm{d}\mu\right)+\log c\right\}\\&=c\int h\log h\,\mathrm{d}\mu+\log c\int h\,\mathrm{d}\mu-c\int h\mathrm{d}\mu\times\log\left(\int h\,\mathrm{d}\mu\right)\\&\qquad-\log c\int h\mathrm{d}\mu\\&=c\mathsf{Ent}_\mu(h)\end{aligned}$$

からわかる. このことから, 一般性を失うことなく, $\int h\,\mathrm{d}\mu=1$ としてよいことがわかる.

<u>(1) の証明</u>. $\int h\,\mathrm{d}\mu=1$ から

$$\mathsf{Ent}_\mu(h)=\int h\log h\,\mathrm{d}\mu$$

となる. $u,v\geq 0$ に対して, 不等式

$$uv\leq u\log u-u+e^v$$

が成立すること [5] に注意する. $\int e^g\,\mathrm{d}\mu\leq 1$ に対して

$$\begin{aligned}\int hg\,\mathrm{d}\mu&\leq\int h\log h\,\mathsf{d}\mu-\underbrace{\int h\,\mathrm{d}\mu}_{=1}+\underbrace{\int e^g\,\mathrm{d}\mu}_{le1}\\&\leq\mathsf{Ent}_\mu(h)\end{aligned}$$

となるので, 上の不等式の左辺に関して sup を取れると

$$\sup\left\{\int hg\,\mathrm{d}\mu;\int e^g\,\mathrm{d}\mu\leq 1\right\}\leq\mathsf{Ent}_\mu(h)\tag{5.16}$$

[5] Young の不等式 1.38 から $uv\leq\frac{1}{2}\big(u^2+v^2\big)$ がわかる. あとは $y\log u-u-\frac{u^2}{2}\geq 0$ と $e^v\geq\frac{v^2}{2}$ からわかる.

がわかる．一方，$g = \log h$ とおくと

$$\int e^g \,\mathrm{d}\mu = \int h \,\mathrm{d}\mu = 1$$

と

$$\int gh \,\mathrm{d}\mu = \mathsf{Ent}(h)$$

なので

$$\sup\left\{\int hg \,\mathrm{d}\mu;\ \int e^g \,\mathrm{d}\mu \le 1\right\} \ge \mathsf{Ent}_\mu(h) \tag{5.17}$$

がわかる．よって，(5.16) と (5.17) から (1) は示せた．
<u>(2) の証明</u>．正値関数上の汎関数 Φ を

$$\Psi(g) = \int\left\{h(\log h - \log g) - (h - g)\right\}\mathrm{d}\mu$$

で定める．ここで，$h \ge 0$ かつ $\int h \,\mathrm{d}\mu = 1$ なので，$h \le 1$ に注意する．すると

$$\begin{aligned}
\Psi(g) &= \int h \log h \,\mathrm{d}\mu - \underbrace{\int h \,\mathrm{d}\mu}_{=1} - \int h \log g \,\mathrm{d}\mu + \int g \,\mathrm{d}\mu \\
&= \mathrm{Ent}_\mu(h) - 1 - \int h \log g \,\mathrm{d}\mu + \int g \,\mathrm{d}\mu \\
&\ge \mathrm{Ent}_\mu(h) - 1 - \int \log g \,\mathrm{d}\mu + \int g \,\mathrm{d}\mu \\
&= \mathsf{Ent}_\mu(h) + \int \underbrace{\left\{g - \log g - 1\right\}}_{\ge 0} \mathrm{d}\mu \\
&\ge \mathsf{Ent}_\mu(h)
\end{aligned}$$

がわかる．上の不等式の左辺の sup をとると

$$\sup_{g>0} \Psi(g) \ge \mathsf{Ent}_\mu(h) \tag{5.18}$$

となる．一方，$g = 1$ と取ると

$$\Psi(g) = \int h(\log h - 1) \,\mathrm{d}\mu - \underbrace{\int h \,\mathrm{d}\mu}_{=0} + 1 = \mathsf{Ent}_\mu(h)$$

となるので

$$\sup_{g>0} \Psi(g) \ge \Psi(1) = \mathsf{Ent}_\mu(h) \tag{5.19}$$

となる．(5.18) と (5.19) を合わせると (2) が証明できた． □

補題 5.19. $\big(\mathbb{U}_j,\, \mathcal{B}(\mathbb{U}_j),\, \mu_j\big)$, $j = 1,\, 2,\, \dots\, m$ を確率空間とし

$$\mu = \mu_1 \otimes \mu_2 \otimes \cdot \otimes \mu_m$$

を直積空間 $\mathbb{U} = \mathbb{U}_1 \times \cdots \times \mathbb{U}_m$ とその上の直積 σ 加法族上の直積確率測度とする. このとき, 任意の非負値可測関数 $h : \mathbb{U} \longrightarrow \mathbb{R}$ で $h \log h \in L_1(\mu)$ なるものに対して

$$\mathsf{Ent}_\mu(h) \le \sum_{j=1}^{m} \int \mathsf{Ent}_{\mu_j}(h_j)\, \mathrm{d}\mu \tag{5.20}$$

が成り立つ. ただし, $x_j \in \mathbb{U}_j$ に対して

$$h_j(\,\cdot\,) = h(x_1,\, \dots,\, x_{j-1},\, \cdot\,, x_{j+1},\, \dots,\, x_m)$$

と定めた.

Proof. 証明は帰納法を用いる.
I. $m = 1$ **の場合**. 自明である.
II. $m - 1$ **の場合**. m を $m - 1$ に置き換えた (5.20) が成立すると仮定する.
III. m **の場合**. $\int h\, \mathrm{d}\mu = 0$ ならば, $h = 0$, $\mu -$ a.s. となるので, この場合には主張は成立するので, $\int h\, \mathrm{d}\mu > 0$ と仮定する. $\Phi(x) = x \log x$ とし

$$\mu_{-m} = \mu_1 \otimes \cdots \otimes \mu_{m-1}$$

とおく. このとき, Fubini(定理 1.42) の定理, (5.12), 帰納法の仮定から

$$\begin{aligned}
&\int \Phi(h)\, \mathrm{d}\mu \\
&\quad \le \int \bigg\{ \Phi\bigg(\int h(\,\cdot\; x_m)\, \mathrm{d}\mu_{-m} \bigg) \\
&\qquad\qquad + \sum_{j=1}^{m-1} \int \mathsf{Ent}_{\mu_j}\big(h_j(\,\cdot\; x_m)\big)\, \mathrm{d}\mu_{-m} \bigg\}\, \mathrm{d}\mu_m(x_m)
\end{aligned} \tag{5.21}$$

が成り立つこと[6] がわかる. ここで

$$g(x_m) = \log\left(\int h(\cdot\ x_m)\,\mathrm{d}\mu_{-m}\right) - \log\left(\int h(\cdot\ x_m)\,\mathrm{d}\mu\right)$$

とおく. 再度, Fubini(定理 1.42) の定理を用いると

$$\begin{aligned}
\int e^g\,\mathrm{d}\mu_m &= \int\left\{\frac{\exp\log\left(\int h(\cdot\ x_m)\,\mathrm{d}\mu_{-m}\right)}{\exp\log\left(\int h\,\mathrm{d}\mu\right)}\right\}\mathrm{d}\mu_m \\
&= \int\left\{\frac{\int h(\cdot\ x_m)\,\mathrm{d}\mu_{-m}}{\int h\,\mathrm{d}\mu}\right\}\mathrm{d}\mu_m \\
&= \frac{1}{\int h\,\mathrm{d}\mu}\int\left\{\int h(\cdot\ x_m)\,\mathrm{d}\mu_{-m}\right\}\mathrm{d}\mu_m \\
&= \frac{\int h\,\mathrm{d}\mu}{\int h\,\mathrm{d}\mu} \\
&= 1
\end{aligned}$$

となるので, Fubini(定理 1.42) の定理と $\mathsf{Ent}_{\mu_m}(h_m)$ は x_m 以外の変数に依存することに注意して, 補題 5.18(1) を用いると

$$\int \mathsf{Ent}_{\mu_m}(h_m)\,\mathrm{d}\mu \geq \int \Phi\left(\int h(\cdot, x_m)\,\mathrm{d}\mu_{-m}\right)\mathrm{d}\mu_m - \Phi\left(\int h\,\mathrm{d}\mu\right) \tag{5.22}$$

[6]実際

$$\begin{aligned}
&\int \Phi(h)\,\mathrm{d}\mu \\
&\quad= \int\left\{\int \Phi\big(h(\cdot\ x_m)\big)\,\mathrm{d}\mu_{-m}\right\}\mathrm{d}\mu_m(x_m) \\
&\quad= \int\left\{\int h(\cdot\ x_m)\,\mathrm{d}\mu_{-m}\times\log\left(\int h(\cdot\ x_m)\,\mathrm{d}\mu_{-m}\right) + \mathsf{Ent}_{\mu_{-m}}((h(\cdot\ x_m))\right\} \\
&\qquad\times\ \mathrm{d}\mu_m(x_m) \\
&\quad= \int\left\{\Phi\left(\int h(\cdot\ x_m)\,\mathrm{d}\mu_{-m}\right) + \mathsf{Ent}_{\mu_{-m}}((h(\cdot\ x_m))\right\}\mathrm{d}\mu_m(x_m) \\
&\quad\leq \int\left\{\Phi\left(\int h(\cdot\ x_m)\,\mathrm{d}\mu_{-m}\right) + \sum_{j=1}^{m-1}\int \mathsf{Ent}_{\mu_j}\big(h_j(\cdot\ x_m)\big)\,\mathrm{d}\mu_{-m}\right\}\mathrm{d}\mu_m(x_m)
\end{aligned}$$

からわかる.

が成り立つこと[7]がわかる. 最後に, (5.21) と (5.22) を合わせると

$$
\begin{aligned}
\int \Phi(h)\,\mathrm{d}\mu &\le \int \left\{ \Phi\left(\int h(\,\cdot\,, x_m)\,\mathrm{d}\mu_{-m} \right\} \mathrm{d}\mu_m + \sum_{j=1}^{m-1} \int \mathsf{Ent}_{\mu_j}\big(h_j\big)\,\mathrm{d}\mu \\
&\le \int \mathsf{Ent}_{\mu_m}(h_m)\,\mathrm{d}\mu + \Phi\left(\int h\,\mathrm{d}\mu \right) + \sum_{j=1}^{m-1} \int \mathsf{Ent}_{\mu_j}\big(h_j\big)\,\mathrm{d}\mu
\end{aligned}
$$

となる. この不等式を整理すると

$$
\begin{aligned}
\sum_{j=1}^{n} \int \mathsf{Ent}_{\mu_j}(h_j)\,\mathrm{d}\mu &\ge \int \Phi(h)\,\mathrm{d}\mu - \Phi\left(\int h\,\mathrm{d}\mu \right) \\
&= \int (h \log h)\,\mathrm{d}\mu - \int h\,\mathrm{d}\mu \log\left(\int h\,\mathrm{d}\mu \right) \\
&= \mathsf{Ent}_{\mu}(h)
\end{aligned}
$$

を得る. よって, (5.20) は証明された. □

---

[7]実際

$$
\begin{aligned}
&\int \mathsf{Ent}_{\mu_m}(h_m)\,\mathrm{d}\mu \\
&\quad = \int \left\{ \int \mathsf{Ent}_{\mu_m}(h_m)\,\mathrm{d}\mu_m \right\} \mathrm{d}\mu_{-m} \\
&\quad = \int \left\{ \mathsf{Ent}_{\mu_m}(h_m) \underbrace{\int \mathrm{d}\mu_m}_{=1} \right\} \mathrm{d}\mu_{-m} \\
&\quad = \int \mathrm{Ent}_{\mu_m}(h_m)\,\mathrm{d}\mu_{-m} \\
&\quad \ge \int \left\{ \int h_m g\,\mathrm{d}\mu_m \right\} \mathrm{d}\mu_{-m} \qquad (\because \text{補題 } 5.18(1)) \\
&\quad = \int \left[\int h(\,\cdot\,, x_m) \left\{ \log\left(\int h(\,\cdot\,, x_m)\,\mathrm{d}\mu_{-m} \right) - \log\left(\int h\,\mathrm{d}\mu \right) \right\} \mathrm{d}\mu_m \right] \mathrm{d}\mu_{-m} \\
&\quad = \int \left[\int h(\,\cdot\,, x_m) \left\{ \log\left(\int h(\,\cdot\,, x_m)\,\mathrm{d}\mu_{-m} \right) - \log\left(\int h\,\mathrm{d}\mu \right) \right\} \mathrm{d}\mu_{-m} \right] \mathrm{d}\mu_m \\
&\quad = \int \left[\log\left(\int h(\,\cdot\,, x_m)\,\mathrm{d}\mu_{-m} \right) \int h(\,\cdot\,, x_m)\,\mathrm{d}\mu_{-m} \right. \\
&\qquad\quad \left. - \log\left(\int h\,\mathrm{d}\mu \right) \int h(\,\cdot\,, x_m)\,\mathrm{d}\mu_{-m} \right] \mathrm{d}\mu_m \\
&\quad = \int \left[\log\left(\int h(\,\cdot\,, x_m)\,\mathrm{d}\mu_{-m} \right) \int h(\,\cdot\,, x_m)\,\mathrm{d}\mu_{-m} \right] \mathrm{d}\mu_m \\
&\qquad\quad - \log\left(\int h\,\mathrm{d}\mu \right) \underbrace{\int \left\{ \int h(\,\cdot\,, x_m)\,\mathrm{d}\mu_{-m} \right\} \mathrm{d}\mu_m}_{=\int h\,\mathrm{d}\mu} \\
&\quad = \int \Phi\left(\int h(\,\cdot\,, x_m)\,\mathrm{d}\mu_{-m} \right) \mathrm{d}\mu_m - \Phi\left(\int h\,\mathrm{d}\mu \right)
\end{aligned}
$$

からわかる.

命題 5.20. P を標本空間 $\mathbb{X}$ 上の確率分布とし, $X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{P}$ とする. Z を $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ の有界な統計量とする. $\{X_j'\}_{j=1}^n$ を $\{X_j\}_{j=1}^n$ の独立複製とし, $Z_j\,(j = 1, 2, \ldots, n)$ を $(X_1, \ldots, X_{j-1}, X_j', X_{j+1}, \ldots, X_n)$ の有界な統計量とする. すなわち, ある可測関数 $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ が存在して

$$Z_j = h(X_1, \ldots, X_{j-1}, X_j', X_{j+1}, \ldots, X_n)$$

と書ける.
(1) 任意の $\lambda \in \mathbb{R}$ と $j = 1, 2, \ldots, n$ に対して

$$\lambda\mathsf{E}\big[Ze^{\lambda Z}\big] - \mathsf{E}\big[e^{\lambda Z}\big]\log\mathsf{E}\big[e^{\lambda Z}\big] \le \sum_{k=1}^n \mathsf{E}\big[e^{\lambda Z}\varphi\big(-\lambda(Z - Z_k)\big)\big] \tag{5.23}$$

が成り立つ. ただし, $\varphi(x) = e^x - x - 1$ である.
(2) $a \in \mathbb{R}$ に対して, $a_+ = \max\{a, 0\}$ と書いたとき

$$\lambda\mathsf{E}\big[Ze^{\lambda Z}\big] - \mathsf{E}\big[e^{\lambda Z}\big]\log\mathsf{E}\big[e^{\lambda Z}\big] \le \sum_{j=1}^n \mathsf{E}\Big[e^{\lambda Z}\psi\big(-\lambda(Z - Z_j)_+\big)\Big] \tag{5.24}$$

が成り立つ. ただし, $\psi(x) = x(e^x - 1)$ である.

Proof. ① (5.23) の証明: $\boldsymbol{X}_{-j} = (X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_n)\,(j = 1, 2, \ldots, n)$ とし, $\mathsf{E}^{(j)}[\,\cdot\,] := \mathsf{E}\big[\,\cdot\,\big|\boldsymbol{X}_{-j}\big]$ と書くことにする. G を $\boldsymbol{X}$ の非負値有界統計量とし, $\Phi(u) = u\log u\,(u > 0)$ とする. まず, テンソル化不等式 (補題 5.19) と (5.12) を用いると

$$\begin{aligned}
\mathsf{E}\big[\Phi(G)\big] - \Phi\big(\mathsf{E}[G]\big) &= \mathsf{Ent}_\mu(G) \\
&\le \sum_{j=1}^n \int \mathsf{Ent}_{\mu_j}(G_j)\,\mathrm{d}\mu \\
&= \mathsf{E}\Bigg[\sum_{j=1}^n \bigg(\mathsf{E}^{(j)}\big[\Phi(G)\big] - \Phi\big(\mathsf{E}^{(j)}[G]\big)\bigg)\Bigg]
\end{aligned} \tag{5.25}$$

を得る. ただし, μ は $\boldsymbol{X}$ の確率測度で, μ_j と G_j は補題 5.19 の記法に準じた. つぎに, 補題 5.18(2) を用いると

$$\mathsf{E}^{(j)}\big[\Phi(G)\big] - \Phi\big(\mathsf{E}^{(j)}[G]\big) = \inf_{g>0}\mathsf{E}^{(j)}\Big[G\big(G - \log g\big) - (G - g)\Big] \tag{5.26}$$

となる. ここで, $G = e^{\lambda Z}$ と $G_j =^{\lambda Z_j}$ $(j = 1, 2, \ldots, , n)$ とおく. すると $\Phi(e^{\lambda Z}) = e^{\lambda Z}\log e^{\lambda Z} = \lambda Ze^{\lambda Z}$ となることに注意すると

$$\lambda\mathsf{E}\big[Ze^{\lambda Z}\big] - \mathsf{E}\big[e^{\lambda Z}\big]\log\mathsf{E}\big[e^{\lambda Z}\big] = \mathsf{E}\big[\Phi(G)\big] - \Phi\big(\mathsf{E}\big[e^{\lambda Z}\big]\big) \tag{5.27}$$

となる. また, G_j は $(X'_j, \boldsymbol{X}_{-j})$ に基づく統計量なので, (5.26) において, X'_j を固定して, $g = G_j$ と取ることができる. さらに, Fubini の定理を用いると

$$\mathsf{E}^{(j)}\big[\Phi(G)\big] - \Phi\big(\mathsf{E}^{(j)}[G]\big) \le \mathsf{E}^{(j)}\Big[G\big(\log G - \log G_j\big) - (G - G_j)\Big] \tag{5.28}$$

を得る. この式に $G = e^{\lambda Z}$, $G_j = e^{\lambda Z_j}$ $(j = 1, 2, \ldots, n)$ を代入すると

$$\begin{aligned}\mathsf{E}^{(j)}\big[\Phi(e^{\lambda Z})\big] - \Phi\big(\mathsf{E}^{(j)}[e^{\lambda Z}]\big) &\le \mathsf{E}^{(j)}\big[e^{\lambda Z}\big(\lambda Z - \lambda Z_j\big) - \big(e^{\lambda Z} - e^{\lambda Z_j}\big)\big] \\ &= \mathsf{E}^{(j)}\big[e^{\lambda Z}\big(\lambda(Z - Z_j) - 1 + e^{\lambda(Z_j - Z)}\big)\big] \\ &= \mathsf{E}^{(j)}\big[e^{\lambda Z}\varphi\big(-\lambda(Z - Z_j)\big)\big] \end{aligned} \tag{5.29}$$

を得る. (5.25), (5.27), (5.29) を合わせると

$$\begin{aligned}\lambda\mathsf{E}\big[Ze^{\lambda Z}\big] - \mathsf{E}\big[e^{\lambda Z}\big]\log\mathsf{E}\big[e^{\lambda Z}\big] &= \mathsf{E}\big[\Phi(G)\big] - \Phi\big(\mathsf{E}\big[e^{\lambda Z}\big]\big) \\ &= \sum_{j=1}^{n}\mathsf{E}\Bigg[\bigg(\mathsf{E}^{(j)}\big[\Phi(G)\big] - \Phi\big(\mathsf{E}^{(j)}[G]\big)\bigg)\Bigg] \\ &\le \sum_{j=1}^{n}\mathsf{E}\Big[\mathsf{E}^{(j)}\big[e^{\lambda Z}\varphi\big(-\lambda(Z - Z_j)\big)\big]\Big] \\ &= \sum_{j=1}^{n}\mathsf{E}\Big[e^{\lambda Z}\varphi\big(-\lambda(Z - Z_j)\big)\Big]\end{aligned}$$

を得る.

② (5.24) の証明: 実数 a に対して, $a_+ = \max\{a, 0\}$, $a_1 = \max\{-a, 0\}$ とおくと $a = a_+ - a_-$ となる. $a = a_+$ ならば, $a_- = 0$ であり, $a = -a_-$ ならば, $a_+ = 0$ であることと $\varphi(0) = 0$ に注意すると

$$\varphi\big(-\lambda(Z - Z_j)\big) = \varphi\big(-\lambda(Z - Z_j)_+\big) + \varphi\big(\lambda(Z - Z_j)_-\big)$$

となる. $\boldsymbol{Z}_{-j}$ を条件付けすると Z と Z_j は同じ分布に従うので

$$\begin{aligned}\mathsf{E}^{(j)}\big[e^{\lambda Z}\varphi\big(\lambda(Z - Z_j)_-\big)\big] &= \mathsf{E}^{(j)}\big[e^{\lambda Z_j}\varphi\big(\lambda(Z_j - Z)_-\big)\big] \\ &\qquad \big(\because Z \text{ と } Z_j \text{ を入れ替えた}\big) \\ &= \mathsf{E}^{(j)}\big[e^{\lambda Z}e^{\lambda(Z_j - Z)}\varphi\big(\lambda(Z_j - Z)_-\big)\big] \\ &= \mathsf{E}^{(j)}\big[e^{\lambda Z}e^{-\lambda(Z_j - Z)_-}\varphi\big(\lambda(Z_j - Z)_-\big)\big] \\ &\qquad \big(\because (Z_j - Z)_- = 0 \text{ のときは } \varphi(0) = 0\big) \\ &= \mathsf{E}^{(j)}\big[e^{\lambda Z}e^{-\lambda(Z - Z_j)_+}\varphi\big(\lambda(Z - Z_j)_+\big)\big] \qquad (5.30) \\ &\qquad \big(\because (-a)_+ = -a_-\big)\end{aligned}$$

となる. $\psi(x) = e^x\varphi(-x) + \varphi(x)$ となること[8]となることと (5.30) から

$$
\begin{aligned}
&\mathsf{E}^{(j)}\bigl[e^{\lambda Z}\varphi\bigl(-\lambda(Z-Z_j)\bigr)\bigr] \\
&\quad = \mathsf{E}^{(j)}\bigl[e^{\lambda Z}\varphi\bigl(-\lambda(Z-Z_j)_+\bigr)\bigr] + \mathsf{E}^{(j)}\bigl[e^{\lambda Z}\varphi\bigl(-\lambda(Z-Z_j)_-\bigr)\bigr] \\
&\quad = \mathsf{E}^{(j)}\bigl[e^{\lambda Z}\psi\bigl(-\lambda(Z-Z_j)_+\bigr)\bigr] + \mathsf{E}^{(j)}\bigl[e^{\lambda Z}e^{-\lambda(Z-Z_j)_+}\varphi\bigl(\lambda(Z-Z_j)_+\bigr)\bigr] \\
&\quad = \mathsf{E}^{(j)}\bigl[e^{\lambda Z}\psi\bigl(-\lambda(Z-Z_j)_+\bigr)\bigr]
\end{aligned}
$$

がわかる. この式を (5.23) の上に代入すると (5.24) が証明できる. □

5.4 Orlicz ノルム

定義 5.21. $\Phi : [0, \infty) \longrightarrow [0, \infty)$ は狭義凸関数で $\Phi(0) = 0$ かつ $\lim_{x\to\infty}\Phi(x) = \infty$ とする[9]. 確率変数 X の Orlicz **ノルム** $\|X\|_\Phi$ を

$$
\|X\|_\Phi := \inf\left\{c > 0;\ \mathsf{E}\left[\Phi\left(\frac{|X|}{c}\right)\right] \le 1\right\} \tag{5.31}
$$

で定める. また

$$
\left\{c > 0;\ \mathsf{E}\left[\Phi\left(\frac{|X|}{c}\right)\right] \le 1\right\} = \varnothing \Leftrightarrow \|X\|_\Phi = \infty
$$

と約束する.

命題 5.22. $a \in \mathbb{R}$ を定数, X, Y を確率変数とする. 以下が成立する.
(1) $\|X\|_\Phi \ge 0$. さらに, $\|X\|_\Phi = 0 \Leftrightarrow X = 0\,\mathrm{a.s.}$
(2) $\|aX\|_\Phi = |a|\|X\|_\Phi$.
(3) $\|X+Y\|_\Phi \le \|X\|_\Phi + \|Y\|_\Phi$.

Proof. (2) の証明: 任意の実数 a に対して

$$
\begin{aligned}
\|af\|_\Phi &= \inf\left\{c > 0;\ \mathsf{E}\left[\Phi\left(\frac{|aX|}{c}\right)\right] \le 1\right\} \\
&= \inf\left\{\frac{c}{|a|} > 0;\ \mathsf{E}\left[\Phi\left(\frac{|aX|}{c/|a|}\right)\right] \le 1\right\} \\
&= \inf\left\{\frac{c}{|a|} > 0;\ \mathsf{E}\left[\Phi\left(\frac{|X|}{c}\right)\right] \le 1\right\} \\
&= |a|\ \inf\left\{c > 0;\ \mathsf{E}\left[\Phi\left(\frac{|X|}{c}\right)\right] \le 1\right\} \\
&= |a|\|X\|_\Phi
\end{aligned}
$$

[8]実際

$$
e^x\varphi(-x)+\varphi(x) = e^x\bigl\{e^{-x}+x-1\bigr\} + e^x - x - 1 = 1 + xe^x - e^x + e^x - x - 1 = x(e^x-1) = \psi(x)
$$

からわかる.

[9]したがって, Φ は狭義単調増加でもある.

からわかる.
(1) の証明: $X=0$ ならば, $\|X\|_\Phi=0$ は明らかである. $\|X\|_\Phi=0$ ならば, $X=0$, a.s. を示すために, 背理法を用いる. すなわち, $\mathsf{Pr}\bigl(|X|>a\bigr)>0$ をみたす $a>0$ の存在を仮定して矛盾を導こう. $A:=\{\omega\in\Omega;\ |X(\omega)|>a\}$ とおく. 任意の $c>0$ に対して

$$\begin{aligned}
\mathsf{E}\left[\Phi\left(\frac{|X|}{c}\right)\right] &= \mathsf{E}\left[\Phi\Bigl(\underbrace{\frac{|X|}{c}}_{\geq \frac{a}{c}}\Bigr)\mathbb{1}_A(X)\right] + \underbrace{\mathsf{E}\left[\Phi\left(\frac{|X|}{c}\right)\mathbb{1}_{A^c}(X)\right]}_{\geq 0} \\
&\geq \Phi\left(\frac{a}{c}\right)\mathsf{E}[\mathbb{1}_A(X)] \\
&= \Phi\left(\frac{a}{c}\right)\mathsf{Pr}\bigl(|X|>a\bigr) \qquad (5.32)
\end{aligned}$$

が成立する. 一方, $\|X\|_\Phi=0$ と (2) から, 任意の $c>0$ に対して

$$0=\frac{1}{|c|}\|X\|_\Phi=\|X/c\|_\Phi=\inf\left\{c>0;\ \mathsf{E}\left[\Phi\left(\frac{|X|}{c}\right)\right]\leq 1\right\}$$

がわかる. よって, $\forall c>0$ に対して

$$\mathsf{E}\left[\Phi\left(\frac{|X|}{c}\right)\right]\leq 1 \qquad (5.33)$$

が成り立つ. (5.32) と (5.33) から

$$\Phi\left(\frac{a}{c}\right)\mathsf{Pr}\bigl(|X|>a\bigr)\leq 1 \qquad (5.34)$$

が成立することがわかる. しかし, (5.34) の左辺の項は $a\to 0$ のとき, $\Phi(\infty)=\infty$ かつ $\mathsf{Pr}\bigl(|X|>a\bigr)>0$ なので, (5.34) の左辺は発散する. よって, 矛盾が生じるので, $\|X\|_\Phi=0$ ならば $X=0$, a.s. が示せた.
③ の証明: $\|X\|_\Phi=0$ または $\|Y\|_\Phi=0$ のときは, 明らかだから, $\|X\|_\Phi>$

0 かつ $\|Y\|_\Phi > 0$ を仮定して, 証明すればよい.

$$
\begin{aligned}
&\mathsf{E}\left[\Phi\left(\frac{|X+Y|}{\|X\|_\Phi + \|Y\|_\Phi}\right)\right] \\
&\quad \le \mathsf{E}\left[\Phi\left(\frac{|X|}{\|X\|_\Phi + \|Y\|_\Phi} + \frac{|Y|}{\|X\|_\Phi + \|Y\|_\Phi}\right)\right] \\
&\qquad\quad (\because \Phi \text{ の単調増加性}) \\
&\quad = \mathsf{E}\left[\Phi\left(\frac{\|X\|_\Phi}{\|X\|_\Phi + \|Y\|_\Phi}\frac{|X|}{\|X\|_\Phi} + \frac{\|Y\|_\Phi}{\|X\|_\Phi + \|U\|_\Phi}\frac{|Y|}{\|Y\|_\Phi}\right)\right] \\
&\quad \le \frac{\|X\|_\Phi}{\|X\|_\Phi + \|Y\|_\Phi}\mathsf{E}\left[\Phi\left(\frac{|X|}{\|X\|_\Phi}\right)\right] \\
&\qquad + \frac{\|Y\|_\Phi}{\|X\|_\Phi + \|Y\|_\Phi}\mathsf{E}\left[\Phi\left(\frac{|Y|}{\|Y\|_\Phi}\right)\right] \qquad (\because \Phi \text{ の凸性}) \\
&\quad \le 1 \qquad (\because (5.31))
\end{aligned}
$$

がわかる. よって

$$\|X\|_\Phi + \|Y\|_\Phi \in \left\{c > 0;\ \mathsf{E}\left[\Phi\left(\frac{X+Y|}{c}\right)\right] \le 1\right\}$$

となる. 一方, $\|\cdot\|_\Phi$ の定義 (5.31) から

$$\|X+Y\|_\Phi = \inf\left\{c > 0;\ \mathsf{E}\left[\Phi\left(\frac{|X+Y|}{c}\right)\right] \le 1\right\}$$

であることに注意すると

$$\|X+Y\|_\Phi \le \|X\|_\Phi + \|Y\|_\Phi$$

がわかる. □

補題 5.23. $\{X_n\}_{n=1}^\infty$, X を確率変数列とし, $\{|X_n|\}_{n=1}^\infty$ は単調増加列で $|X_n| \overset{\text{a.s.}}{\to} |X|\,(n \to \infty)$ とする. このとき

$$\|X_n\|_\Phi \overset{n\to\infty}{\longrightarrow} \|X\|_\Phi$$

が成立する.

Proof. $\underline{\|X\|_\Phi < \infty \text{ の場合の証明}}$: 補題の仮定と Φ の単調性から

$$\|X_n\|_\Phi \le \|X\|_\Phi\,(n = 1, 2, \ldots) \tag{5.35}$$

となること[10] がわかる．すなわち $\{\|X_n\|_\Phi\}_{n=1}^{\infty}$ は有界単調列となる．これから, ある $s\in\mathbb{R}$ が存在して, $\|X_n\|_{\ \Phi} \overset{n\to\infty}{\longrightarrow} s$ となる．(5.35) と合わせると $s \le \|X\|_\Phi$ がわかるので

$$s \ge \|X\|_\Phi \tag{5.36}$$

を示せば, $s=\|X\|_\Phi$ がわかる．$s=0$ の場合, $X_n=0\,(\text{a.s.}, n\ge 1)$ から $X=0\,(\text{a.s.})$ となり, 不等式は成立するので, $s>0$ として, 補題の主張を証明しよう．$\|X_n\|_\Phi < s$ とノルム $\|\cdot\|_\Phi$ の定義 (5.31) から

$$s \ge \|X_n\|_\Phi = \inf\left\{c>0;\, \mathsf{E}\left[\Phi\left(\frac{|X_n|}{c}\right)\right] \le 1\right\}$$

から

$$\mathsf{E}\left[\frac{|X_n|}{s}\right] \le 1$$

がわかる．ここで単調収束定理を用いると

$$1 \ge \lim_{n\to\infty}\mathsf{E}\left[\Phi\left(\frac{|X_n|}{s}\right)\right] = \mathsf{E}\left[\lim_{n\to\infty}\Phi\left(\frac{|X_n|}{s}\right)\right] = \mathsf{E}\left[\Phi\left(\frac{|X|}{s}\right)\right]$$

を得る．すなわち

$$s \in \left\{c>0;\, \mathsf{E}\left[\Phi\left(\frac{|X|}{c}\right)\right] \le 1\right\} \tag{5.37}$$

である．再度, ノルム $\|\cdot\|_\Phi$ の定義 (5.31) と (5.37) に注意すると

$$s \ge \|X\|_\Phi = \inf\left\{c>0;\, \mathsf{E}\left[\Phi\left(\frac{|X|}{c}\right)\right] \le 1\right\}$$

がわかる．よって, (5.36) が示せたので, この場合について補題の主張が確認できた．

$\|X\|_\Phi=\infty$ の場合の証明: $\|X_n\| \overset{n\to\infty}{\longrightarrow} \infty$ となることを示せばよい．数列 $\{\|X_n\|_\Phi\}_{n=1}^{\infty}$ は有界単調増加列と仮定して矛盾を導く．背理法の仮定から

$$\lim_{n\to\infty}\|X_n\|_\Phi = s < \infty$$

---

[10] $|X_n| \le |X|\,(\text{a.s.})$ とノルム $\|\cdot\|_\Phi$ の定義から $\Phi\left(\frac{|X|}{\|X_n\|_\Phi}\right) \ge 1 \ge \Phi\left(\frac{|X_n|}{\|X_n\|_\Phi}\right)$ である．一方, ノルム $\|\cdot\|_\Phi$ の定義から $\Phi\left(\frac{|X|}{\|X\|_\Phi}\right) \le 1$ なので, $\Phi\left(\frac{|X|}{\|X_n\|_\Phi}\right) \ge \Phi\left(\frac{|X|}{\|X\|_\Phi}\right)$ となる．関数 Φ の単調性から, $\frac{|X|}{\|X_n\|_\Phi} \ge \frac{|X|}{\|X\|_\Phi} \Rightarrow \|X\|_\Phi \ge \|X_n\|_\Phi$ がわかる．

とおくことができる．すると

$$\inf\left\{c > 0;\ \mathsf{E}\left[\Phi\left(\frac{|X_n|}{c}\right)\right] \leq 1\right\} = \|X_n\|_\Phi \leq s$$

から

$$1 \geq \mathsf{E}\left[\Phi\left(\frac{|X_n|}{s}\right)\right]$$

がわかる．ここで, 単調収束定理を用いると

$$\begin{aligned}1 &\geq \lim_{n\to\infty} \mathsf{E}\left[\Phi\left(\frac{|X_n|}{s}\right)\right] = \mathsf{E}\left[\lim_{n\to\infty} \Phi\left(\frac{|X_n|}{s}\right)\right] = \mathsf{E}\left[\Phi\left(\frac{\lim_{n\to\infty}|X_n|}{s}\right)\right] \\ &= \mathsf{E}\left[\Phi\left(\frac{|X|}{s}\right)\right]\end{aligned}$$

を得る．よって, ノルム $\|\cdot\|_\Phi$ の定義から

$$s \geq \|X\|_\Phi$$

となり, $\|X\|_\Phi$ は有限となる．このことは $\|X\|_\Phi = \infty$ の仮定と矛盾する．以上の議論から

$$\lim_{n\to\infty} \|X_n\|_\Phi = \infty$$

がわかる．よって, 補題は証明された． □

補題 5.24. $\{X\}_{n=1}^\infty$, X を確率変数列とする．このとき

$$\|X_n\|_\Phi \xrightarrow{n\to\infty} \|X\|_\Phi < \infty \Rightarrow X_n \xrightarrow{\mathrm{P}} X \qquad (n \to \infty)$$

となる．

Proof. 補題を証明するために

$$\|X_n\|_\Phi \xrightarrow{n\to\infty} 0 \Rightarrow X_n \xrightarrow{\mathrm{P}} 0 \tag{5.38}$$

を示せばよいことに注意する．任意の $\delta > 0$ を取る．するとある $N \in \mathbb{N}$ が存在して

$$\forall n \geq N \Rightarrow \|X_n\|_\Phi \leq \delta$$

となる．ノルム $\|\cdot\|_\Phi$ の定義 (定義 5.21) から

$$\mathsf{E}\left[\Phi\left(\frac{|X_n|}{\delta}\right)\right] \leq 1$$

がわかる. Φ は $[0,\infty)$ 上の狭義単調増加関数で $\lim_{x\to\infty}\Phi(x)=\infty$ なので, 任意の $\epsilon>0$ に対して, 十分大きな $M>0$ を取ると

$$\Phi(M)\geq\frac{1}{\epsilon}$$

とできる. $\|X_n\|_\Phi\leq\delta$ に注意すると

$$\begin{aligned}1&\geq\mathsf{E}\left[\Phi\left(\frac{|X_n|}{\delta}\right)\right]\geq\mathsf{E}\left[\mathbb{1}_{(M\delta,\infty)}(|X_n|)\Phi\left(\frac{|X_n|}{\delta}\right)\right]\\&\geq\mathsf{E}\left[\mathbb{1}_{(M\delta,\infty)}(|X_n|)\Phi\left(\frac{M\delta}{\delta}\right)\right]\geq\Phi(M)\mathsf{E}\left[\mathbb{1}_{(M\delta,\infty)}(|X_n|)\right]\\&\geq\Phi(M)\mathsf{Pr}\bigl(|X_n|>M\delta\bigr)\end{aligned}$$

から

$$\mathsf{Pr}\bigl(|X_n|>M\delta\bigr)\leq\epsilon$$

がわかる. $\delta>0$ は任意だったので, $M\delta=\epsilon$ になるように取れば

$$\mathsf{Pr}\bigl(|X_n|>\epsilon\bigr)\leq\epsilon$$

がわかる. よって, 補題は証明された. □

5.4.1 $\Phi_2(x)=e^{x^2}-1$ の場合

例 5.25. $\Phi(x)=x^p\,(p\geq1)$ のとき, L^p ノルムとなる. 他の重要な例として, 劣 Gauss 確率変数に対する $\Phi_2(x)=e^{x^2}-1$ と劣指数確率変数に対する $\Phi_1(x)=e^x-1$ がある.

例 5.26. $1\leq p<\infty$ とし

$$\Phi_p(x)=e^{x^p}-1\qquad(x\geq0)\tag{5.39}$$

とおく. すると

$$\Phi_p^{-1}(x)=\sqrt[p]{\log(1+x)}$$

となる. したがって

$$\Phi_p^{\vee}(x)=\sup_{y>0}\bigl\{xy-\Phi(y)\bigr\}=x\Phi^{-1}(x)-x=x\sqrt[p]{\log(1+x)}-x$$

となる. □

補題 5.27. (5.39) において, $p = 2$ を考える. 確率変数 X に対して, $\|X\|_{\Phi_2} < \infty$ としたとき, $q = 1, 2, \ldots$ に対して

$$\left\{\mathsf{E}\left[|X|^{2q}\right]\right\}^{1/(2q)} \le \left(q!\right)^{1/(2q)} \big\|X\big\|_{\Phi_2} \tag{5.40}$$

が成り立つ. 特に

$$\mathsf{E}\left[|X|\right] \le \|X\|_{\Phi_2}$$

が成り立つ.

Proof. ① $q \ge 1$ の場合の証明: まず

$$\Phi_2(x) = e^{x^2} - 1 = x^2 + \frac{x^4}{2} + \cdots > \frac{x^{2q}}{q!} \qquad (x > 0,\ q = 1, 2, \ldots)$$

に注意する. $c > 0$ に対して

$$\frac{1}{q!}\mathsf{E}\left[\left(\frac{|X|}{c}\right)^{2q}\right] < \mathsf{E}\left[\Phi_2\left(\frac{|X|}{c}\right)\right] \tag{5.41}$$

となる. $c = \dfrac{\left\{\mathsf{E}[|X|^{2q}]\right\}^{1/(2q)}}{q!}$ を (5.41) に代入すると

$$\mathsf{E}\left[\Phi_2\left(\frac{|X|}{\left\{\mathsf{E}[|X|^{2q}]\right\}^{1/(2q)}/q!}\right)\right] > 1$$

となる. したがって

$$\frac{\left\{\mathsf{E}[|X|^{2q}]\right\}^{1/(2q)}}{q!} \notin \left\{c > 0;\ \mathsf{E}\left[\Phi\left(\frac{|X|}{c}\right)\right] \le 1\right\}$$

となる. この不等式とノルム $\|\cdot\|_{\Phi_2}$ の定義 (5.21) から

$$\frac{\left\{\mathsf{E}[|X|^{2q}]\right\}^{1/(2q)}}{(q!)^{1/(2q)}} \le \|X\|_{\Phi_2} = \inf\left\{c > 0;\ \mathsf{E}\left[\Phi_2\left(\frac{|X|}{c}\right)\right] \le 1\right\}$$

がわかる.
② $\mathsf{E}[|X|] \le \|X\|_{\Phi_2}$ の証明: Cauchy-Schwarz の不等式と ① の $q = 1$ の場合の不等式から

$$\mathsf{E}[|X|] \le \sqrt{\mathsf{E}[X^2]}\sqrt{\mathsf{E}[1]} = \sqrt{\mathsf{E}[X^2]} \le \|X\|_{\Phi_2}$$

がわかる. □

命題 5.28. X を確率変数とし, $p \geq 1$ とする. つぎの 2 つは同値である.
(1) $\|X\|_{\Phi_p} < \infty$.
(2) ある C, $K > 0$ が存在して

$$\mathsf{Pr}\big(|X| > t\big) \leq Ke^{-Ct^p} \qquad (\forall t > 0)$$

となる.

Proof. (1) ⇒ (2) の証明: $0 < \|X\|_{\Phi_p} < \infty$ として証明すればよい. $t > 0$ に対して

$$\begin{aligned}\mathsf{Pr}\big(|X| > t\big) &= \mathsf{Pr}\left(\Phi_p\left(\frac{|X|}{\|X\|_\Phi}\right) > \Phi_p\left(\frac{t}{\|X\|_\Phi}\right)\right) \\ &\leq \left\{\Phi_p\left(\frac{t}{\|X\|_\Phi}\right)\right\}^{-1} \mathsf{E}\left[\Phi_p\left(\frac{|X|}{\|X\|_\Phi}\right)\right] \qquad (\because \text{Markov の不等式 (命題 1.33)}) \\ &\leq \left\{\Phi_p\left(\frac{t}{\|X\|_\Phi}\right)\right\}^{-1} \qquad (\because (5.31))\end{aligned}$$

となる. したがって

$$\mathsf{Pr}\big(|X| > t\big) \leq \min\left[1,\ \left\{\Phi_p\left(\frac{t}{\|X\|_\Phi}\right)\right\}^{-1}\right] \quad \leq 2\exp\left(-\frac{t^p}{\|X\|_{\Phi_p}^p}\right)$$

がわかる. 最後の不等号は, 任意の $u > 0$ に対して

$$\min\left[1,\ \frac{1}{\Phi_p(u)}\right] = \min\left[1,\ \frac{1}{e^{u^p} - 1}\right] \leq \frac{2}{e^{u^p}}$$

が成り立つこと [11] からわかる.

[11] 任意の $0 < x < 1$ のとき

$$\min\left[1,\ \frac{1}{x-1}\right] = \frac{1}{x-1} \leq \frac{2}{x}$$

である. 一方, $1 \leq x < 2$ のとき

$$\min\left[1,\ \frac{1}{x-1}\right] = 1 \leq \frac{2}{x}$$

となる. $x > 2$ のとき

$$\min\left[1,\ \frac{1}{x-1}\right] = \frac{1}{x-1} \leq \frac{2}{x}$$

となる. 以上から

$$\min\left[1,\ \frac{1}{x-1}\right] \leq \frac{2}{x}$$

がわかる. あとは, 上の不等式に $x = e^{-u^p}$ を代入すればよい.

(2) $\Rightarrow$ (1) の証明:

$$
\begin{aligned}
\mathsf{E}\left[\Phi_p\left(\frac{|X|}{c}\right)\right] &= \mathsf{E}\left[\exp\left(\frac{|X|^p}{c^p}\right) - 1\right] \\
&= \mathsf{E}\left[\int_0^{\|X\|_{\Phi_p}} \frac{e^{s/c^p}}{c^p}\,\mathrm{d}s\right] \\
&= \int_0^\infty \mathsf{Pr}\big(|X| > s^{1/p}\big)\frac{e^{s/c^p}}{c^p}\,\mathrm{d}s \quad (\text{この等号の証明がわからない!}) \\
&\leq \int_0^\infty Ke^{-Cs}\frac{e^{s/c^p}}{c^p}\,\mathrm{d}s \\
&= \frac{K}{c^p}\int_0^\infty \exp\left\{-\left(C - \frac{1}{c^p}\right)s\right\}\mathrm{d}s \\
&= \frac{K}{c^p}\frac{1}{C - \frac{1}{c^p}}
\end{aligned}
$$

となる. この不等式から

$$
c \geq \left(\frac{1+K}{C}\right)^{1/p}
$$

のとき

$$
\mathsf{E}\left[\Phi_p\left(\frac{|X|}{c}\right)\right] \leq \frac{K}{c^p}\frac{1}{C - \frac{1}{c^p}} = \frac{K}{c^pC - 1} \leq \frac{K}{\frac{1+K}{C}C - 1} = 1
$$

となる. したがって

$$
\left[\left(\frac{1+K}{C}\right)^{1/p}, \infty\right] \subset \left\{c > 0;\ \mathsf{E}\left[\Phi_p\left(\frac{|X|}{c}\right)\right] \leq 1\right\}
$$

なので, ノルム $\|\cdot\|_{\Phi_2}$ の定義に注意すると

$$
\left(\frac{1+K}{C}\right)^{1/p} \geq \inf\left\{c > 0;\ \mathsf{E}\left[\Phi_p\left(\frac{|X|}{c}\right)\right] \leq 1\right\} = \|X\|_{\Phi_p}
$$

がわかる. □

つぎに Φ_2 に関する最大不等式を導出する.

定理 5.29. $\{X_n\}_{n=1}^\infty$ を確率変数列 (独立でなくともよい) とする. このとき

$$
\left\|\sup_n \frac{|X_n|}{\Phi_2^{-1}(n)}\right\|_{\Phi_2} \leq C\sup_n \|X_n\|_{\Phi_2}
$$

が成立する. $C > 0$ は Φ_2 にのみ依存する定数である.

Proof. $\Phi_2(x) = e^{x^2} - 1\,(x \geq 0)$ から $\Phi_2^{-1}(x) = \sqrt{\log(1+x)}\,(x \geq 0)$ である. $X_n / \sup_n \|X_n\|_{\Phi_2}$ を考えることで, 一般性を失わず

$$\|X_n\|_{\Phi_2} \leq \sup_n \|X_n\|_{\Phi_2} = 1$$

と仮定してよい. すると

$$\begin{aligned}
\|X_n\|_{\Phi_2} \leq 1 \ &\Leftrightarrow\ \mathsf{E}\left[\Phi\left(\frac{|X_n|}{1}\right)\right] \leq 1 \\
&\Leftrightarrow\ \mathbb{E}\big[e^{X_n^2} - 1\big] \leq 1 \\
&\Leftrightarrow\ \mathbb{E}\big[e^{X_n^2}\big] \leq 2
\end{aligned} \tag{5.42}$$

がわかる. $t \geq \dfrac{3}{2}$ とする. $n \geq 18$ に対して

$$\frac{1}{\log n} + \frac{1}{\log t} \leq \frac{1}{\log 18} + \frac{1}{\log(3/2)} \leq 3$$

となるので

$$3(\log n)(\log t) \geq \log n + \log t = \log(nt)$$

である. したがって

$$\begin{aligned}
\mathsf{Pr}\left(\exp\left\{\sup_{n\geq 18}\left(\frac{|X_n|}{\sqrt{6\log n}}\right)^2\right\} > t\right) &= \mathsf{Pr}\left(\sup_{n\geq 18}\frac{|X_n|}{\sqrt{6\log n}} > \sqrt{\log t}\right) \\
&= \mathsf{Pr}\left(\sup_{n\geq 18}\frac{|X_n|}{\sqrt{6(\log n)(\log t)}} > 1\right) \\
&\leq \sum_{n=18}^{\infty}\mathsf{Pr}\left(|X_n| > \sqrt{6(\log n)(\log t)}\right) \\
&\leq \sum_{n=18}^{\infty}\mathsf{Pr}\left(e^{X_n^2} > e^{6(\log n)(\log t)}\right) \leq \sum_{n=18}^{\infty}\frac{\mathsf{E}[e^{X_n^2}]}{e^{6(\log n)(\log t)}} \\
&\leq \sum_{n=18}^{\infty}\frac{2}{e^{6(\log n)(\log t)}} \leq \sum_{n=18}^{\infty}\frac{2}{e^{2\log n + 2\log t}} \leq \sum_{n=18}^{\infty}\frac{2}{n^2 t^2} \leq \frac{1}{4t^2}
\end{aligned}$$

がわかる. よって

$$\begin{aligned}
\mathsf{E}\left[\exp\left\{\sup_{n\geq 18}\left(\frac{|X_n|}{\sqrt{6\log n}}\right)^2\right\}\right] &= \int_0^\infty \mathsf{Pr}\left(\exp\left\{\sup_{n\geq 18}\left(\frac{|X_n|}{\sqrt{6\log n}}\right)^2\right\} > t\right)\mathrm{d}t \\
&= \int_0^\infty \frac{1}{4t^2}\,\mathrm{d}t = \int_0^{3/2}\frac{1}{4t^2}\,\mathrm{d}t + \int_{3/2}^\infty \frac{1}{4t^2}\,\mathrm{d}t \\
&\leq \frac{3}{2} + \int_{3/2}^\infty \frac{1}{4t^2} < 2
\end{aligned}$$

がわかる. よって, (5.42) に注意すると

$$\left\| \sup_{n\geq 18} \frac{|X_n|}{\sqrt{6\log n}} \right\|_{\Phi_2} \leq 1 = \sup_n \|X_n\|_{\Phi_2}$$

がわかる. よって

$$\begin{aligned}\left\| \sup_n \frac{|X_n|}{\Phi_2^{-1}(n)} \right\|_{\Phi_2} &= \left\| \sup_n \frac{|X_n|}{\sqrt{\log(1+n)}} \right\|_{\Phi_2} \leq \sqrt{6}\left\| \sup_n \frac{|X_n|}{\sqrt{6\log n}} \right\|_{\Phi_2} \\ &= 18\sqrt{6}\sup_n \|X_n\|_{\Phi_2}\end{aligned}$$

がわかる. よって, 定理は証明された. □

系 5.30. $N \in \mathbb{N}$ とし, $\{X_n\}_{n=1}^N$ を確率変数列 (独立でなくともよい) とする. このとき

$$\left\| \max_{1\leq n\leq N} |X_n| \right\|_{\Phi_2} \leq C\Phi_2^{-1}(N) \max_{1\leq n\leq N} \|X_n\|_{\Phi_2}$$

が成り立つ. さらに, $q = 1, 2, \ldots$ に対して

$$\left(\mathsf{E}\left[\max_{1\leq n\leq N} |X_n|^{2q}\right]\right)^{1/(2q)} \leq C'\sqrt{\log(1+N)} \max_{1\leq n\leq N} \|X_n\|_{\Phi_2}$$

が成り立つ. ただし, $C > 0$ と C' は Φ_2 にのみ依存する定数である.

Proof. $X_n = X_N\ (n \geq N+1)$ とおくと定理 5.29 から

$$\begin{aligned}\left\| \max_{1\leq n\leq N} |X_n| \right\|_{\Phi_2} &\leq \left\| \Phi_2^{-1}(N) \max_{1\leq n\leq N} \frac{|X_n|}{\Phi_2^{-1}(n)} \right\|_{\Phi_2} \\ &= \Phi_2^{-1}(N) \left\| \max_{1\leq n\leq N} \frac{|X_n|}{\Phi_2^{-1}(n)} \right\|_{\Phi_2} \\ &= \Phi_2^{-1}(N) \left\| \sup_n \frac{|X_n|}{\Phi_2^{-1}(n)} \right\|_{\Phi_2} \\ &\leq C\Phi_2^{-1}(N) \sup_n \|X_n\|_{\Phi_2} \\ &= C\Phi_2^{-1}(N) \max_{1\leq n\leq N} \|X_n\|_{\Phi_2}\end{aligned}$$

を得る.

次に, (5.40) から

$$\begin{aligned}\left(\mathsf{E}\left[\max_{1\leq n\leq N} |X_n|^{2q}\right]\right)^{2q} &\leq (q!)\left\| \max_{1\leq n\leq N} |X_n| \right\|_{\Phi_2} \\ &\leq (q!)C\Phi_2^{-1}(N) \max_{1\leq n\leq N} \|X_n\|_{\Phi_2} \\ &\leq C'\sqrt{\log(1+N)} \max_{1\leq n\leq N} \|X_n\|_{\Phi_2}\end{aligned}$$

がわかる. □

節 ?? および節 ?? と同様の評価式を導出できるのでは? (2025/01/15)

注意 5.31. 劣 Gauss 性の仮定のもとではより精緻な評価が得られる. $X_1, X_2, \ldots, X_N$ を定数 $\nu\,(\nu > 0)$ の劣 Gauss 列とする. すなわち, $\lambda > 0$ に対して

$$\mathsf{E}\left[e^{\lambda X_j}\right] \leq \exp\left(\frac{\nu\lambda^2}{2}\right) \qquad (j = 1, 2, \ldots, N) \tag{5.43}$$

をみたす. このとき, $\forall t > 0$ と $\lambda > 0$ に対して

$$\begin{aligned}
\mathsf{Pr}\left(X_j > t\right) &= \mathsf{Pr}\left(e^{\lambda X_j} \geq e^{\lambda t}\right) \\
&\leq e^{-\lambda t}\mathsf{E}\left[e^{\lambda X_j}\right] \qquad (\because \text{Markov の不等式 (命題 1.33)}) \\
&\leq e^{-\lambda t}\exp\left(\frac{\nu\lambda^2}{2}\right) \qquad (\because (5.43)) \\
&= \exp\left\{\frac{\nu}{2}\left(\lambda^2 - \frac{2\lambda t}{\nu}\right)\right\} \\
&= \exp\left\{\frac{\nu}{2}\left(\lambda - \frac{t}{\nu}\right)^2 - \frac{t^2}{2\nu}\right\}
\end{aligned}$$

から

$$\mathsf{Pr}\left(X_j > t\right) \leq \exp\left(-\frac{t^2}{2\nu}\right)$$

を得る. よって

$$\begin{aligned}
\mathsf{Pr}\left(\max_{1\leq j\leq N} X_j > t\right) &= \mathsf{Pr}\left(\bigcup_{j=1}^{N}\{X_j > t\}\right) \\
&\leq \sum_{j=1}^{N}\mathsf{Pr}(X_j > t) \qquad (\because \text{union bound}) \\
&\leq N\exp\left(-\frac{t^2}{2\nu}\right) = \exp\left(-\frac{t^2}{2\nu} + \log N\right)
\end{aligned}$$

を得る. □

5.5 基本的な情報不等式