

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

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

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

$$Z := Z_n = f_n(X_1, X_2, \dots, X_n), \quad f_n: \mathbb{R}^n \rightarrow \mathbb{R}$$

としたとき

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

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 $E[X^2] < \infty$ なる確率変数とする。このとき

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

が成立する。

Proof. Markov の不等式 (命題 1.29) を用いると

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

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

$$[0, \lambda_0) \ni \lambda \mapsto \Phi_Z(\lambda) := \log E[\exp(\lambda Z)]$$

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

$$\Phi_Z^\vee(t) := \sup_{\lambda \geq 0} \{\lambda t - \Phi_Z(\lambda)\} \quad (t > 0)$$

で定める.

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

$$\Pr(Z \geq t) \leq \exp\left(-\Phi_Z^\vee(t)\right) \quad (t > 0)$$

が成り立つ.

Proof. 原点を含むある近傍に含まれる $\lambda \geq 0$ を取る. Markov の不等式 (命題 1.29) を用いると

$$\begin{aligned} \Pr(Z \geq t) &= \Pr(\exp(\lambda Z) \geq \exp(\lambda t)) \\ &\leq \frac{1}{\exp(\lambda t)} E[\exp(\lambda Z)] \\ &= \exp\left\{-\{\lambda t - \log E[\exp(\lambda Z)]\}\right\} \\ &= \exp\left\{-\{\lambda t - \Phi_Z(\lambda)\}\right\} \end{aligned}$$

となる. $\lambda \geq 0$ は任意だったので

$$\begin{aligned} \Pr(Z \geq t) &\leq \exp\left\{-\sup_{\lambda \geq 0} \{\lambda t - \Phi_Z(\lambda)\}\right\} \\ &= \exp\left\{-\Phi_Z^\vee(t)\right\} \end{aligned}$$

を得る. □

5.1.2 例

例 5.3 (正規分布). $Z \sim N(0, \sigma^2)$ ($\sigma > 0$) とする. このとき, $\lambda \geq 0$ に対して

$$\begin{aligned} E[e^{\lambda Z}] &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\lambda z - \frac{z^2}{2\sigma^2}\right\} dz \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(z - \sigma^2\lambda)^2 + \frac{\sigma^2\lambda^2}{2}\right\} dz \\ &= \exp\left(\frac{\sigma^2\lambda^2}{2}\right) \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(z - \sigma^2\lambda)^2\right\} dz}_{=1} \\ &= \exp\left(\frac{\sigma^2\lambda^2}{2}\right) \end{aligned}$$

となる. よって

$$\Phi_Z(\lambda) = \exp\left(\frac{\sigma^2\lambda^2}{2}\right)$$

である. これより

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

となる. よって, 系 4.5 を用いると

$$\Pr(Z \geq t) \leq \exp\left(-\frac{t^2}{2\sigma^2}\right) \quad (t > 0)$$

を得る. □

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

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

となる. これより

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

となる. ここで

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

とおくと

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

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

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

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

$$\Phi_X(\lambda) \leq \frac{\lambda^2 \nu}{2} \quad (\lambda \in \mathbb{R})$$

であることをいう. これを $X \sim \text{SubG}(\nu)$ 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

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

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

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

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

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 1.29) を用いる. $X \sim \text{subG}(\sigma^2)$ であることより

$$\Pr(X \geq t) = \Pr(e^{sX} \geq e^{st}) \leq \frac{\mathbb{E}[e^{sX}]}{e^{st}} \leq \exp\left(\frac{\sigma^2}{2}s^2 - st\right) \quad (5.1)$$

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 (5.1) は任意の $s > 0$ で成立しているので

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

を得る. □

補題 5.7. 確率変数 X は

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

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

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

となる. 特に

$$(\mathbb{E}[|X|^k])^{1/k} \leq \sigma e^{1/e} \sqrt{k}, \quad (k \geq 2)$$

と

$$\mathbb{E}[|X|] \leq \sigma \sqrt{2\pi}$$

である.

Proof. 命題 1.28 と補題 5.6 より, $t > 0$ に対して

$$\begin{aligned} \mathbb{E}[|X|^k] &= \int_0^\infty \Pr(|X| > t^{1/k}) dt \quad (\because \text{命題 1.28}) \\ &\leq 2 \int_0^\infty \exp\left(-\frac{t^{2/k}}{2\sigma^2}\right) dt \quad (\because \text{補題 5.6}) \\ &= (2\sigma^2)^{k/2} k \int_0^\infty e^{-u} u^{(k/2)-1} du \quad \left(\because u = \frac{t^{2/k}}{2\sigma^2} \text{ と変換}\right) \\ &= (2\sigma^2)^{k/2} k \Gamma\left(\frac{k}{2}\right) \end{aligned}$$

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

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

を用いると

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

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 $(\Omega, \mathcal{F}, \Pr)$ 上の確率変数で以下をみたすものとする.

- (1) $E[X] = 0$.
- (2) $\forall t > 0$ に対して

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

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

$$E[e^{sX}] \leq e^{4\sigma^2 s^2}$$

となる.

Proof. 指数関数に対する Taylor 展開を用いる.

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

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.

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

となる. 最後の不等式はほとんど明らかだが, どのように評価してよいか思いつかない. □

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

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

が成り立つ.

Proof. $\Phi_X(\lambda) \leq \frac{\nu\lambda^2}{2(1-c\lambda)}$ 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.

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

となる¹. ただし

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

である. □

補題 5.11. $0 < b \leq \infty$ とする. $\Phi : [0, b) \rightarrow [0, \infty)$ 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

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

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

$$\{\Phi^\vee\}^{-1}(y) = \inf_{\lambda \in (0, b)} \left[\frac{y + \Phi(\lambda)}{\lambda} \right] \quad (5.2)$$

となる.

1

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

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.

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 を含む非正值関数の最大なので, $\Phi^\vee(0) = 0$ がわかる.

② Φ^\vee の凸性の証明:

③ Φ^\vee の狭義単調増加性の証明:

④ (5.2) の証明:

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

とおく. 任意の $t \geq 0$ に対して

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

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

$$\{\Phi^\vee\}^{\leftarrow}(y) = \sup\{t \geq 0; \Phi^\vee(t) \leq y\} = u$$

となる. さらに, Φ^\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, \dots, Z_N は確率変数列で, $\forall \lambda \in (0, b)$ に対して

$$\Phi_{Z_j}(\lambda) \leq \Phi(\lambda) \quad (j = \{1, 2, \dots, N\})$$

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

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

が成り立つ.

Proof. 任意の $\lambda \in (0, b)$ に対して

$$\exp \left(\lambda \mathbb{E} \left[\max_{j=1, 2, \dots, N} Z_j \right] \right) \leq \sum_{j=1}^n \mathbb{E} [\exp(\lambda Z_j)] \leq N \exp(\Phi(\lambda))$$

となる. この不等式の辺々に対数をとると

$$\lambda \mathbb{E} \left[\max_{j=1,2,\dots,N} Z_j \right] \leq \log N + \Phi(\lambda)$$

を得る. $\lambda \in (0, b)$ は任意だったことと (5.2) に注意すると

$$\mathbb{E} \left[\max_{j=1,2,\dots,N} Z_j \right] \leq \inf_{\lambda \in (0,b)} \left[\frac{\log N + \Phi(\lambda)}{\lambda} \right] = \{\Phi^\vee\}^{-1}(\log N)$$

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 $\mathbb{E}[X] = 0$ 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

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

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

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

$$\begin{aligned} \dot{\Phi}_X(s) &:= \frac{d}{ds} \Phi_X(s) = \frac{\mathbb{E}[X e^{sX}]}{\mathbb{E}[e^{sX}]}; \\ \ddot{\Phi}_X(s) &:= \frac{d^2}{ds^2} \Phi_X(s) = \frac{\mathbb{E}[X^2 e^{sX}]}{\mathbb{E}[e^{sX}]} - \left\{ \frac{\mathbb{E}[X e^{sX}]}{\mathbb{E}[e^{sX}]} \right\}^2 \end{aligned} \quad (5.3)$$

となる. しかし, ほとんど確実に $X \in [a, b]$ なので

$$\text{Var}[X] = \text{Var}\left[X - \frac{a+b}{2}\right] \leq \mathbb{E}\left[\left(X - \frac{a+b}{2}\right)^2\right] \leq \frac{(b-a)^2}{4}$$

となる. このこととキュムラント関数の性質²から

$$\ddot{\Phi}_X(s) \leq \frac{(b-a)^2}{4}$$

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

$$\begin{aligned} \Phi_X(s) &= \Phi_X(s) - \Phi_X(0) = \int_0^s \dot{\Phi}_X(u) du = \int_0^s \left\{ \dot{\Phi}_X(u) - \dot{\Phi}_X(0) \right\} du \\ &= \int_0^s \left\{ \int_0^u \ddot{\Phi}_X(r) dr \right\} du = \int_0^s \int_0^u \ddot{\Phi}_X(r) dr du \\ &\leq \frac{(b-a)^2}{4} \int_0^s \int_0^u dr du = \frac{(b-a)^2}{4} \int_0^s u du \\ &= \frac{(b-a)^2}{8} s^2 \end{aligned}$$

を得る. □

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

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

としたとき, $\forall t > 0$ に対して

$$\begin{aligned} \Pr(\bar{X} - E[\bar{X}] \geq t) &\leq \exp\left(-\frac{2n^2 t^2}{\sum_{j=1}^n (b_j - a_j)^2}\right), \\ \Pr(\bar{X} - E[\bar{X}] \leq -t) &\leq \exp\left(-\frac{2n^2 t^2}{\sum_{j=1}^n (b_j - a_j)^2}\right) \end{aligned}$$

を得る.

² X の p.d.f./p.m.f. を p としたとき, $q_s(x) = e^{sx - \kappa(s)} p(x)$ は指数型分布族となる. ただし, $\kappa(s) = \int_a^b e^{-sx} p(x) dx$ (連続型の場合) である. この分布に従う確率変数を \tilde{X} と記す. すると (5.3) と指数型分布族の性質から

$$\ddot{\Phi}_X(s) = \ddot{\kappa}(x) = \text{Var}[\tilde{X}]$$

となる. 一方, $\Pr(a \leq \tilde{X} \leq b) = 1$ 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

$$\text{Var}[\tilde{X}] \leq \text{Var}\left[\tilde{X} - \frac{a+b}{2}\right] \leq \frac{(b-a)^2}{4}$$

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

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

$$E[e^{sY_j}] \leq \exp\left(\frac{s^2}{8}(b_j - a_j)^2\right) \quad (5.4)$$

を得る. よって

$$\begin{aligned} \Pr(\bar{X} - E[\bar{X}] \geq t) &= \Pr\left(\sum_{j=1}^n Y_j \geq nt\right) = \Pr\left(\exp\left(s \sum_{j=1}^n Y_j\right) \geq e^{nst}\right) \\ &\leq E\left[\exp\left(s \sum_{j=1}^n Y_j\right)\right] e^{-nst} \\ &\quad (\because \text{Markov の不等式 (命題 1.29)}) \\ &= \prod_{j=1}^n E[\exp(sY_j)] e^{-nst} \\ &\leq \prod_{j=1}^n \exp\left(\frac{s^2}{8}(b_j - a_j)^2\right) e^{-nst} \quad (\because (5.4)) \\ &= \exp\left(\frac{s^2}{8} \sum_{j=1}^n (b_j - a_j)^2 - nst\right) \\ &= \exp\left\{\frac{\sum_{j=1}^n (b_j - a_j)^2}{8} \left(s - \frac{4nt}{\sum_{j=1}^n (b_j - a_j)^2}\right)^2 - \frac{2n^2t^2}{\sum_{j=1}^n (b_j - a_j)^2}\right\} \end{aligned}$$

を得る. ここで

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

とおくと

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

を得る. □

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

$$\Pr(\bar{X} \geq p + t) \leq \exp(-2nt^2)$$

を得る.

たとえば, $p = \frac{1}{2}$ とすると

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

となる. また

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

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

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, \dots, X_n は独立な確率変数列で, $X_j \leq b$ ($j = 1, 2, \dots, n$) とする.

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

とおく. このとき

$$\Phi_S(\lambda) = \log \mathbb{E}[e^{\lambda S}] \leq \frac{v}{b^2} \phi(b\lambda) \quad (\lambda \geq 0) \quad (5.5)$$

が成り立つ. さらに, $\forall t \geq 0$ に対して

$$\Pr(S \geq t) \leq \exp\left\{-\frac{v}{b^2} h\left(\frac{tb}{v}\right)\right\} \quad (5.6)$$

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

Proof. ① (5.5) の証明: X_j/b を X_j とすることで $b = 1$ とできる. 関数 $u \mapsto \frac{\phi(u)}{u^2}$ は \mathbb{R} 上の減少関数である. $X_j \leq 1$ ($j = 1, 2, \dots, n$) 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

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

がわかる. さらに, 上の不等式の両辺に対数をとると

$$\Phi_{X_j}(\lambda) = \log \mathbb{E}[e^{\lambda X_j}] \leq \log \left(1 + \lambda \mathbb{E}[X_j] + \mathbb{E}[X_j^2] \phi(\lambda) \right)$$

を得る. このとき

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

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

② (5.6) の証明: $\lambda \geq 0$ に対して

$$\begin{aligned} \Pr(S \geq t) &= \Pr(e^{\lambda S} \geq e^{\lambda t}) \leq e^{-\lambda t} \mathbb{E}[e^{\lambda S}] = \exp \left\{ - \left(\lambda t - \Phi_S(\lambda) \right) \right\} \\ &\leq \exp \left\{ - \left(\lambda t - \frac{v}{b^2} \phi(b\lambda) \right) \right\} \end{aligned}$$

となる. 最後の不等号は (5.5) からわかる. $\lambda (\geq 0)$ は任意だったので

$$\Pr(S \geq t) \leq \exp \left\{ - \sup_{\lambda \geq 0} \left(\lambda t - \frac{v}{b^2} \phi(b\lambda) \right) \right\}$$

となる. いま

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

とおく. すると

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

となるので

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

となるので

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

がわかる. □

5.2 分散不等式

5.2.1 Efron-Stein の不等式

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

$$\text{Var}[Z] = \sum_{j=1}^{\infty} \text{Var}[X_j] \tag{5.7}$$

が成立する.

いま, $f: \mathbb{R}^n \rightarrow \mathbb{R}$ を可測関数とし, $Z = f(X_1, X_2, \dots, X_n)$ とする. Efron-Stein の不等式は, (5.7) に対応する $\text{Var}[Z]$ 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, \dots, X_n を独立な確率変数列とし, $Z = f(X_1, X_2, \dots, X_n)$ とする. ただし, $f: \mathbb{R}^n \rightarrow \mathbb{R}$ を可測関数である. さらに, $E[Z^2] < \infty$ を仮定する. このとき

$$\text{Var}[Z] \leq \sum_{j=1}^n E[\text{Var}^{(j)}[Z]]$$

が成り立つ. ただし, $j = 1, 2, \dots, n$ に対して

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

である.

Proof. $E_0[\cdot] = E[\cdot]$ とし, $j = 1, 2, \dots, n$ に対して

$$E_j[\cdot] = E[\cdot | X_1, X_2, \dots, X_j]$$

とし

$$\Delta_j = E_j[Z] - E_{j-1}[Z]$$

とおく. すると

$$\begin{aligned} Z - E[Z] &= \underbrace{(E_n[Z] - E_{j-1}[Z])}_{=Z} + (E_{j-1}[Z] - E_{j-2}[Z]) + \dots \\ &\quad + (E_1[Z] - \underbrace{E_0[Z]}_{=E[Z]}) \\ &= \sum_{j=1}^n \Delta_j \end{aligned}$$

と書ける. これより

$$\text{Var}[Z] = E\left[\left(\sum_{j=1}^n \Delta_j\right)^2\right] = \sum_{j=1}^n E[\Delta_j^2] + 2 \sum_{\ell > j} E[\Delta_j \Delta_\ell] \quad (5.8)$$

となる. $\ell > j$ のとき

$$\begin{aligned} E_j[\Delta_\ell] &= E[E_\ell[Z] - E_{\ell-1}[Z] | X_1, X_2, \dots, X_j] \\ &= E[E_\ell[Z] | X_1, X_2, \dots, X_j] - E[E_{\ell-1}[Z] | X_1, X_2, \dots, X_j] \\ &= E\left[E[Z | X_1, X_2, \dots, X_\ell] \Big| X_1, X_2, \dots, X_j\right] \\ &\quad - E\left[E[Z | X_1, X_2, \dots, X_{\ell-1}] \Big| X_1, X_2, \dots, X_j\right] \\ &= E[Z | X_1, X_2, \dots, X_j] - E[Z | X_1, X_2, \dots, X_j] \\ &\quad (\because \text{towering property}) \\ &= 0 \end{aligned}$$

となる. 上の式から, $\ell > j$ のとき

$$E_j[\Delta_\ell \Delta_j] = \Delta_j E_j[\Delta_\ell] = 0$$

がわかる. 再度, tower property を用いると

$$E[\Delta_j \Delta_\ell] = E\left[E[\Delta_j \Delta_\ell | X_1, X_2, \dots, X_j]\right] = E\left[E_j[\Delta_j \Delta_\ell]\right] = 0 \quad (5.9)$$

がわかる. (5.9) を (5.8) に代入すると

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

を得る.

次に

$$\Delta_j = E_j[Z] - E_{j-1}[Z] = E_j[Z] - E_j[E^{(j)}[Z]] = E_j\left[Z - E^{(j)}[Z]\right]$$

と書きかえて, Jensen の不等式を用いると

$$\Delta_j^2 = \left\{E_j\left[Z - E^{(j)}[Z]\right]\right\}^2 \leq E_j\left[\left\{Z - E^{(j)}[Z]\right\}^2\right]$$

を得る. したがって

$$\begin{aligned} E[\Delta_j^2] &\leq E\left[E_j\left[\left\{Z - E^{(j)}[Z]\right\}^2\right]\right] \\ &= E\left[\left\{Z - E^{(j)}[Z]\right\}^2\right] \quad (\because \text{towering property}) \\ &= E\left[E^{(j)}\left[\left\{Z - E^{(j)}[Z]\right\}^2\right]\right] \quad (\because \text{towering property}) \\ &= E\left[\text{Var}^{(j)}[Z]\right] \end{aligned}$$

がわかる. よって, 定理は証明された. □

5.3 Orlicz ノルム

定義 5.18. $\Phi : [0, \infty) \rightarrow [0, \infty)$ は狭義凸関数で $\Phi(0) = 0$ かつ $\lim_{x \rightarrow \infty} \Phi(x) = \infty$ とする³. 確率変数 X の Orlicz ノルム $\|X\|_\Phi$ を

$$\|X\|_\Phi := \inf \left\{ c > 0; E\left[\Phi\left(\frac{|X|}{c}\right)\right] \leq 1 \right\} \quad (5.10)$$

³したがって, Φ は狭義単調増加でもある.

で定める. また

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

と約束する.

命題 5.19. $a \in \mathbb{R}$ を定数, X, Y を確率変数とする. 以下が成立する.

- (1) $\|X\|_{\Phi} \geq 0$. さらに, $\|X\|_{\Phi} = 0 \Leftrightarrow X = 0$ a.s..
- (2) $\|aX\|_{\Phi} = |a| \|X\|_{\Phi}$.
- (3) $\|X + Y\|_{\Phi} \leq \|X\|_{\Phi} + \|Y\|_{\Phi}$.

Proof. (2) の証明: 任意の実数 a に対して

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

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, $\Pr(|X| > a) > 0$ をみたく $a > 0$ 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

$$\begin{aligned} \mathbb{E} \left[\Phi \left(\frac{|X|}{c} \right) \right] &= \mathbb{E} \left[\underbrace{\Phi \left(\frac{|X|}{c} \right) \mathbb{1}_A(X)}_{\geq \frac{a}{c}} \right] + \underbrace{\mathbb{E} \left[\Phi \left(\frac{|X|}{c} \right) \mathbb{1}_{A^c}(X) \right]}_{\geq 0} \\ &\geq \Phi \left(\frac{a}{c} \right) \mathbb{E}[\mathbb{1}_A(X)] \\ &= \Phi \left(\frac{a}{c} \right) \Pr(|X| > a) \end{aligned} \tag{5.11}$$

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

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

がわかる. よって, $\forall c > 0$ に対して

$$\mathbb{E} \left[\Phi \left(\frac{|X|}{c} \right) \right] \leq 1 \quad (5.12)$$

が成り立つ. (5.11) と (5.12) から

$$\Phi \left(\frac{a}{c} \right) \Pr(|X| > a) \leq 1 \quad (5.13)$$

が成立することがわかる. しかし, (5.13) の左辺の項は $a \rightarrow 0$ のとき, $\Phi(\infty) = \infty$ かつ $\Pr(|X| > a) > 0$ なので, (5.13) 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.

$$\begin{aligned} & \mathbb{E} \left[\Phi \left(\frac{|X+Y|}{\|X\|_{\Phi} + \|Y\|_{\Phi}} \right) \right] \\ & \leq \mathbb{E} \left[\Phi \left(\frac{|X|}{\|X\|_{\Phi} + \|Y\|_{\Phi}} + \frac{|Y|}{\|X\|_{\Phi} + \|Y\|_{\Phi}} \right) \right] \\ & \quad (\because \Phi \text{ の単調増加性}) \\ & = \mathbb{E} \left[\Phi \left(\frac{\|X\|_{\Phi}}{\|X\|_{\Phi} + \|Y\|_{\Phi}} \frac{|X|}{\|X\|_{\Phi}} + \frac{\|Y\|_{\Phi}}{\|X\|_{\Phi} + \|Y\|_{\Phi}} \frac{|Y|}{\|Y\|_{\Phi}} \right) \right] \\ & \leq \frac{\|X\|_{\Phi}}{\|X\|_{\Phi} + \|Y\|_{\Phi}} \mathbb{E} \left[\Phi \left(\frac{|X|}{\|X\|_{\Phi}} \right) \right] \\ & \quad + \frac{\|Y\|_{\Phi}}{\|X\|_{\Phi} + \|Y\|_{\Phi}} \mathbb{E} \left[\Phi \left(\frac{|Y|}{\|Y\|_{\Phi}} \right) \right] \quad (\because \Phi \text{ の凸性}) \\ & \leq 1 \quad (\because (5.10)) \end{aligned}$$

がわかる. よって

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

となる. 一方, $\|\cdot\|_{\Phi}$ の定義 (5.10) から

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

であることに注意すると

$$\|X+Y\|_{\Phi} \leq \|X\|_{\Phi} + \|Y\|_{\Phi}$$

がわかる. □

補題 5.20. $\{X_n\}_{n=1}^\infty$, X を確率変数列とし, $\{|X_n|\}_{n=1}^\infty$ は単調増加列で $|X_n| \xrightarrow{\text{a.s.}} |X| (n \rightarrow \infty)$ とする. このとき

$$\|X_n\|_\Phi \xrightarrow{n \rightarrow \infty} \|X\|_\Phi$$

が成立する.

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

$$\|X_n\|_\Phi \leq \|X\|_\Phi (n = 1, 2, \dots) \quad (5.14)$$

なので, $\{\|X_n\|_\Phi\}_{n=1}^\infty$ は有界単調列となる. したがって, ある $s \in \mathbb{R}$ が存在して, $\|X_n\|_\Phi \xrightarrow{n \rightarrow \infty} s$ となる. (5.14) と合わせると $s \leq \|X\|_\Phi$ がわかるので

$$s \geq \|X\|_\Phi \quad (5.15)$$

を示せば, $s = \|X\|_\Phi$ がわかる. $s = 0$ の場合, $X_n = 0$ (a.s., $n \geq 1$) から $X = 0$ (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 (5.10) から

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

から

$$\mathbb{E} \left[\frac{|X_n|}{s} \right] \leq 1$$

がわかる. ここで単調収束定理を用いると

$$1 \geq \lim_{n \rightarrow \infty} \mathbb{E} \left[\Phi \left(\frac{|X_n|}{s} \right) \right] = \mathbb{E} \left[\lim_{n \rightarrow \infty} \Phi \left(\frac{|X_n|}{s} \right) \right] = \mathbb{E} \left[\Phi \left(\frac{|X|}{s} \right) \right]$$

を得る. 再度, $\|\cdot\|_\Phi$ の定義 (5.10) に注意すると

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

がわかる. よって, (5.15) 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: $\|X_n\|_\Phi \xrightarrow{n \rightarrow \infty} \infty$ 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

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

とおくことができる. すると

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

から

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

がわかる. ここで, 単調収束定理を用いると

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

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

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

がわかる. よって, 補題は証明された. \square

補題 5.21. $\{X\}_{n=1}^{\infty}$, X を確率変数列とする. このとき

$$\|X_n\|_{\Phi} \xrightarrow{n \rightarrow \infty} \|X\|_{\Phi} < \infty \Rightarrow X_n \xrightarrow{P} X$$

となる.

Proof. 補題を証明するために

$$\|X_n\|_{\Phi} \xrightarrow{n \rightarrow \infty} 0 \Rightarrow X_n \xrightarrow{P} 0 \quad (5.16)$$

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 5.18) から

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

がわかる. Φ は $[0, \infty)$ 上の狭義単調増加関数で $\lim_{x \rightarrow \infty} \Phi(x) = \infty$ 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

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

から

$$\Pr(|X_n| > M\delta) \leq \epsilon$$

がわかる. δ を $M\delta = \epsilon$ になるように取ればは

$$\Pr(|X_n| > \epsilon) \leq \epsilon$$

がわかる. よって, 補題は証明された. □

5.3.1 $\Phi_2(x) = e^{x^2} - 1$ の場合

例 5.22. $\Phi(x) = x^p$ ($p \geq 1$) 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 5.23. $1 \leq p < \infty$ とし

$$\Phi_p(x) = e^{x^p} - 1 \quad (x \geq 0) \tag{5.17}$$

とおく. すると

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

となる. したがって

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

となる. □

補題 5.24. (5.17) において, $p = 2$ を考える. 確率変数 X に対して, $\|X\|_{\Phi_p} < \infty$ としたとき, $p = 1, 2, \dots$ に対して

$$\left(\mathbb{E}[|X|^{2p}] \right)^{1/(2p)} \leq (p!)^{1/(2p)} \|X\|_{\Phi_p} \quad (5.18)$$

が成り立つ. 特に

$$\mathbb{E}[|X|] \leq \|X\|_{\Phi_2}$$

が成り立つ.

Proof. ① $p \geq 1$ の場合の証明: まず

$$\Phi_2(x) = e^{x^2} - 1 = x^2 + \frac{x^4}{2} + \dots > \frac{x^{2p}}{p!} \quad (x > 0, p = 1, 2, \dots)$$

に注意する. $c > 0$ に対して

$$\frac{1}{p!} \mathbb{E} \left[\left(\frac{|X|}{c} \right)^{2p} \right] < \mathbb{E} \left[\Phi_2 \left(\frac{|X|}{c} \right) \right] \quad (5.19)$$

となる. $c = \frac{\mathbb{E}[|X|^{2p}]^{1/(2p)}}{p!}$ を (5.19) に代入すると

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

となる. この不等式とノルム $\|\cdot\|_{\Phi_2}$ の定義 (5.18) から

$$\frac{\mathbb{E}[|X|^{2p}]^{1/(2p)}}{(p!)^{1/(2p)}} \geq \|X\|_{\Phi_2} = \inf \left\{ c > 0; \mathbb{E} \left[\Phi_2 \left(\frac{|X|}{c} \right) \right] \leq 1 \right\}$$

がわかる.

② $\mathbb{E}[|X|] \leq \|X\|_{\Phi_2}$ の場合の証明: Cauchy-Schwarz の不等式と ① の $p = 1$ の場合の不等式から

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

がわかる. □

命題 5.25. 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

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

となる.

Proof. (1) \Rightarrow (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

$$\begin{aligned} \Pr(|X| > t) &= \Pr\left(\Phi_p\left(\frac{|X|}{\|X\|_{\Phi}}\right) > \Phi_p\left(\frac{t}{\|X\|_{\Phi}}\right)\right) \\ &\leq \left\{\Phi_p\left(\frac{t}{\|X\|_{\Phi}}\right)\right\}^{-1} \mathbb{E}\left[\Phi_p\left(\frac{|X|}{\|X\|_{\Phi}}\right)\right] \quad (\because \text{Markov の不等式 (命題 1.29)}) \\ &\leq \left\{\Phi_p\left(\frac{t}{\|X\|_{\Phi}}\right)\right\}^{-1} \quad (\because (5.10)) \end{aligned}$$

となる. したがって

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

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⁴からわかる.

⁴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:

$$\begin{aligned} \mathbb{E} \left[\Phi_p \left(\frac{|X|}{c} \right) \right] &= \mathbb{E} \left[\exp \left(\frac{|X|^p}{c^p} \right) - 1 \right] = \mathbb{E} \left[\int_0^{\|X\|_{\Phi_p}} \frac{e^{s/c^p}}{c^p} ds \right] \\ &= \int_0^\infty \Pr(|X| > s^{1/p}) \frac{e^{s/c^p}}{c^p} ds \\ &\leq \int_0^\infty K e^{-Cs} \frac{e^{s/c^p}}{c^p} ds \\ &= \frac{K}{c^p} \int_0^\infty \exp \left\{ - \left(C - \frac{1}{c^p} \right) s \right\} ds \\ &= \frac{K}{c^p} \frac{1}{C - \frac{1}{c^p}} \end{aligned}$$

となる.

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

のとき

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

となる. したがって

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

なので

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

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

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

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. $X_n / \sup_n \|X\|_{\Phi_2}$ 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

$$\begin{aligned} \|X_n\|_{\Phi_2} \leq 1 &\Leftrightarrow \mathbb{E}\left[\Phi\left(\frac{|X_n|}{1}\right)\right] \leq 1 \\ &\Leftrightarrow \mathbb{E}[e^{X_n^2} - 1] \leq 1 \\ &\Leftrightarrow \mathbb{E}[e^{X_n^2}] \leq 2 \end{aligned} \quad (5.20)$$

がわかる. $t \geq \frac{3}{2}$ 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

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

がわかる. よって

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

がわかる. よって, (5.20) 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. \square

系 5.27. $N \in \mathbb{N}$ 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, $p = 1, 2, \dots$ に対して

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

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 5.26 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, (5.18) から

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

がわかる. \square

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 5.28. 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, \dots, 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

$$\mathbb{E}[e^{\lambda X_j}] \leq \exp\left(\frac{\nu \lambda^2}{2}\right) \quad (j = 1, 2, \dots, N) \quad (5.21)$$

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

$$\begin{aligned} \Pr(X_j > t) &= \Pr(e^{\lambda X_j} \geq e^{\lambda t}) \\ &\leq e^{-\lambda t} \mathbb{E}[e^{\lambda X_j}] \quad (\because \text{Markov の不等式 (命題 1.29)}) \\ &\leq e^{-\lambda t} \exp\left(\frac{\nu \lambda^2}{2}\right) \quad (\because (5.21)) \\ &= \exp\left\{\frac{\nu}{2}\left(\lambda^2 - \frac{2\lambda t}{\nu}\right)\right\} \\ &= \exp\left\{\frac{\nu}{2}\left(\lambda - \frac{t}{\nu}\right)^2 - \frac{t^2}{2\nu}\right\} \end{aligned}$$

から

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

を得る. よって

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

を得る. □