統計數學特別講義第二 (2025/05/22)

本日の内容

確率集中不等式

- Chernoff 限界
- 劣Gaussと劣Gamma確率変数列
- Hoeffdingの不等式
- Bernettの不等式
- 分散不等式

目標 … 独立な確率変数列の期待値からの偏差

を確率的に評価したい.

確率変数 Z の不等式

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

ただし、$X_1, \ldots, X_n$ は i.i.d. 確率変数列で
(独立でよい)

$g: \mathbb{R} \to \mathbb{R}$ はある関数。

$$Z := f_n(X_1, \ldots, X_n), \quad f_n: \mathbb{R}^n \longrightarrow \mathbb{R}.$$

のとき、 $\Pr(Z - E[Z] \geq t) \leq g_n(t)$

5.1 Chernoff 限界 ⇒ 凸関数の双対性

5.1.1 発想

系5.1 (Chebyshevの不等式) Xをr.v.とし, $E[X^2]<\infty$ とする. このとき

$$\Pr(|X-E[X]|\ge t)\le \frac{\mathrm{Var}[X]}{t^2} \qquad (t>0)$$

↑自明な上限「1」は省略

が成立.

<u>証明</u>　Yを非負値確率変数で，$E[Y] < \infty$ としたとき

$$E[Y \geq a] \leq \frac{E[Y]}{a} \qquad (a > 0)$$

を思い出す．　← Markovの不等式

すると，$t > 0$ に対し

$$\Pr(|X - E[X]| \geq t) = \Pr(\underbrace{|X - E[X]|^2}_{=:Y} \geq t^2)$$

$$\leq \frac{E[(X - E[X])^2]}{t^2} = \frac{\mathrm{Var}[X]}{t^2}.$$

□

Chebyshevの不等式では、変換

$$x \longmapsto x^2$$

を用いた。この変換を

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

に変えてみる。ただし、$\lambda > 0$ は任意の変数。

期待値が存在すると?

$\lambda_0 > 0$ とし、Z を確率変数とする。このとき、関数

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

λ の関数

を Z の Cramér-Chernoff 変換という

∵) Hölderの不等式(定理1.35(1)). $p, q>0$ と $\frac{1}{p}+\frac{1}{q}=1$ とする.

$E[|X|^p]<\infty$, $E[|Y|^q]<\infty$ のとき, $E[|XY|]<\infty$ で

$$E[|XY|] \leq \{E[|X|^p]\}^{1/p} \{E[|Y|^q]\}^{1/q}$$

を思い出そう.

$^\forall \lambda \in [0, \lambda_0)$ に対して, $E[\exp(\lambda Z)]<\infty$ と仮定する.

このとき, $^\forall \lambda_1, \lambda_2 \in [0, \lambda_0)$ と $0 \leq a \leq 1$ に対して.

$$\Psi_Z(a\lambda_1+(1-a)\lambda_2) \leq a\Psi_Z(\lambda_1)+(1-a)\Psi_Z(\lambda_2)$$

を示せばよい.

実際

$$E[|X|^p]^{1/p} \to \left\{E\left[\left\{(\exp(\lambda_1 Z))^a\right\}^{\frac{1}{a}}\right]\right\}^a = \{E[\exp(\lambda_1 Z)]\}^a$$

$$\Phi_Z(a\lambda_1 + (1-a)\lambda_2) = \log E\left[\exp\{(a\lambda_1 + (1-a)\lambda_2) Z\}\right]$$

$$= \log E\Big[\underbrace{\{\exp(\lambda_1 Z)\}^{\overset{=\frac{1}{p}}{a}}}_{=X}\ \underbrace{\{\exp(\lambda_2 Z)\}^{\overset{=\frac{1}{q}}{1-a}}}_{=Y}\Big]$$

$$\le \log\left[\{E[\exp(\lambda_1 Z)]\}^a \{E[\exp(\lambda_2 Z)]\}^{1-a}\right]$$

$$= a\Phi_Z(\lambda_1) + (1-a)\Phi_Z(\lambda_2)$$

$\square$

次に、Φ_Z の Fenchel-Legendre 変換を

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

で定める。

注　$\displaystyle\lim_{\lambda\to\infty} \frac{\Phi_Z(\lambda)}{\lambda} = +\infty$ ならば $\Phi_Z^\vee(t) < \infty \quad (t > 0)$.

$\Phi^\vee$ を凸関数 Φ の双対関数という。

$$\Phi \xrightarrow{\text{Fenchel}} \Phi^\vee \xrightarrow{\text{Fenchel}} (\Phi^\vee)^\vee = \Phi.$$

系5.2 ← Markovの不等式の系

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

$$\Pr(Z \geq t) \leq \exp\left(-\psi_Z^*(t)\right) \qquad (t > 0)$$

が成立する。

<u>証明</u> 原点を含むある近傍に含まれる$\lambda \geq 0$を考える。

Markovの不等式を用いて

非負値

$$\Pr(Z \geq t) = \Pr\left(\boxed{e^{\lambda Z}} \geq e^{\lambda t}\right)$$

$$\leq \frac{1}{e^{\lambda t}} E[e^{\lambda Z}]$$

$$= \exp\left[- \left\{ \lambda t - \log E[e^{\lambda Z}] \right\} \right]$$

$$= \exp\left[- \left\{ \lambda t - \Phi_Z(\lambda) \right\} \right].$$

よって

$$\Pr(Z \geq t) \leq \exp\left[- \sup_{\lambda \geq 0} \left\{ \lambda t - \Phi_Z(\lambda) \right\} \right]$$

← $\Phi_Z(\lambda) = \infty$ のことも入れてよい

⤳ $-\Phi_Z(\lambda) = -\infty$

$$= \exp\left[- \Phi_Z^{\vee}(t) \right].$$

$\square$

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

$$E[e^{\lambda Z}] = \exp\left(\frac{\sigma^2 \lambda^2}{2}\right).$$

よって

$$\Psi_Z(\lambda) = \log E[e^{\lambda Z}] = \frac{\sigma^2 \lambda^2}{2}.$$

これより

$$\Psi_Z^{\vee}(t) = \sup_{\lambda \geq 0} \left\{ \lambda t - \frac{\sigma^2 \lambda^2}{2} \right\} = \exp\left(\frac{t^2}{2\sigma^2}\right).$$

系4.5より

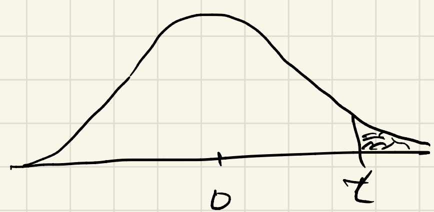
$$\Pr(Z \geq t) \leq \exp\left(-\frac{t^2}{2\sigma^2}\right) \qquad (t>0)$$

である. □

簡単のために, $\sigma=1$とする.

$$\Pr(Z \geq t) = \int_t^\infty \frac{1}{\sqrt{2\pi}} e^{-s^2/2}\, ds \leq \exp\left(-\frac{t^2}{2}\right)$$

と主張している.



5.1.3 劣Gauss と 劣Gamma 確率変数 I

定義5.5 (1) r.v. X は定数 $\nu > 0$ の劣Gauss であるとは

$$\Phi_X(\lambda) = E[e^{\lambda X}] \leq \frac{\lambda^2 \nu}{2} \qquad (\lambda \in \mathbb{R})$$

であることをいう。これを $X \sim \mathrm{subG}(\nu)$ と記す。

(2) r.v. X は定数 $\nu > 0$ と $c > 0$ の右から の劣Gamma であるとは

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

であるときをいう。$X \sim \Gamma_+(\nu, c)$ と記す。

さらに $-X \in \Gamma_+(\nu, c)$ のとき, $X \in \Gamma_-(\nu, c)$ とかく

$\Gamma(\nu, c) := \Gamma_+(\nu, c) \cap \Gamma_-(\nu, c)$ とする.

注意　$X \sim N(0, \sigma^2)$ の記号「$\sim$」とは異なる意味で使っている. 記号の乱用.

$$X \sim subG(\sigma^2) \Longleftrightarrow \text{裾の確率 } \Pr(|Z| > t) \le 2\exp\left(-\frac{t^2}{2\sigma^2}\right)$$

ココが甘くなる.

正規分布のスピード

定義5.6 $X \sim \mathrm{subG}(\sigma^2)$ とする．ただし，$0 < \sigma < \infty$ である．

このとき，

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

証明 $\lambda > 0$ とする．

$$\Pr(X \geq t) = \Pr(e^{\lambda X} \geq e^{\lambda t}) \leq \frac{E[e^{\lambda X}]}{e^{\lambda t}} \overset{\text{劣 Gauss}}{\leq} \frac{\exp\left(\frac{\lambda^2 \sigma^2}{2}\right)}{e^{\lambda t}}$$

$$= \exp\left(\frac{\sigma^2}{2}\lambda^2 - \lambda t\right).$$

よって，$\lambda > 0$ は任意であったので

$$\Pr(X \geq t) \leq \inf_{\lambda > 0} \exp\left(\frac{\sigma^2}{2}\lambda^2 - \lambda t\right) = \exp\left(-\frac{t^2}{2\sigma^2}\right).$$

$\lambda > 0$ に対して

$$\Pr(X \leq -t) = \Pr(-X \geq t) = \Pr(e^{-\lambda X} \geq e^{\lambda t})$$

$$\leq \frac{E[e^{-\lambda X}]}{e^{\lambda t}} \leq \frac{1}{e^{\lambda t}} \exp\left(\frac{(-\lambda)^2 \sigma^2}{2}\right) \leq \exp\left(-\frac{t^2}{2\sigma^2}\right)$$

最後に

$$\Pr(|X| \geq t) = \Pr\left(\{X \geq t\} \cup \{X \leq -t\}\right)$$

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

補題5.7 r.v. X は

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

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

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

ただし、$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x}\, dx \quad (a > 0)$.

証明　まず，非負値 r.v. Y に対し

$$E[Y] = \int_0^\infty \Pr(Y > t)\,dt \quad \longleftarrow \text{命題 1.28}$$

だった。$t > 0$ に対して

$$E[\underbrace{|X|^k}_{=Y}] = \int_0^\infty \Pr(|X|^k > t)\,dt$$

$$= \int_0^\infty \Pr(|X| > t^{1/k})\,dt$$

$$\leq 2\int_0^\infty \exp\left(-\frac{t^{2/k}}{2a^2}\right)dt \quad \because) \text{補題 5.6}$$

$$=: (※)$$

ここで

$$u = \frac{t^{2/k}}{2\sigma^2} \iff 2\sigma^2 u = t^{2/k} \iff (2\sigma^2)^{\frac{k}{2}} u^{\frac{k}{2}} = t$$

$$\curvearrowright \quad dt = (2\sigma^2)^{\frac{k}{2}} \frac{k}{2} u^{\frac{k}{2}-1} \, du$$

$$(※) = (2\sigma^2)^{\frac{k}{2}} k \int_0^{\infty} e^{-u} u^{\frac{k}{2}-1} \, du$$

$$= (2\sigma^2)^{\frac{k}{2}} k \, \Gamma\left(\frac{k}{2}\right). \qquad \square$$

補題5.8 以下を仮定.

(1) $E[X]=0$;

(2) $^\forall t>0$ に対して

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

このとき, $^\forall s>0$ に対して

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

証明

$$E[e^{sX}] \leq 1 + \sum_{k=1}^{\infty} \frac{s^k E[|X|^k]}{k!}$$

$$\leq 1 + \sum_{k=1}^{\infty} \frac{(2\sigma^2 s^2)^{\frac{k}{2}} k \Gamma\left(\frac{k}{2}\right)}{k!}$$

$$\vdots$$

$$\leq e^{4\sigma^2 s^2}$$

$\square$

5.1.4 Hoeffdingの不等式

補題 5.12 $E[X]=0$ かつ $\Pr(a \leq X \leq b)=1$ とする.

このとき, $^\forall \lambda > 0$ に対して

$$E[e^{\lambda X}] \leq \exp\left(\frac{\lambda^2}{8}(b-a)^2\right).$$

したがって, $X \sim \mathrm{subG}\left(\frac{(b-a)^2}{4}\right)$.

証明　$\Psi_X(s) = \log E[e^{sX}]$ $(s \in \mathbb{R})$ とする.

このとき

$$\dot{\Psi}_X(s) = \frac{E[Xe^{sX}]}{E[e^{sX}]} \; ; \; \ddot{\Psi}_X(s) = \frac{E[X^2 e^{sX}]}{E[e^{sX}]} - \left\{ \frac{E[Xe^{sX}]}{E[e^{sX}]} \right\}^2$$

となる. しかし、$\Pr(a \leq X \leq b) = 1$ なので.

平行移動する

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

$$\leq \frac{(b-a)^2}{4}.$$

一方, $\tilde{X}$ の p.d.f. を $p(x)$ としたことを

$$g_\lambda(x) = p(x)\exp\{\lambda x - \psi(\lambda)\};\quad \psi(\lambda) = \log\int e^{\lambda x} p(x)\,dx$$

$$= \log E[e^{\lambda X}].$$

に~~　~~指数型分布族になる.

$$\ddot{\Psi}_X(\lambda) = \mathrm{Var}[\tilde{X}] = \ddot{\psi}(\lambda).$$

また, $\Pr(a \le \tilde{X} \le b) = 1$ なので.

$$\mathrm{Var}[\tilde{X}] = \mathrm{Var}\left[\tilde{X} - \frac{a+b}{2}\right]$$

$$\le E\left[\left(\tilde{X} - \frac{a+b}{2}\right)^2\right]$$

$$\le \frac{(b-a)^2}{4}.$$

よって

$$\Phi_X(s) = \int_0^s \int_0^u \ddot{\Phi}_X(r)\, dr\, du$$

$$\leq \frac{(b-a)^2}{8} s^2$$

$\square$

$$E[e^{st(x-\tilde{x})} \mid x, \tilde{x}] \leq E\left[\exp\left(\frac{s^2(x-\tilde{x})}{2}\right) \mid x, \tilde{x}\right]$$

ここで、$|x-\tilde{x}| \leq b-a$ に注意すると

$$E[e^{st(x-\tilde{x})} \mid x, \tilde{x}] \leq \exp\left(\frac{s^2(b-a)}{2}\right)。$$

命題5.13 (Hoeffdingの不等式) $X_1, \ldots, X_n$ を i.i.d. r.v's で $\Pr(X_j \in [a_j, b_j]) = 1$ $(j = 1, \ldots, n,\ a_j < b_j)$ とする.
$\overline{X} = n^{-1} \sum_{j=1}^{n} X_j$ とおいたとき, $\forall t > 0$ に対して

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

を得る.

証明　$Y_j := X_j - E[X_j]$ $(j=1,\ldots,n)$ とおくと

$$\Pr(Y_j \in [a_j - E[X_j], b_j - E[X_j]]) = 1 \text{ かつ } E[Y_j]=0$$

となる。Y_j に補題5.12を適用すると，$s \in \mathbb{R}$ に対して

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

よって

$$\Pr(\bar{X} - E[\bar{X}] \geq t)$$

$$= \Pr\left(\sum_{j=1}^{n} Y_j \geq nt\right)$$

$$= \Pr\left(\exp\left(s\sum_{j=1}^{n} Y_j\right) \geq e^{ns\tau}\right)$$

$$\leq e^{-ns\tau}\, E\left[\exp\left(s\sum_{j=1}^{n} Y_j\right)\right] \quad \because \text{Markov}$$

$$= e^{-ns\tau}\prod_{j=1}^{n} E\left[e^{sY_j}\right] \quad \because)\ \text{独立性}$$

$$\leq e^{-ns\tau}\prod_{j=1}^{n}\exp\left(\frac{s^2}{8}(b_j-a_j)^2\right) \quad \because)\ \text{補題 5.12}$$

$$= \exp\left[\underline{\frac{s^2}{8}\sum_{j=1}^{n}(b_j-a_j)^2 - ns\tau}\right)$$

sの2次関数をsで最小化

より

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

例　$X_1, \ldots, X_n$ は i.i.d で

$$\Pr(X_1 = 1) = p, \quad \Pr(X_1 = 0) = 1 - p, \quad 0 < p < 1$$

とする. $\bar{X} = n^{-1} \sum_{j=1}^{n} X_j$ とおく. $n\bar{X}$ は $\mathrm{Bino}(n, p)$ である.

命題 5.13 より

$$\Pr(\bar{X} - E[\bar{X}] \geq t) \leq e^{-2nt^2}$$ を得る.

$$\| $$

$$\Pr(\bar{X} \geq p + t)$$

Chernoff-Okamoto の不等式.

$\square$

5.3 Orlicz ノルム

$\forall x, y > 0$ と $0 < a < 1$ に対し

$\Phi(ax + (1-a)y) < a\Phi(x) + (1-a)\Phi(y)$

定義 5.16 $\Phi : [0, \infty) \to [0, \infty)$ は狭義凸関数で $\Phi(0) = 0$ とする. 確率変数 X の Orlicz ノルム $\|X\|_\Phi$ を

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

で定める. 特に

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

と約束する.

命題 5.17 X, Y を r.v's で a は定数とする.

(1) $\|X\|_\Phi \geq 0$. さらに, $\|X\|_\Phi = 0 \Leftrightarrow Pr(X=0)=1$. ($\to X=0,\ a.s.$)

(2) $\|aX\|_\Phi = |a|\,\|X\|_\Phi$.

(3) $\|X+Y\|_\Phi \leq \|X\|_\Phi + \|Y\|_\Phi$.

(3) の証明 $\|X\|_\Phi > 0$ かつ $\|Y\|_\Phi > 0$ の場合を示せばよい.

まず

$$E\left[\Phi\left(\frac{|X+Y|}{\|X\|_\Phi + \|Y\|_\Phi}\right)\right]$$

$$\leq E\left[\Phi\left(\frac{|X|}{\|X\|_\Phi + \|Y\|_\Phi} + \frac{|Y|}{\|X\|_\Phi + \|Y\|_\Phi}\right)\right]$$

$\because$ $|X+Y| \leq |X| + |Y|$ と Φ の単調増加性

$$= E\left[\Phi\left(\underbrace{\frac{\|X\|_\Phi}{\|X\|_\Phi + \|Y\|_\Phi}}_{=a} \cdot \frac{|X|}{\|X\|_\Phi} + \underbrace{\frac{\|Y\|_\Phi}{\|X\|_\Phi + \|Y\|_\Phi}}_{=1-a} \cdot \frac{|Y|}{\|Y\|_\Phi}\right)\right]$$ 33

$$\leq \frac{\|X\|_\Phi}{\|X\|_\Phi + \|Y\|_\Phi} \underbrace{E\left[\Phi\left(\frac{|X|}{\|X\|_\Phi}\right)\right]}_{\leq 1 \ \because \ ノルムの定義 (5.9)} + \frac{\|Y\|_\Phi}{\|X\|_\Phi + \|Y\|_\Phi} \underbrace{E\left[\Phi\left(\frac{|Y|}{\|Y\|_\Phi}\right)\right]}_{\leq 1}$$

$$= 1$$

がわかる。よって (5.9) より

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

一方

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

d-i), $\|X+Y\|_{\Phi} \leq \|X\|_{\Phi} + \|Y\|_{\Phi}$. $\square$

補題 5.18 $\{X_n\}_{n=1}^{\infty}$ は単調増加な r.v.'s とする.

このとき

$$|X_n| \xrightarrow{a.s.} |X| \ (n \to \infty) \Rightarrow \|X_n\|_{\Phi} \xrightarrow{n\to\infty} \|X\|_{\Phi}.$$

補題 5.19

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