本日の講義内容

- 2.1 正規分布の裾確率と積率母関数
- 2.2 劣Gauss型確率変数列と
  その裾確率評価
- 2.3 劣指数型確率変数列
- 2.4 最大値の評価

2.1 正規分布の確率と確率母関数

定義 2.1 $X \sim N(\mu, \sigma^2)$.

$N(0,1)$を標準正規分布

$\Leftrightarrow$

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (-\infty < x < \infty).$$

ただし、$\mu = E[X]$, $\sigma^2 = Var[X]$ $(0 < \sigma < \infty)$.

$(\Omega, \mathcal{A}, P_r)$: 確率空間

${}^{\forall} a < b$ に対して

$$P_r(\omega \in \Omega ; a < X(\omega) \leq b) = P_r(a < X \leq b)$$
$$= \int_a^b p(x)\,dx.$$

命題2.3 (Millの不等式) $X \sim N(\mu, \sigma^2)$.

$^\forall t > 0$ に対して

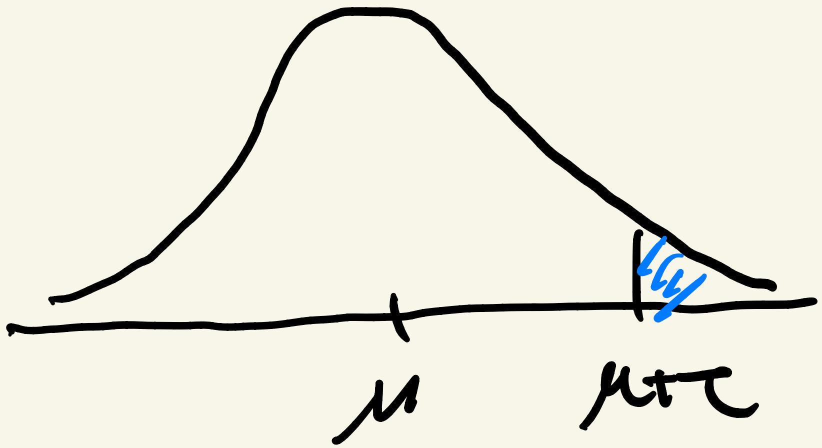
$$\Pr(X - \mu > t) \leqq \frac{\sigma}{t\sqrt{2\pi}} \exp\left(-\frac{t^2}{2\sigma^2}\right).$$

$$\Pr(X - \mu < -t) <$$

$$\text{〃}$$

$$\Pr(|X - \mu| > t)$$

$$\leqq \frac{\sigma}{t}\sqrt{\frac{2}{\pi}} \exp\left(-\frac{t^2}{2\sigma^2}\right).$$



証明　$\mu=0,\ \sigma=1$ の場合を示せばよい．

$Z \sim N(0,1)$ とする．

(1) $$\Pr(Z>t) = \frac{1}{\sqrt{2\pi}}\int_t^\infty \exp\left(-\frac{z^2}{2}\right)dz$$

$$\leq \frac{1}{\sqrt{2\pi}}\int_t^\infty \underbrace{\frac{z}{t}}_{\geq 1}\exp\left(-\frac{z^2}{2}\right)dz$$

$$= \frac{1}{t\sqrt{2\pi}}\int_t^\infty \left\{-\frac{\partial}{\partial z}\exp\left(-\frac{z^2}{2}\right)\right\}dz$$

$$= \frac{1}{t\sqrt{2\pi}}\left[-\exp\left(-\frac{z^2}{2}\right)\right]_t^\infty = \frac{1}{t\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)$$

(2) 2番目の不等号は $p(z)=\frac{1}{\sqrt{2\pi}}e^{-z^2/2}$ は偶関数であることより

No. 5

$$\Pr(Z<-t)=\Pr(Z>t)$$

よりわかる.

(3) $\Pr(|Z|>t)=\Pr(\{Z>t\}\cup\{Z<-t\})$

$$=\Pr(Z>t)+\Pr(Z<-t)$$

$$=2\Pr(Z>t)$$

□

$Z \sim N(0,1)$ の積率母関数. $s \in \mathbb{R}$ に対して　　No. 6

$$M_Z(s) := E\left[e^{sZ}\right]$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{sz} e^{-z^2/2}\, dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{ -\frac{(z-s)^2}{2} + \frac{s^2}{2} \right\} dz$$

$$= \exp\left(\frac{s^2}{2}\right) \int_{-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(z-s)^2}{2}\right) dz}_{N(s,1)\text{ の p.d.f } = 1}$$

$$= \exp\left(\frac{s^2}{2}\right).$$

$X \sim N(\mu, \sigma^2)$ のとき，$s \in \mathbb{R}$ に対して No.7

$$M_X(s) = E[e^{sX}]$$

$$= E[e^{s(\sigma Z + \mu)}]$$

$$= e^{s\mu} E[e^{(s\sigma)Z}]$$

$$= e^{s\mu} M_Z(s\sigma)$$

$$= \exp\left\{ \mu s + \frac{\sigma^2}{2} s^2 \right\}.$$

□

2.2 多 Gauss 型確率変数列とその収縮評価 No.8

$$X_1, X_2, \dots, X_n \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$$

$$\bar{X} := \frac{1}{n}\sum_{j=1}^{n} X_j \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

すると $s \in \mathbb{R}$ に対して

$$M_{\bar{X}-\mu}(s) = \exp\left(\frac{\sigma^2}{2n} s^2\right)$$

以下で求める補題2.10を用いると　No.9

$^\forall t>0$に対して

$$\Pr(|\bar{X}-\mu|>t)\leq 2\exp\left(-\frac{nt^2}{2\sigma^2}\right)$$

となることがわかる. すると、十分小さな$\delta>0$に対して

$$\Pr\left(\bar{X}-\sigma\sqrt{\frac{2\log(2/\delta)}{n}}\leq\mu\leq\bar{X}+\sigma\sqrt{\frac{2\log(2/\delta)}{n}}\right)>1-\delta$$

を得る

No. 16

$$\Pr(|\bar{X}-\mu|>t) \leq \underbrace{2\exp\left(-\frac{nt^2}{2\sigma^2}\right)}$$

$= \delta$ とおく.

$$2\exp\left(-\frac{nt^2}{2\sigma^2}\right)=\delta \iff -\frac{nt^2}{2\sigma^2}=\log\left(\frac{\delta}{2}\right)$$

$$\iff t^2=\frac{2\sigma^2}{n}\log\left(\frac{2}{\delta}\right)$$

$$\iff t=\pm\sigma\sqrt{\frac{2\log(2/\delta)}{n}}$$

$$\Pr\left(|\bar{X}-\mu|>\sigma\sqrt{\frac{2\log(2/\delta)}{n}}\right)\leq\delta$$

補事象を取る $\Pr(A)\leq\delta \Rightarrow \Pr(A^c)>1-\delta$

No. 11

$$\Pr\left(|\bar{X} - \mu| \leq \sigma \sqrt{\frac{2\log(2/\delta)}{n}} \right) > 1 - \delta$$

$$\parallel$$

$$\Pr\left(\bar{X} - \sqrt{\frac{2\log(2/\delta)}{n}} \leq \mu \leq \bar{X} + \sqrt{\frac{2\log(2/\delta)}{n}} \right)$$

定義2.4　$X \sim \mathrm{subG}(\sigma^2)$

$\Longleftrightarrow$

(1) $E[X] = 0$

(2) $\forall \lambda \in \mathbb{R}$ に対して

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

$N(0, \sigma^2)$ の積率母関数

注意2.5 $\mathrm{subG}(\sigma^2)$ は分布族を表すので、

「$X \sim \mathrm{subG}(\sigma^2)$」は記号の乱用であることに注意せよ。

「$X \sim N(\mu, \sigma^2)$」とはちがう。

定義2.6 $\underline{X}$ を $\mathbb{R}^d$ 値の確率べクトルとする。

$\underline{X} \sim \mathrm{subG}_d(\sigma^2)$

$\iff$

(1) $E[\underline{X}] = \underline{0}_d$ (2) $\forall \underline{u} \in S^{d-1}$ に対して

$\underline{u}^T \underline{X} \sim \mathrm{subG}(\sigma^2)$

ただし、$S^{d-1} = \{\underline{u} \in \mathbb{R}^d ; \underline{u}^T \underline{u} = 1\}$

$|\cdot|_{2,d}$ を $\mathbb{R}^d$ の Euclid norm

定義2.8　$\underline{X}$を $d \times T$ のランダム行列とする.　No.13

$$\underline{X} \sim \mathrm{subG}_{d\times T}(\sigma^2)$$

$\Leftrightarrow$

(1) $E[\underline{X}] = \underline{O}_{d\times T}$

(2) ${}^\forall \underline{u} \in S^{d-1}, \underline{v} \in S^{T-1}$ に対して

$$\underline{u}^\top \underline{X} \underline{v} \sim \mathrm{subG}(\sigma^2)$$

補題 2.9 (Markovの不等式)

X は非負値確率変数とする。すなわち、$\Pr(X \geq 0) = 1$。
さらに、$E[X] < \infty$ とする。このとき、$\forall a > 0$ に対して

$$\Pr(X > a) \leq \frac{E[X]}{a}.$$

証明

$A = (a, \infty) \subset \mathbb{R}$ とする。

$$\mathbb{1}_A(x) = \begin{cases} 1 & (x \in A) \\ 0 & (x \notin A) \end{cases}$$

とおく。

このとき

$$E[X] = E[X\{\underbrace{\mathbb{1}_A(x) + \mathbb{1}_{A^c}(x)}_{=1}\}]$$

$$= E[X\mathbb{1}_A(x)] + E[\underbrace{X\mathbb{1}_{A^c}(x)}_{\geq 0}]$$

$$\geq E[\underbrace{X}_{\geq a}\mathbb{1}_A(x)]$$

$$\geq aE[\mathbb{1}_A(x)] = aP_r(X>a)$$

補題2.10

$X \sim \mathrm{subG}(\sigma^2)$ とする.

${}^\forall t > 0$ に対して

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

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

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

証明　Chernoff 限界のテクニックを用いる.

$\forall \lambda > 0$ に対して

$$\Pr(X > t) = \Pr\left(e^{\lambda X} > e^{\lambda t}\right)$$

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

subG(σ^2)

Markov

$$\leq \exp\left(\boxed{\frac{\sigma^2}{2}\lambda^2 - \lambda t}\right) =: \phi(\lambda)$$

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

$$\Pr(X > t) \leq \inf_{\lambda > 0} \exp\{\Phi(\lambda)\}$$

$$= \exp\left(-\frac{t^2}{2\alpha^2}\right)$$

$\square$

補題　Xを非負値確率変数で$E[X]<\infty$ とする。このとき

No 19

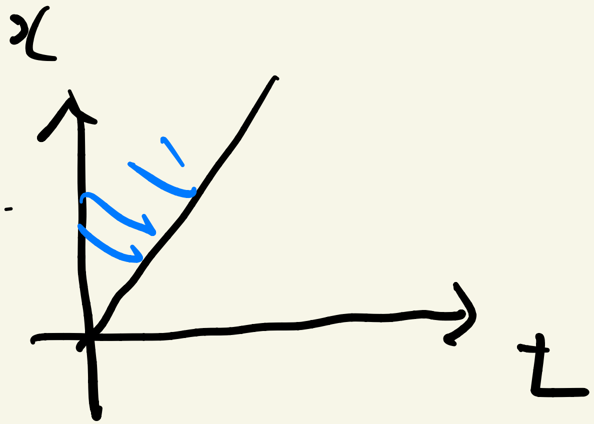
$$E[X] = \int_0^\infty \Pr(X>t)\,dt$$

$$\because \int_0^\infty \Pr(X>t)\,dt = \int_0^\infty \left\{ \int_t^\infty p(x)\,dx \right\} dt$$

$$= \int_0^\infty \left\{ \int_0^x p(x)\,dt \right\} dx$$

Fubini

$$= \int_0^\infty p(x) \left\{ \int_0^x dt \right\} dx = \int_0^\infty x\,p(x)\,dx = E[X]$$



No.20

$Z \sim N(0, \sigma^2)$ と $k \in \mathbb{N}$ に対して

$$E[|Z|^k] = \frac{1}{\sqrt{2\pi}}(2\sigma^2)^{k/2}\Gamma\left(\frac{k+1}{2}\right)$$

ただし

$$\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx$$

No.20

補題2.11 $X \sim \mathrm{subG}(\sigma^2)$とする.

このとき, $k \in \mathbb{N}$に対して

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

特に

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

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

No.21

証明 補題2.10より, $^\forall t>0$に対して

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

となる. 補題より

$$E[|X|^k] = \int_0^\infty \Pr(|X|^k > t)\,dt$$

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

$$\leq \int_0^\infty 2\exp\left(-\frac{t^{2/k}}{2\sigma^2}\right)dt$$

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

$$\left(\begin{array}{l} u = \dfrac{t^{2/k}}{2\sigma^2} \curvearrowright\ t = (2\sigma^2)^{\frac{k}{2}}\, u^{\frac{k}{2}} \\[2ex] dt = \dfrac{k}{2}(2\sigma^2)^{\frac{k}{2}}\, u^{\frac{k}{2}-1} \end{array} \right)$$

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

No.23

不等式

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

を用いると

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

$$\leq e^{1/e} \sigma \sqrt{k}$$

$k=1$ について

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

$X \sim \mathrm{subG}(\sigma^2) \iff E[e^{\lambda X}] \le \exp\left(\frac{\sigma^2}{2}\lambda^2\right)$

$X \sim \mathrm{subG}(\sigma^2) \Rightarrow \Pr(X > t) \le \exp\left(-\frac{\sigma^2}{2}\lambda^2\right)$

であった。

逆に

$P(X > t) \le \exp\left(-\frac{\sigma^2}{2}t^2\right) \Rightarrow E[e^{\lambda X}] \le \exp(4\sigma^2\lambda^2)$

$P(X < -t) \le$ 〃 $\qquad X \sim \mathrm{sub}(8\sigma^2)$

を示そう

補題2.12 以下を仮定する.

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

(2) $^\forall t>0$ に対して

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

ただし、$0<\sigma<\infty$.

このとき, $^\forall \lambda>0$ に対して

$$E[e^{\lambda X}]\leq e^{4\sigma^2\lambda^2}$$

証明

$$E[e^{\lambda X}] = E\left[1 + \sum_{k=1}^{\infty} \frac{(\lambda X)^k}{k!}\right]$$

$$= 1 + \sum_{k=1}^{\infty} \frac{\lambda^k E[X^k]}{k!}$$

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

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

∵ 補題 2.11

$$= 1 + \sum_{k=1}^{\infty} \frac{(2\sigma^2\lambda^2)^{k}(2k)\Gamma(k)}{(2k)!}$$

偶数項

$$+ \sum_{k=1}^{\infty} \frac{(2\sigma^2\lambda^2)^{k+(1/2)}(2k+1)\Gamma(k+0.5)}{(2k+1)!}$$

奇数項

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

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

$\because\ 2(k!)^2 \leq (2k)!$

No.23

$$= 1 + \left(1 + \sqrt{\frac{\sigma^2\lambda^2}{2}}\right) \sum_{k=1}^{\infty} \frac{(2\sigma^2\lambda^2)^k}{k!}$$

$$= 1 + \left(1 + \sqrt{\frac{\sigma^2\lambda^2}{2}}\right)\left(e^{2\sigma^2\lambda^2} - 1\right)$$

$$= e^{2\sigma^2\lambda^2} + \sqrt{\frac{\sigma^2\lambda^2}{2}}\left\{e^{2\sigma^2\lambda^2} - 1\right\}$$

$$\le e^{4\sigma^2\lambda^2}$$

注文 2.13

$\square$

2.2.1 独立 Gauss 確率変数列の和

$$X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$$

$\underline{a} = (a_1, \ldots, a_n) \in \mathbb{R}^n$; 定数ベクトル

$$\sum_{j=1}^{n} a_j X_j \sim N(0, |\underline{a}|_{2,n}^2 \sigma^2)$$

ただし、$|\underline{a}|_{2,n} = \sqrt{\underline{a}^T \underline{a}} = \sqrt{\sum_{j=1}^{n} a_j^2}$

命題2.14 $X_1, X_2, \dots, X_n \overset{\text{i.i.d.}}{\sim} \mathrm{subG}(\sigma^2)$

$\underline{X} = (X_1, \dots, X_n)^T$ とする. このとき,

$$\underline{X} \sim \mathrm{subG}_n(\sigma^2).$$

∵) $\underline{u} = (u_1, \dots, u_n)^T \in S^{n-1}$ と $s>0$ に対して

$$E[\exp\{s\underline{u}^T\underline{X}\}] = \prod_{j=1}^{n} E[e^{su_jX_j}]$$

$$\leq \prod_{j=1}^{n} \exp\left(\frac{\sigma^2 s^2 u_j^2}{2}\right)$$

$$= \exp\left(\frac{1}{2}\sigma^2 s^2 |\underline{u}|_{2,n}^2\right) = \exp\left(\frac{\sigma^2}{2}s^2\right)$$

□

系2.15 $X_1, X_2, \ldots, X_n$ は独立で

$X_j \sim \mathrm{subG}(\sigma^2)$ $(j=1,2,\ldots,n)$ とする.

$\underline{a}=(a_1,\ldots,a_n)^T \in \mathbb{R}^n$ と $t>0$ に対して

$$P_r\left(\sum_{j=1}^{n} a_j X_j > t\right) \leq \exp\left(-\frac{t^2}{2\sigma^2 |\underline{a}|_{2,n}^2}\right)$$

$$P_r\left(\sum_{j=1}^{n} a_j X_j < -t\right) \leq \exp\left(-\frac{t^2}{2\sigma^2 |\underline{a}|_{2,n}^2}\right).$$

特に

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

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

証明 命題2.15を用いる. $\forall \lambda > 0$に対して

$$\Pr\left(\sum_{j=1}^{n} a_j X_j > t\right) = \Pr\left(\exp\left(\lambda \sum_{j=1}^{n} a_j X_j\right) > e^{\lambda t}\right)$$

$$\leq e^{-\lambda t} E\left[\exp\left(\lambda \sum_{j=1}^{n} a_j X_j\right)\right]$$

Markov

No.28

$$\leq e^{-\lambda t} \exp\left(\frac{1}{2}\lambda^2 |\underline{a}|_{2,n}^2 \sigma^2\right)$$

$$= \exp\left\{ \overbrace{\boxed{\frac{\sigma^2 |\underline{a}|_{2,n}^2}{2}\lambda^2 - \lambda t}}^{\phi(\lambda)} \right\}$$

$$= \exp\left[\frac{\sigma^2 |\underline{a}|_{2,n}^2}{2} \left\{ \lambda - \frac{t}{\sigma^2 |\underline{a}|_{2,n}^2} \right\}^2 - \frac{t^2}{2\sigma^2 |\underline{a}|_{2,n}^2} \right]$$

よって

$$\Pr\left(\sum_{j=1}^{n} a_j X_j > t\right) \leq \inf_{\lambda > 0} \exp\{\phi(\lambda)\}$$

$$= \exp\left(-\frac{t^2}{2\sigma^2 |\underline{a}|_{2,n}^2}\right)$$

□

補題2.16 (Hoeffdingの補題) $a < b$ とする.

$E[X] = 0$ 　 $Pr(a \le X \le b) = 1$ とする

このとき, ${}^\forall \lambda \in \mathbb{R}$ に対して

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

よって, $X \sim subG\left(\frac{(b-a)^2}{4}\right)$ である

証明 関数 $\mathbb{R} \ni t \mapsto e^{\lambda t} \in (0, \infty)$ No.30

の凸性を用いる.

$$x = \gamma x + (1-\gamma) x ; \quad \gamma = \frac{x-a}{b-a} \quad (0 \leq \gamma \leq 1)$$

とおく. すると

$$e^{\lambda x} \leq \gamma e^{\lambda b} + (1-\gamma) e^{\lambda a}$$

$$= \frac{x-a}{b-a} e^{\lambda b} + \frac{b-x}{b-a} e^{\lambda a}$$

よって

$$E[e^{\lambda X}] \leq \frac{E[X]-a}{b-a} e^{\lambda b} + \frac{b-E[X]}{b-a} e^{\lambda a}$$

$$= -\frac{a}{b-a} e^{\lambda b} + \frac{b}{b-a} e^{\lambda a}$$

$$=: g(u)$$

・ $g(u) = -cu + \log\{1 - c + ce^u\}$

$u = s(b-a)$; $c = -\dfrac{a}{b-a}$; $1 - c = \dfrac{b}{b-a}$

あとは $g(u)$ を評価 $(u>0)$

$$g(0) = \dot{g}(0) = 0; \quad \ddot{g}(u) \leq \frac{1}{4} \quad (u>0)$$

Taylor 展開より

$$g(u) = g(0) + u\dot{g}(0) + \frac{u^2}{2}\ddot{g}(\xi)$$

$$(0 < \xi < u)$$

$$\leq \frac{u^2}{2}\ddot{g}(\xi) \leq \frac{u^2}{8} = \frac{s^2(b-a)^2}{8}$$

□

命題2.17 (Hoeffding) $X_1, X_2, \ldots, X_n$ は独立で

$Pr(a_j \leq X_j \leq b_j) = 1 \quad (j = 1, 2, \ldots, n)$

とする。$^\forall t > 0$ に対して

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

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

証明 $Y_j = X_j - E[X_j]$ とおく

すると

$$E[Y_j]=0,\ Pr(a_j - E[X_j] \leq \xi \leq b_j - E[X_j]) = 1$$

補題2.16を用いると

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

すると

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

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

No.34

$$= \Pr\left(\exp\left\{ s \sum_{j=1}^{n} \tau_j \right\} > e^{nst} \right)$$

$$\leq e^{-nst} E\left[\exp\left\{ s \sum_{j=1}^{n} \tau_j \right\} \right]$$

Markov

$$= e^{-nst} \prod_{j=1}^{n} E\left[e^{s\tau_j} \right]$$

独立性

$$\leq e^{-nst} \prod_{j=1}^{n} \exp\left\{ \frac{s^2}{8} (b_j - a_j)^2 \right\}$$

補題2.16

$$= \exp\left\{ \frac{\sum_{j=1}^{n}(b_j - a_j)^2}{8}\left(\lambda - \frac{4n\tau}{\sum_{j=1}^{n}(b_j - a_j)^2}\right)^2 - \frac{2n^2\tau^2}{\sum_{j=1}^{n}(b_j - a_j)^2} \right\}$$

ここで、 $\lambda = 4n\tau / \sum_{j=1}^{n}(b_j - a_j)^2$ とおくと

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

例 2.18 $\Pr(X=1) = \Pr(X=-1) = \frac{1}{2}$

このとき, ${}^\forall s \in \mathbb{R}$ に対して

$$E[e^{sX}] = \frac{e^{s}+e^{-s}}{2} = \cosh(s) \leq \exp\left(\frac{s^2}{2}\right)$$

Hoeffding の補題

$a = -1$; $b = 1$; $E[X]$

より

$$E[e^{sX}] \leq \exp\left(\frac{s^2}{2}\right)$$

□

2.3 劣指数型確率変数

両側 Laplace 分布

$$X \sim \mathrm{Lap}(1) \iff p(x) = \frac{1}{2}e^{-|x|} \quad (x \in \mathbb{R})$$

$X \sim \mathrm{Lap}(1)$ とする. $\forall t > 0$ に対して

$$\Pr(|X| > t) = 1 - \Pr(|X| \leq t) = 1 - \int_{-1}^{1} \frac{1}{2} e^{-|s|} ds$$

$$= 1 - \int_{0}^{1} e^{-s} ds = e^{-t}$$

注 $\underbrace{\text{正規分布}}_{(e^{-t^2})}$のように裾確率は急激に減少していない！

$X \sim \mathrm{Lap}(1)$ の m.g.f.　$|s|<1$ に対して

$$E[e^{sX}] = \frac{1}{2}\int_{-\infty}^{\infty} e^{-|x|}\cdot e^{sx}\,dx$$

$$= \frac{1}{2}\left\{\int_0^{\infty} e^{-(1-s)x}\,dx + \int_{-\infty}^{0} e^{(1+s)x}\,dx\right\}$$

$$= \frac{1}{2}\left\{\left[-\frac{1}{1-s}e^{-(1-s)x}\right]_0^{\infty} + \left[\frac{1}{1+s}e^{(1+s)x}\right]_{-\infty}^{0}\right\}$$

$$= \frac{1}{2}\left\{\frac{1}{1-s} + \frac{1}{1+s}\right\} = \frac{1}{1-s^2}$$

$\square$

$X \sim \mathrm{Lap}(1)$ に対して

$$E[e^{sX}] \leq 2e^{s^2} \quad \left(|s| < \frac{1}{2}\right)$$

である。　注 原点近辺では正規分布と同じ

補題 2.20　X は $E[X]=0$ をみたす確率変数で、ある $\lambda>0$ が存在して、$^\forall t>0$ に対して

$$\Pr(|X|>t) \leq 2\exp\left(-\frac{2t}{\lambda}\right)$$

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

$$E[|X|^k] \leq \lambda^k k!.$$

さらに

$$\{E[|X|^k]\}^{1/k} \leq 2\lambda k.$$

と

$$E[e^{\alpha X}] \leq \exp(2\alpha^2\lambda^2) \quad \left(^\forall |\alpha| \leq \frac{1}{2\lambda}\right).$$

証明の概略

$k \in \mathbb{N}$ と $^\forall t > 0$ に対して

$$E[|X|^k] = \int_0^\infty \Pr(|X|^k > t)\, dt$$

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

$$\leq \int_0^\infty 2\exp\left(-\frac{2t^{1/k}}{\lambda}\right) dt \quad (\because \text{仮定})$$

$$= 2\left(\frac{\lambda}{2}\right)^k k \int_0^\infty e^{-u} u^{k-1}\, du$$

$$\left(u = \frac{2t^{1/k}}{\lambda} \Leftrightarrow t = \left(\frac{\lambda}{2}\right)^k u^k \rightsquigarrow dt = \left(\frac{\lambda}{2}\right)^k k u^{k-1}\, du \right)$$

$$\leq \lambda^k k \Gamma(k) = \lambda^k k!$$

$|s| < 1/(2\lambda)$ に対して

$$E[e^{sX}] = E\left[1 + \sum_{k=1}^{\infty} \frac{(sX)^k}{k!}\right]$$

$$= 1 + \sum_{k=2}^{\infty} \frac{E[(sX)^k]}{k!}$$

$$\leq 1 + \sum_{k=2}^{\infty} \frac{|s|^k E[|X|^k]}{k!}$$

$$\leq 1 + \sum_{k=2}^{\infty} (|s|\lambda)^k \qquad (\because \text{1番目の主張})$$

$$= 1 + s^2\lambda^2 \sum_{k=0}^{\infty} (|s|\lambda)^k$$

$$\leq 1 + s^2\lambda^2 \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k \qquad \because)\ |s| \leq \frac{1}{2\lambda}$$

No. 42

$$= 1 + A^2 \lambda^2 \lim_{k \to \infty} \frac{1 - \left(\frac{1}{2}\right)^{k+1}}{1 - \frac{1}{2}}$$

$$= 1 + 2A^2 \lambda^2 \leq e^{2A^2 \lambda^2}$$

$\boxtimes$

定義 2.21 $\lambda > 0$ とする.

$X \sim \mathrm{subE}(\lambda)$

$\Longleftrightarrow$

(1) $E[X] = 0$; (2) ${}^\forall |\alpha| < 1/\lambda$ に対して, $E[e^{\alpha X}] \leq \exp\left(\frac{\alpha^2 \lambda^2}{2}\right)$.

補題 2.22 $X \sim \mathrm{subG}(\sigma^2)$ とし, $Z := X^2 - E[X^2]$ とおく. すると $Z \sim \mathrm{subE}(16\sigma^2)$ となる

証明の説明は省略

2.3.1 Bernstein の不等式

命題2.33 $X_1, X_2, \dots, X_n$ は独立で,

$$E[X_j]=0 \ ; \ X_j \sim \mathrm{subE}(\lambda) \quad (j=1,2,\dots,n)$$

とする. このとき, $^\forall t>0$ に対して

$$\Pr(\bar{X}>t) \vee \Pr(\bar{X}<-t) \le \exp\left[-\frac{n}{2}\left(\frac{t^2}{\lambda^2} \wedge \frac{t}{\lambda}\right)\right].$$

ただし, $a \vee b = \max\{a, b\}$; $a \wedge b = \min\{a, b\}$.

証明の概略　一般性を失わず，$\lambda=1$としてよい．

$\forall s>0$に対して

$$\Pr(\bar{X}>t)=\Pr\left(\sum_{j=1}^{n}X_j>nt\right)$$

$$=\Pr\left(\exp\left(s\sum_{j=1}^{n}X_j\right)>e^{nst}\right)$$

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

$$=e^{-nst}\prod_{j=1}^{n}E\left[e^{sX_j}\right]$$

$$\leq e^{-nst}\prod_{j=1}^{n}e^{s^2/2}=\exp\left(\frac{ns^2}{2}-snt\right)$$

No.46

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

よって

$$\Pr(\bar{X} > t) = \inf_{0 < s \le 1} \exp\left\{ \frac{n}{2}(s-\tau) - \frac{\tau^4}{2} \right\}$$

$$= \exp\left(-\frac{n}{2}(t^2 \wedge \tau) \right).$$

$\square$

2.4 最大値の確率不等式

2.4.1 有限集合上の最大値

命題 2.24 $X_j \sim \mathrm{subG}(\sigma^2)$ $(j=1,2,\ldots,n)$ とする.

このとき

$$E\left[\max_{1\le j\le n} X_j\right] \le \sigma\sqrt{2\log n}$$

$$E\left[\max_{1\le j\le n} |X_j|\right] \le \sigma\sqrt{2\log(2n)}.$$

${}^\forall t>0$に対して

$$\Pr\left(\max_{1\le j\le n} X_j > t\right) \le n\exp\left(-\frac{t^2}{2\sigma^2}\right)$$

$$\Pr\left(\max_{1\le j\le n} |X_j| > t\right) \le 2n\exp\left(-\frac{t^2}{2\sigma^2}\right).$$

証明の概略　${}^\forall s>0$に対して

$$E\left[\max_{1\le j\le n} X_j\right] = \frac{1}{s}E\left[\log\exp\left(s\max_{1\le j\le n} X_j\right)\right]$$

$$= \frac{1}{s}E\left[\log\left(\max_{1\le j\le n} e^{sX_j}\right)\right]$$

$$\leq \frac{1}{\lambda} \log \left\{ E\left[\max_{1 \leq j \leq n} e^{\lambda X_j} \right] \right\}$$

$\because$ Jense

$$\leq \frac{1}{\lambda} \log \left\{ \sum_{j=1}^{n} E\left[e^{\lambda X_j} \right] \right\}$$

$$\leq \frac{1}{\lambda} \log \left\{ \sum_{j=1}^{n} \exp\left(\frac{\lambda^2 \sigma^2}{2} \right) \right\}$$

$\because$ $X_j \sim \mathrm{subG}(\sigma^2)$

$$= \frac{\log n}{\lambda} + \frac{\sigma^2 \lambda^2}{2}$$

ここで

$$\lambda = \sqrt{\frac{\sigma^2}{2 \log n}}$$ とおくと

$$\frac{\log n}{\lambda} + \frac{\lambda^2 \sigma^2}{2} = \sigma\sqrt{2\log n}$$

となるので

$$E\left[\max_{1 \le j \le n} X_j\right] \le \sigma\sqrt{2\log n}.$$

次に

$$\Pr\left(\max_{1\le j\le n} X_j > t\right) = \Pr\left(\bigcup_{j=1}^{n} \{X_j > t\}\right)$$

$$\le \sum_{j=1}^{n} \Pr(X_j > t) \qquad \because) \text{ union bound}$$

$$\le n\exp\left(-\frac{t^2}{2\sigma^2}\right) \qquad \because) \text{ 補題2.10.}$$

$\square$

2.4.3 ℓ_2 球上の最大値

$\mathbb{R}^d$ の ℓ_2 球

$$B_2 := \left\{ \underline{x} = (x_1, \dots, x_d)^T \in \mathbb{R}^d ; \sum_{j=1}^{d} x_j^2 \leq 1 \right\}.$$

定義 2.27 $K \subset \mathbb{R}^d$ を部分集合とする. 集合 $N \subset \mathbb{R}^d$ は $\mathbb{R}^d$ 上の距離 $d(\cdot, \cdot)$ に関する ε 網 $(\varepsilon > 0)$ であるとは, $N \subset K$ であって, $^\forall \underline{x} \in K$ に対して ある $\underline{z} \in N$ が存在して

$$d(\underline{x}, \underline{z}) \leq \varepsilon$$

となっているときをいう.

補題2.29 $0 < \varepsilon < 1$ を固定する.

単位球 B_2 は d 次元 Euclid 距離に関する ε 網 N で

$$\#(N) \leq \left(\frac{3}{\varepsilon}\right)^d \quad (※)$$

をみたすものが存在する. ただし, $\#(N)$ は N の元の個数である.

証明 再帰的に (※) をみたす ε 網を構成する.

$\underline{x}_1 = \underline{0}_d$ を取る. $\forall i \in \mathbb{N}\ (i \geq 2)$ において. $\underline{x}_i \in B_2$ は, $\underline{x} \in B_2$ で

$$|\underline{x} - \underline{x}_j|_{2,d} > \varepsilon \ (j = 1, 2, \ldots, i-1)$$

をみたすような点の1つを $\boldsymbol{x}_j$ とする.

このような $\boldsymbol{x}$ が存在しないとき、この操作を止める. ← B_2 はコンパクトなので必ず止まる!

すると、操作が止まれば ε網となる. これを N とおく.

$\{\boldsymbol{x}_i\}$ の選び方より、$\forall \boldsymbol{x}, \boldsymbol{y} \in N$ について

$$|\boldsymbol{x}-\boldsymbol{y}|_{2,d} > \varepsilon$$

なので、$\boldsymbol{x} \in N$ を中心とする半径 $\varepsilon/2$ の Euclid 球

は互いに排反である. さらに

$$\bigcup_{\boldsymbol{z}\in N}\left\{\boldsymbol{z}+\frac{\varepsilon}{2}B_2\right\} \subset \left(1+\frac{\varepsilon}{2}\right)B_2$$

である. ただし $\boldsymbol{z}+\varepsilon B_2 := \{\boldsymbol{z}+\varepsilon\boldsymbol{x} \,;\, \boldsymbol{x}\in B_2\}$ である

よって．体積を評価すると　　No. 58

$$\mathrm{vol}\left(\left(1+\frac{\varepsilon}{2}B_2\right)\right) \geq \mathrm{vol}\left(\bigcup_{z\in N}\left\{z+\frac{\varepsilon}{2}B_2\right\}\right)$$

$$=\sum_{z\in N}\mathrm{vol}\left(z+\frac{\varepsilon}{2}B_2\right).$$

これは

$$\left(1+\frac{\varepsilon}{2}\right)^d \geq \#(N)\left(\frac{\varepsilon}{2}\right)^d$$

$$\curvearrowright \quad \#(N) \leq \left(1+\frac{2}{\varepsilon}\right)^d \leq \left(\frac{3}{\varepsilon}\right)^d.$$

$\square$

No. 5

命題 2.30. $\underline{X} \in \mathbb{R}^d \sim \mathrm{subG}_d(\sigma^2)$ とする.

このとき

$$E\left[\max_{\underline{\theta} \in B_2} \underline{\theta}^T \underline{X}\right] = E\left[\max_{\underline{\theta} \in B_2} |\underline{\theta}^T \underline{X}|\right] \leq 4\sqrt{d}.$$

さらに, $\forall \delta > 0$ に対して

$$\Pr\left(\max_{\underline{\theta} \in B_2} \underline{\theta}^T \underline{X} = \max_{\underline{\theta} \in B_2} |\underline{\theta}^T \underline{X}| \leq 4\sigma\sqrt{d} + 2\sigma\sqrt{2\log\left(\frac{1}{\delta}\right)}\right) \geq 1 - \delta.$$

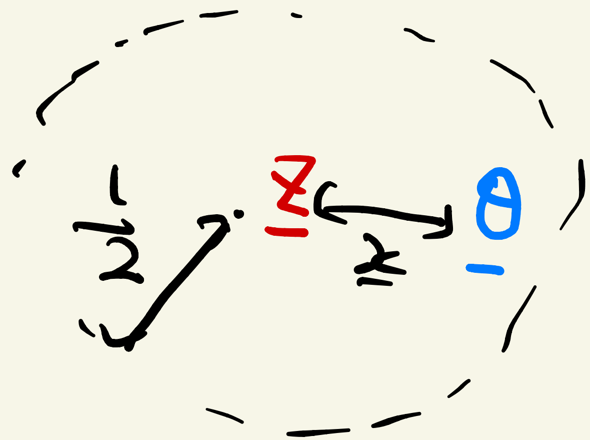
証明の概略　N を B_2 の $\frac{1}{2}$ 網とする.　No.56

すると補題2.29 ($\varepsilon = 1/2$) より

$$\#(N) \le 6^d.$$

$\underline{\theta} \in B_2$ に対して, ある $^\exists \underline{z} \in N$ と $\underline{x} \in \mathbb{R}^d$ で $\|\underline{x}\|_{2,d} \le \frac{1}{2}$ なるものが存在して

$$\underline{\theta} = \underline{z} + \underline{x}$$



と書ける. しかし

$$\max_{\underline{\theta} \in (1/2)B_2} \underline{\theta}^T \underline{X} = \frac{1}{2} \max_{\underline{\theta} \in B_2} \underline{\theta}^T \underline{X}$$

なので.

$$\max_{\underline{\theta} \in B_2} \underline{\theta}^T \underline{X} \leq \max_{\underline{z} \in N} \underline{z}^T \underline{X} + \max_{\underline{x} \in (1/2) B_2} \underline{x}^T \underline{X}$$

$$= \max_{\underline{z} \in N} \underline{z}^T \underline{X} + \frac{1}{2} \max_{\underline{\theta} \in B_2} \underline{\theta}^T \underline{X}$$

よって

$$\max_{\underline{\theta} \in B_2} \underline{\theta}^T \underline{X} \leq 2 \max_{\underline{z} \in N} \underline{z}^T \underline{X}.$$

よって

$$E\left[\max_{\underline{\theta} \in B_2} \underline{\theta}^T X \right] \leq E\left[2 \max_{\underline{z} \in N} \underline{z}^T X \right]$$

No. 58

$$\leq 2a\sqrt{2\log(\#(N))}$$ ∵) 命題 2.24

$$\leq 2a\sqrt{2d\log 6}$$ ∵) 補題 2.29

$$\leq 4a\sqrt{d}.$$

$t>0$ に対して

$$\Pr\left(\max_{\underline{\theta}\in B_2}\underline{\theta}^T\underline{x}>t\right)\leq\Pr\left(2\max_{\underline{z}\in N}\underline{z}^T\lambda>t\right)$$

$$=\Pr\left(\max_{\underline{z}\in N}z^T\lambda>\frac{t}{2}\right)$$

$$\Pr\left(\max_{1\leq j\leq n}|x_j|>t\right)\leq n\exp\left(-\frac{t^2}{2a^2}\right)$$

$$\uparrow \quad t\Rightarrow\frac{t}{2}$$

$$\leq\#(N)\exp\left(-\frac{t^2}{8a^2}\right)$$

$$\leq 6^d \exp\left(-\frac{\tau^2}{8\sigma^2}\right)$$

ここに

$$t := 4\sigma\sqrt{d} + 2\sigma\sqrt{2\log\left(\frac{1}{\delta}\right)}$$

とおくと

$$6^d \exp\left(-\frac{\tau^2}{8\sigma^2}\right) \leq \delta$$

となるので

$$P_r\left(\max_{\underline{\theta}\in B_2} \underline{\theta}^T \underline{x} > 4\sigma\sqrt{d} + 2\sigma\sqrt{2\log\left(\frac{1}{\delta}\right)}\right)$$

$$\leq P_r\left(\max_{\underline{\theta}\in B_2} \underline{\theta}^T \underline{x} > \tau\right) \leq 6^d \exp\left(-\frac{\tau^2}{8\sigma^2}\right)$$

$$\le \exp\left(-\frac{\tau^2}{8\sigma^2} + d\log 6\right) \le$$

$$= 6^d \exp\left(-\frac{\tau^2}{8\sigma^2}\right) \le \delta$$

よって

$$\Pr\left(\max_{\underline{\theta}\in B_2} \underline{\theta}^T \underline{x} \le 4\sigma\sqrt{d} + 2\sigma\sqrt{2\log\left(\frac{1}{\delta}\right)}\right) \ge 1-\delta.$$