確率数理要論4

- 極値分布

$X_1, \dots, X_n \sim G$ (i.i.d.) のとき、
$M_n := \max\{X_1, \dots, X_n\}$ の分布に興味がある。
(応用例：リスク分析)

M_n の分布
$$F_n(x) := P(M_n \leqq x) = \prod_{i=1}^{n} P(X_i \leqq x) = G^n(x)$$
を極値分布と言う。

<u>Ex.</u> (Gumbel分布)
$$G(x) = \begin{cases} 1 - e^{-\alpha x} & (x \geq 0) \\ 0 & (x < 0) \end{cases} \qquad (\alpha \geq 0) \text{ ：指数分布}$$
$$\Rightarrow F_n(x) = \begin{cases} (1 - e^{-\alpha x})^n & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

ここで、$^\forall x$ で $F_n(x) \longrightarrow 0 \quad (n \to \infty)$ なので、F_n は何らかの分布関数に弱収束しない。
そこで、
$$M_n - \alpha^{-1} \log(n) \qquad (\text{平均をずらす})$$
の分布を考える。

$$P(M_n - \alpha^{-1}\log(n) \leqq x)$$
$$= P(M_n \leqq x + \alpha^{-1}\log(n))$$
$$= \begin{cases} \left(1 - e^{-(\alpha x + \log(n))}\right)^n & (x \geqq -\frac{1}{\alpha}\log(n)) \\ 0 & (x < -\frac{1}{\alpha}\log(n)) \end{cases}$$
$$= \begin{cases} \left(1 - \frac{1}{n}e^{-\alpha x}\right)^n & (\quad ; \quad) \\ 0 & (\quad ; \quad) \end{cases}$$
ここで、
$$\left(1 - \frac{1}{n}e^{-\alpha x}\right)^n \longrightarrow e^{-e^{-\alpha x}}$$
がわかる。
$$F(x) := e^{-e^{-\alpha x}} \quad : \underline{\text{Gumbel 分布}}$$

$$M_n - \alpha^{-1}\log(n) \rightsquigarrow \text{Gumbel 分布}.$$

__Ex.__ (Fréchet 分布)
$$G(x) = \begin{cases} 1 - x^{-\alpha} & (x \geqq 1) \\ 0 & (x < 1) \end{cases} \quad (\alpha > 0) : \text{パレート分布}$$
このとき、
$$F_n(x) = (1 - x^{-\alpha})^n \quad (x \geqq 1)$$
これも $F_n(x) \longrightarrow 0 \quad (\forall x)$
代わりに、$\dfrac{M_n}{n^{\frac{1}{\alpha}}}$ を考える。
$$P\left(\frac{M_n}{n^{\frac{1}{\alpha}}} \leqq x\right) = F_n\left(M_n \leqq n^{\frac{1}{\alpha}}x\right)$$
$$= \left(1 - \frac{1}{n}x^{-\alpha}\right)^n$$
$$\xrightarrow{n \to \infty} e^{-x^{-\alpha}}$$
$$F(x) = \begin{cases} e^{-x^{-\alpha}} & (x \geqq 0) \\ 0 & (x < 0) \end{cases} \quad : \text{Fréchet 分布}.$$

Ex. (Weibull分布)

$$G(x) = \begin{cases} 1 & (x \geq 1) \\ 1-(1-x)^{\alpha} & (0 \leq x < 1) \\ 0 & (x < 0) \end{cases}$$: Beta分布の特殊例)

$(\alpha > 0)$

このとき、

$$F_n(x) = \{1-(1-x)^{\alpha}\}^n \quad (0 \leq x \leq 1)$$

これは、1に退化した分布に収束（1点のpoint massになる）

そこで、$n^{\frac{1}{\alpha}}(M_n - 1)$ を代わりに考えると、

$$P\left(n^{\frac{1}{\alpha}}(M_n-1) \leq x\right) \xrightarrow{n\to\infty} F(x) = \begin{cases} 1 & (x \geq 0) \\ e^{-(-x)^{\alpha}} & (x < 0) \end{cases}$$

: Weibull分布

※実際は $-X$ を Weibull分布と呼ぶことが多い。

Thm

$X_1, \dots, X_n \sim G$ (i.i.d.)

$M_n = \max\{X_1, \dots, X_n\}$

$\exists a_n, b_n \in \mathbb{R}$ s.t. $\dfrac{M_n - b_n}{a_n}$ が退化しない分布に収束（1点に収束しない）

$\Longrightarrow$ 分布の収束先は Gumbel, Fréchet, Weibull のどれかになる。

つまり、裾の落ち方が
- 指数オーダー $\longrightarrow$ Gumbel
- 多項式オーダー $\longrightarrow$ Fréchet
- 有界サポート $\longrightarrow$ Weibull

と、落ち方の違いで決まる。

○ 特性関数

Def

X: r.v. に対し、
$\phi(t) = E[e^{itX}]$ $(t \in \mathbb{R})$
を特性関数と呼ぶ。 (characteristic function)

※ $|e^{itX}| \leq 1$ より、期待値は常に存在。
※ $\mathbb{R}^k$-値 r.v. に対しても $\phi(t) = E[e^{it^TX}]$ $(t \in \mathbb{R}^k)$ とすれば
同様に特性関数を定義できる。以下の話は $\mathbb{R}^k$ 値でも成り立つ。

Lem

$\phi(0) = 1$, $|\phi(t)| \leq 1$ $({}^\forall t \in \mathbb{R})$, $\phi(t)$ は一様連続 //

Lem

X: r.v.
ϕ: 特性関数
ある $k > 0$ で $E[|X|^k] < \infty$ なら
ϕ は ${}^\forall t \in \mathbb{R}$ で k 回連続微分可能で $\phi^{(k)}(0) = i^k E[X^k]$ である。 //

(Proof) $\frac{\partial^k}{\partial t^k} e^{itX} = (iX)^k e^{itX}$ で、
$\left| \frac{\partial^k}{\partial t^k} e^{itX} \right| \leq |X|^k$ かつ $|X|^k$ は可積分
よって微分と積分が交換できて、(講義第2回目ノート参照)
$\frac{\partial^k}{\partial t^k} E[e^{itX}] = E\left[\frac{\partial^k}{\partial t^k} e^{itX}\right] = E[(iX)^k e^{itX}]$

特に $t=0$ のとき、右辺 $= E[(iX)^k]$ である。
さらに、右辺は t について連続であることも優収束定理より示せる。 //

Ex.

(1) 正規分布 ($N(\mu, \sigma^2)$)

$$f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad : \text{p.d.f.}$$

$$\phi(t) = \int_{-\infty}^{\infty} e^{itx} f(x|\mu,\sigma^2)\,dx$$
$$= \exp\left(it\mu - \frac{\sigma^2 t^2}{2}\right) \quad \text{(コーシーの積分定理)}$$

(2) Cauchy 分布

$$f(x|c) = \frac{1}{\pi}\frac{c}{x^2+c^2} \quad (c>0) \quad \text{(平均は存在しない)}$$
$$\phi(t) = e^{-c|t|}$$
(留数定理)

Lem

$X_1, \dots, X_n$：独立な r.v.
$\phi_1(t), \dots, \phi_n(t)$：その特性関数
$S_n = X_1 + X_2 + \dots + X_n$ の特性関数は
$$\phi(t) = E[e^{itS_n}] = \prod_{i=1}^{n} \phi_i(t)$$

(Proof) $e^{itS_n} = \prod_{i=1}^{n} e^{itX_i}$ と独立な確率変数の積になるので、
独立な r.v. の積の期待値が各期待値の積になることより従う。//

Thm

P, Q：2つの分布
$$P = Q \iff \phi_P(t) = \phi_Q(t) \quad (^\forall t \in \mathbb{R})$$
ただし、ϕ_P, ϕ_Q は P, Q に対する特性関数 //

($\Rightarrow$) は自明
($\Leftarrow$) を示すため、いくつかの補題を用意する。

Lem

$X \sim P,\ Y \sim Q$，XとYは独立な r.v. とする．
$X+Y$の分布は畳み込み分布 $P*Q$ に従う：

$$(P*Q)(A) = \int P(A-x)\, dQ(x). \qquad (A-x := \{x'-x \mid x' \in A\})$$

また，Xが密度$f(x)$を持つなら
$P*Q$も密度を持ち、$h(x) = \int f(x-y)\, dQ(y)$ で与えられる．

Proof

$Z = X+Y$ とすると、$Z \in A$ である確率は

$$\begin{aligned} E[\mathbb{1}_A(Z)] &= E[\mathbb{1}_A(X+Y)] \\ &= \iint \mathbb{1}_A(x+y)\, dP(x)\, dQ(y) \qquad (X\text{と}Y\text{は独立, Fubini}) \\ &= \iint \mathbb{1}_{A-y}(x)\, dP(x)\, dQ(y) \\ &= \int P(A-y)\, dQ(y) \end{aligned}$$

$\begin{cases} Z \in A \Rightarrow 1 \\ Z \notin A \Rightarrow 0 \end{cases}$

特に，Xが密度fを持つとき，

$$\begin{aligned} \text{左辺} &= \iint \mathbb{1}_A(x+y)\, f(x)\, dx\, dQ(y) \\ &= \iint \mathbb{1}_A(u)\, f(u-y)\, du\, dQ(y) \\ &= \int_A \underbrace{\left(\int f(u-y)\, dQ(y)\right)}_{= h(u)} du \qquad (\text{Fubini}) \end{aligned}$$

よって．hは$P*Q$の密度になっている． //

$X \sim P,\ Y \sim N(0, \sigma^2)$ のとき（X, Yは独立）
$Z = X+Y$ の分布は $P^{(\sigma)} = P*N(0,\sigma^2)$ で与えられる．

Lem

Pの特性関数が$g(t) = \int e^{itx}\, dP(x)$ で与えられる時，
$P^{(\sigma)}$は密度

$$f^{(\sigma)}(x) = \frac{1}{2\pi}\int g(t) \exp\left(-itx - \frac{\sigma^2 t^2}{2}\right) dt$$

を持つ．

(Proof)
↳ 次ページ

先の Lem より

$$f^{(\sigma)}(x) = \int \overbrace{\varphi_\sigma(x-y)}^{N(0,\sigma^2)\text{の密度}} dP(y)$$

$$= \int \left(\frac{1}{2\pi} \int \exp\left[i(y-x)t - \frac{\sigma^2 t^2}{2} \right] dt \right) dP(y)$$

$$= \frac{1}{2\pi} \int \underbrace{\left(\int e^{ity} dP(y) \right)}_{= g(t)} e^{-itx - \frac{\sigma^2 t^2}{2}} dt$$

$$= \frac{1}{2\pi} \int g(t) e^{-itx - \frac{\sigma^2 t^2}{2}} dt$$

//

Lem

$\sigma \to 0$ で $P^{(\sigma)} \rightsquigarrow P$ である.

(Proof)

$X \sim P,\ Y \sim N(0, \sigma^2)$ のとき.

$X + Y \xrightarrow{P} X \quad (\sigma \to 0)$ はすぐわかる. $\left(P(|X+Y-X| \geq \varepsilon) = P(|Y| \geq \varepsilon) \xrightarrow[\sigma \to 0]{} 0 \right)$

特に.

$X + Y \rightsquigarrow X$

である. $X + Y \sim P^{(\sigma)},\ X \sim P$ より示された.

//

(Thm の $\Leftarrow$ の証明)

$\phi_P = \phi_Q$ なら, $P^{(\sigma)}$ と $Q^{(\sigma)}$ は同じ密度関数 $f^{(\sigma)}$ を持つ.

(先の Lem より g に ϕ_P, ϕ_Q を代入.)

よって. $P^{(\sigma)} = Q^{(\sigma)}$ がわかる.

また. $P^{(\sigma)} \rightsquigarrow P,\ Q^{(\sigma)} \rightsquigarrow Q$ より.

${}^\forall f$: 有界連続 に対し,

$$\lim_{\sigma \to 0} \int f(x)\, dP^{(\sigma)}(x) = \int f(x)\, dP(x)$$

$$\|$$

$$\lim_{\sigma \to 0} \int f(x)\, dQ^{(\sigma)}(x) = \int f(x)\, dQ(x)$$

つまり. $\int f dP = \int f dQ$ (${}^\forall f$: 有界連続).

特に $P = Q$

//

※ この証明は多変量ver. でも使える.

Ex. (安定分布)

ある a_n が存在して. i.i.d.列 X_n に対し. $a_n \sum_{i=1}^{n} X_i$ の分布が n に対して変わらないとき. 安定分布 と言う.

(1) 正規分布:

$$X_1, \dots, X_n \sim N(0,1) \Longrightarrow \sum_{i=1}^{n} X_i / \sqrt{n} \sim N(0,1) \quad ({}^{\forall} n)$$

$$(\because) \ E\left[e^{it\left(\frac{\sum X_i}{\sqrt{n}}\right)}\right] = \prod_{i=1}^{n} E\left[e^{i\frac{t}{\sqrt{n}}X_i}\right] = \prod_{i=1}^{n} e^{-\frac{t^2}{2n}} = e^{-\frac{t^2}{2}}$$

$N(0,1)$ の c.f.

(2) Cauchy分布:

$$X_1, \dots, X_n \sim \mathrm{Cauchy}(c) \Longrightarrow \frac{\sum_{i=1}^{n} X_i}{n} \sim \mathrm{Cauchy}(c)$$

$$(\because) \ E\left[e^{it\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)}\right] = \prod_{i=1}^{n} E\left[e^{i\frac{t}{n}X_i}\right] = \prod_{i=1}^{n} e^{-|c|\frac{t}{n}} = e^{-|c|t}$$

Cauchy(c) の c.f. //

Thm (Lévyの連続性定理)

$(X_n)_{n=1}^{\infty}$ を r.v. の列として, $\phi_n(t) = E[e^{itX_n}]$ とする.

1. ある r.v. X が存在して, その c.f. が $\phi(t)$ であるとする.
   このとき,
   $$X_n \xrightarrow{w} X \Longrightarrow \underbrace{\phi_n(t) \longrightarrow \phi(t) \quad (n \to \infty) \quad ({}^{\forall} t \in \mathbb{R})}_{\text{各点収束}}$$
2. ある関数 $\phi : \mathbb{R} \to \mathbb{C}$ が存在して, ${}^{\forall} t \in \mathbb{R}$ で $\phi_n(t) \longrightarrow \phi(t)$ かつ, ϕ が $t=0$ で連続なら, ϕ はある r.v. X の特性関数で $X_n \xrightarrow{w} X$ である. //

(Proof) 1. は Portmanteau の定理より示せる.
2. は難しい. (証明は補足資料) //

そこで2の代わりに以下を考える。

Lévyの連続性定理2′:

$(X_n)_n$ が緊密(タイト)で、$\phi_n(t) \to \phi(t)$ $({}^{\forall}t \in \mathbb{R})$ なら ϕ はある r.v. X の c.f. で $X_n \xrightarrow{w} X$ である。

(2′の proof)

Lem (Prokhorovの定理)

1. $X_n \xrightarrow{w} X$ なら $(X_n)_n$ は緊密。
2. $(X_n)_n$ が緊密なら部分列 $(X_{n_j})_{j=1}^{\infty}$ が存在して、ある r.v. X へ弱収束 $X_{n_j} \xrightarrow{w} X$ する。

※ メトリックスペースのコンパクト集合は点列コンパクト
($\mathbb{R}$上の有界点列は収束部分列を持つ。)
(証明は後述. webの補足資料を参照)

$(X_n)_n$ は $O_p(1)$ (緊密)なので、任意の部分列 $(X_{n_j})_{j=1}^{\infty}$ も $O_p(1)$ で
さらに部分列 $(X_{n_{j_k}})_{k=1}^{\infty}$ が存在して、ある r.v. X に $X_{n_{j_k}} \xrightarrow{w} X$ となる。
(by Prokhorov)

よって、X の c.f. を $\tilde{\phi}(t)$ と書くと、$\phi_{n_{j_k}}(t) \to \tilde{\phi}(t)$ $({}^{\forall}t \in \mathbb{R})$ となるが、
仮定より $\phi_n(t) \to \phi(t)$ でもあるので、$\tilde{\phi}(t) = \phi(t)$ がわかる。
つまり、ϕ は X の c.f. (これは部分列の取り方によらずユニーク。)

ここで、もし $X_n \xrightarrow{w} X$ でなければ、${}^{\exists}f$: 有界連続で $E[f(X_n)] \not\to E[f(X)]$.
特に、ある部分列をとってきて、$|E[f(X_{n_j})] - E[f(X)]| \geq \varepsilon$ $({}^{\forall}j)$ とできる。しかし、さらに部分列 $(X_{n_{j_k}})_{k=1}^{\infty}$ をとって来ると、
$X_{n_{j_k}} \xrightarrow{w} X$ つまり $|E[f(X_{n_{j_k}})] - E[f(X)]| \longrightarrow 0$ とできてしまうので、
矛盾である。よって、$X_n \xrightarrow{w} X$.

//

Thm (中心極限定理)

$(X_n)_{n=1}^{\infty}$: 平均 μ, 分散 σ^2 の i.i.d. r.v. 列

このとき、$S_n = X_1 + \cdots + X_n$ に対し、

$$\frac{\sqrt{n}}{\sigma}\left(\frac{S_n}{n} - \mu\right) \rightsquigarrow N(0,1)$$

//

※ $\sqrt{n}\left(\frac{S_n}{n} - \mu\right) \rightsquigarrow N(0, \sigma^2)$ でもある。

※ $\frac{S_n}{n} - \mu \xrightarrow{a.s.} 0$ はわかっている。$\sqrt{n}$ 倍して膨らませると、ちょうど正規分布になる。

(Proof) $Z_n = \frac{\sqrt{n}}{\sigma}\left(\frac{S_n}{n} - \mu\right)$ とすると、

$$E[Z_n^2] = \frac{n}{\sigma^2} E\left[\left(\sum_{i=1}^{n} \frac{X_i - \mu}{n}\right)^2\right] = \frac{n}{\sigma^2}\frac{1}{n^2}\sum_{i=1}^{n} E[(X_i - \mu)^2]$$

$$= 1$$

よって、$P(|Z_n| \geqq M) \leqq \frac{E[|Z_n|^2]}{M^2} \leqq \frac{1}{M^2}$ となるので、

Z_n は緊密である。

よって、Lévy の連続性定理 2′ より、Z_n の c.f. が $N(0,1)$ の c.f. に各点収束すれば良い。

$$\underset{\substack{\parallel \\ \phi_n(t)\text{ とおく}}}{Z_n \text{ の c.f.}} = E[e^{itZ_n}] = E\left[e^{i\frac{t}{\sqrt{n}}\sum_{i=1}^{n}\frac{X_i - \mu}{\sigma}}\right] = \prod_{i=1}^{n} \phi\left(\frac{t}{\sqrt{n}}\right)$$

ϕ ← $\frac{X_i - \mu}{\sigma}$ の c.f.

ここで、$\frac{X_i - \mu}{\sigma}$ の平均は 0 で分散は 1 なので、

ϕ は 2 階連続微分可能で、

$$\phi(t) = 1 + t\underbrace{\phi'(0)}_{0} + \frac{t^2}{2}\underbrace{\phi''(0)}_{-1} + o(t^2)$$

$$= 1 - \frac{t^2}{2} + o(t^2)$$

と展開できる。よって、$\phi\left(\frac{t}{\sqrt{n}}\right) = 1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)$ であり、

$$\phi_n(t) = \phi\left(\frac{t}{\sqrt{n}}\right)^n = \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n \xrightarrow{n\to\infty} e^{-\frac{t^2}{2}}$$

$e^{-\frac{t^2}{2}}$ ← $N(0,1)$ の c.f.

よって示された。

//

Prokhorovの定理の証明.

1. は簡単なので省略

2. F_nをX_nの分布関数とする.

$\mathbb{Q}$を有理数の集合とする. $\mathbb{Q}$は可算集合なので, $\mathbb{Q}$の各元は,
$q_1, q_2, \ldots$ のように番号を振ることができる.
今 $(F_n(q_1))_{n=1}^{\infty}$ は $[0,1]$ に含まれる数列なので, ある部分列 $(F_{n_j^{(1)}}(q_1))_{j=1}^{\infty}$ が存在して, ある値に収束するようにできる.
同様に, ある部分列 $(n_j^{(2)})_{j=1}^{\infty} \subset (n_j^{(1)})_{j=1}^{\infty}$ が存在して,
$(F_{n_j^{(2)}}(q_2))_{j=1}^{\infty}$ も収束させることができる. 以下, 同様にして,

$$(n_j^{(1)})_{j=1}^{\infty} \supset (n_j^{(2)})_{j=1}^{\infty} \supset (n_j^{(3)})_{j=1}^{\infty} \supset \cdots$$

なる数列の列を得る.
ここで, $n_j = n_j^{(j)}$ とすると,
任意の $q \in \mathbb{Q}$ において, $(F_{n_j}(q))_{j=1}^{\infty}$ は収束する.
この収束先を $G(q)$ とする. Gの構成法より $G(q) \leq G(q')$ $(^\forall q \leq q')$.

しかし, Gは$\mathbb{Q}$上でのみ定義されているので, $\mathbb{R}$上に拡張する:

$$F(x) = \inf\{G(q) \mid q > x\} = \lim_{q \searrow x} G(q).$$

$\geq$ではなく$>$であることに注意.

すると, Fは単調非減少で, 右連続.
なぜなら, $^\forall \varepsilon > 0$に対し, $^\exists q \in \mathbb{Q}$で $q > x$ かつ $|F(x) - G(q)| \leq \varepsilon$ とできる.
このとき, $x \leq {}^\forall y < q$ において $F(x) \leq F(y) \leq G(q)$ かつ, $|F(x) - F(y)| \leq \varepsilon$
である. つまり, 右連続.
さらに, $\lim_{x \to \infty} F(x) = 1$, $\lim_{x \to -\infty} F(x) = 0$ でもある. これは, $(X_n)_n$ が
緊密であることから, $^\forall \varepsilon > 0$において $^\exists M > 0$で, $\sup_j F_{n_j}(-M) \leq \varepsilon$ かつ
$\inf_j F_{n_j}(M) \geq 1 - \varepsilon$ である. よって, $F(-M) \leq \varepsilon$, $F(M) \geq 1 - \varepsilon$ でもある.
$\varepsilon \searrow 0$とすることで, $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$ がわかる
以上より, Fはある分布の分布関数となる. Xをその分布に従うr.v.とする.

↳ 次ページに続く.

あとは、Fの連続点xにおいて、$F_{n_j}(x) \longrightarrow F(x)$が示せれば、

$X_{n_j} \rightsquigarrow X$が言える。

xにおいてFは連続なので、$\forall \varepsilon > 0$に対し、$\exists q < x < \exists q'$で、

$|G(q') - G(q)| \leq \varepsilon$かつ、$G(q) \leq F(x) \leq G(q')$とできる。

よって、

$$G(q) = \lim_{j \to \infty} F_{n_j}(q) \leq \liminf_{j \to \infty} F_{n_j}(x) \leq \limsup_{j \to \infty} F_{n_j}(x) \leq \lim_{j \to \infty} F_{n_j}(q') = G(q')$$

が成り立つ。よって、$\varepsilon \searrow 0$とすることで、$\lim_{j \to \infty} F_{n_j}(x) = F(x)$を得る。

//