Resnick 翻訳メモ

Sidney Resnick 氏の著書 Heavy-Tail Phenomena (Springer) は数学的には自明でない議論が含まれている。その中には著者の論文を含め、専門的な研究論文からの引用も少なくない。訳者は極値理論や確率過程についてこれまで行われてきた膨大な研究成果に精通している専門家であるとは云えないので、すべての内容を十分に理解できているとは言えないが、かなりの飛躍と判断した箇所、弱・漠収束に関して補足のメモを読者のために用意した。なおこのメモは不完全であるから適宜、更新するものとする。(2019.9.13)

翻訳メモ (国友)

1. **On Laplace Functional and Theorem 5.1(定理 5.1)**
(原著 page 134-135)

(i) $f = \lambda 1_A$ のとき $P(N(A) = k) = e^{\mu(A)}\mu(A)/k!$ より

$$
\begin{aligned}
L_n(f) &= \mathbf{E}[e^{-N(f)}] \\
&= \mathbf{E}[e^{-\lambda N(A)}] \\
&= \mathbf{E}[\sum_{k=0}^{\infty} e^{-\lambda k} e^{-\mu(A)} \frac{\mu(A)^k}{k!}] \\
&= \mathbf{E}[e^{-\mu(A)} \sum_{k=0}^{\infty} \frac{(e^{-\lambda}\mu(A))^k}{k!}] \\
&= \mathbf{E}[e^{-\mu(A)} e^{\mu(A)e^{-\lambda}}] \\
&= \mathbf{E}[\exp\{-(1 - e^{-\lambda 1_s})\mu(x)\}] \\
&= \mathbf{E}[\exp\{-\int_E (1 - e^{-\lambda f(x)})\mu(dx)\}]
\end{aligned}
$$

となる。
(ii) 一般に $1_{A_i}(x) = 0$ $(x \in A_j, j \neq i)$ より

$$
-\int_E \sum_{i=1}^{k} (1 - e^{-\lambda_i 1_{A_i}})\mu(dx) = -\int_E (1 - e^{-\sum_{i=1}^{k} \lambda_i 1_{A_i}})\mu(dx) = -\int_E (1 - e^{-f(x))})\mu(dx)
$$

となる。

(iii) (5.11) は $\tau \sim Poisson(\lambda)$ より

$$\begin{aligned}
\mathbf{E}[(\mathbf{E}(e^{-f(X_1)}))^{\tau}] &= \mathbf{E}[\sum_{k=0}^{\infty}(\int_E (e^{-f(X_1)}\mathrm{P}(X_1 \in de))^k \frac{e^{-\lambda}}{k!}\lambda^k] \\
&= e^{-\lambda}\sum_{k=0}^{\infty}\frac{1}{k!}[\lambda\int_E e^{-f(x)}\mathrm{P}(X_1 \in dx)]^k \\
&= \exp\{-\lambda + \lambda[1 - \int_E (1 - e^{-f(x)})\mathrm{P}(X_1 \in dx)]\} \\
&= \exp\{-\int_E (1 - e^{-f(x)})\lambda\mathrm{P}(X_1 \in dx)\}
\end{aligned}$$

より得られる。

2. **On Proof of Theorem 2.1(定理 2.1)**
(原著 page 35-38)

Karamata の定理の証明は $\rho < -1$ の場合について全く触れられていない。次のような議論により $\rho \geq -1$ の議論と同様に結果が得られる。この場合には $V(x) = \int_x^{\infty} U(t)dt$ とおくと。

$$\lim_{x\to\infty} V(x) < +\infty$$

である。(s が大きければ $U(2s) \leq 2^{-1-\delta}$ より

$$\int_{2^{n_0}}^{\infty} f(s)ds = \sum_{k=n_0}^{\infty}\int_{2^k}^{2^{k+1}} f(s)ds \leq \sum_{k=n_0}^{\infty} 2^{-\delta(k-n_0)}\int_{2^{n_0}}^{2^{n_0+1}} f(s)ds < \infty$$

となる。)

$$\frac{\int_x^{\infty} U(t)dt}{xU(x)} = \int_1^{+\infty}\frac{U(xt)}{U(x)}dt$$

より $V(x) \in RV_{\rho+1}$ である。$\rho < -1$ のとき積分を評価すると

$$\int_1^{\infty} s^{\rho}ds = [\frac{s^{\rho+1}}{\rho+1}]_1^{\infty} \to -\frac{1}{\rho+1}$$

より結果が得られる。

証明補足メモ (栗栖)

1. (Proposition 6.2 の証明補足) まず距離空間 $M_+(\mathbb{E})$($M_+(\mathbb{E})$ には漠収束により距離が導入できる) の開集合族が生成するボレル σ-代数上の確率測度

の集合 (例えば $\mathrm{Pr}(M_+(\mathbb{E}))$ とする) を考える．その集合の中の確率測度の列の弱収束により $\mathrm{Pr}(M_+(\mathbb{E}))$ 導入される距離を d_W としたときの距離空間 $(\mathrm{Pr}(M_+(\mathbb{E})), d_W)$ が本書において $\mathrm{PR}(M_+(\mathbb{E}))$ と定義されていることに注意する．また $M_p(\mathbb{E})$ は $M_+(\mathbb{E})$ の閉部分集合であることにも注意する．
Proposition 6.2 の証明において, ほとんどすべての ω に対して

$$\mathbb{P}\left[\sum_{i=1}^{m(n')} \epsilon_{Z_i^*/b_m(n')} \in \cdot | Z_1, \ldots, Z_{n'}\right] \Rightarrow \mathbb{P}[\mathrm{PRM}(\nu) \in \cdot] \text{ in } \mathrm{PR}(M_p(\mathbb{E}))$$

が成り立つ ($\{n'\}$ は $\{n\}$ の任意の部分列 $\{n''\}$ に対してとることのできるある部分列)．ここで $\mathbb{P}[\mathrm{PRM}(\nu) \in \cdot]$ は $\mathrm{PR}(M_p(\mathbb{E}))$ における点 (定数) であり, (同一確率空間上で定義された確率変数列の) 定数への弱収束はその定数への確率収束と同値であるので, 上記の結果は

$$\mathbb{P}\left[\sum_{i=1}^{m(n')} \epsilon_{Z_i^*/b_m(n')} \in \cdot | Z_1, \ldots, Z_{n'}\right] \xrightarrow{p} \mathbb{P}[\mathrm{PRM}(\nu) \in \cdot] \text{ in } \mathrm{PR}(M_p(\mathbb{E}))$$

を意味する．また確率収束列の任意の部分列 $\{n'_k\}$ は概収束する部分列 $\{n'_{k_\ell}\}$ をもつので,

$$\mathbb{P}\left[\sum_{i=1}^{m(n'_{k_\ell})} \epsilon_{Z_i^*/b_m(n'_{k_\ell})} \in \cdot | Z_1, \ldots, Z_{n'_{k_\ell}}\right] \xrightarrow{a.s.} \mathbb{P}[\mathrm{PRM}(\nu) \in \cdot] \text{ in } \mathrm{PR}(M_p(\mathbb{E}))$$

が成り立つ．ここで, $\{n''\}$ は $\{n\}$ の任意の部分列で, 概収束する部分列 $\{n'_{k_\ell}\} \subset \{n''\}$ を含むので, もともとの列 $\{n\}$ についての確率収束

$$\mathbb{P}\left[\sum_{i=1}^{m} \epsilon_{Z_i^*/b_m} \in \cdot | Z_1, \ldots, Z_n\right] \xrightarrow{p} \mathbb{P}[\mathrm{PRM}(\nu) \in \cdot] \text{ in } \mathrm{PR}(M_p(\mathbb{E}))$$

が成り立つ．特に, $Y_n := \mathbb{P}\left[\sum_{i=1}^{m} \epsilon_{Z_i^*/b_m} \in \cdot | Z_1, \ldots, Z_n\right]$ は一様可積分なので, その期待値は $\mathrm{PR}(M_p(\mathbb{E}))$ において $Y := \mathbb{P}[\mathrm{PRM}(\nu) \in \cdot] = E[Y]$ に収束する．よって, $\lim_{n\to\infty} d_W(E[Y_n], E[Y]) = 0$. 即ち,

$$\mathbb{P}\left[\sum_{i=1}^{m} \epsilon_{Z_i^*/b_m} \in \cdot\right] \Rightarrow \mathbb{P}[\mathrm{PRM}(\nu) \in \cdot] \text{ in } \mathrm{PR}(M_p(\mathbb{E})).$$

2. (Proposition 7.6 の証明補足)

   (a) 確率変数 Y について, $1 - F_Y(\cdot) = P(Y > \cdot) \in \mathrm{RV}_{-\alpha_Y}$ とする．また $b_Y(t) = \left(\frac{1}{1-F_Y}\right)^{\leftarrow}(t)$ とする．このとき, $x \to \infty$ として $\frac{1}{1-F_Y(x)} \to \infty$ なので, Proposition 2.6 (v) より $b_Y(t) \in \mathrm{RV}_{1/\alpha_Y}$.

(b) $b_Y(t)b_Z(t)\to\infty(t\to\infty)$ なので, 任意の $t>0$ に対してある $t_0\in\mathbb{R}$ が存在して $b_Y(t_0-1)b_Z(t_0-1)<t\le b_Y(t_0)b_Z(t_0)$. よって,

$$
\begin{aligned}
&(b_Yb_Z)^{\leftarrow}(t)\mathbb{P}\left[\frac{YZ}{t}\in[\mathbf{0},\mathbf{x}]^c\right]\\
&\le(b_Yb_Z)^{\leftarrow}(b_Y(t_0)b_Z(t_0))\mathbb{P}\left[\frac{YZ}{b_Y(t_0-1)b_Z(t_0-1)}\in[\mathbf{0},\mathbf{x}]^c\right]\\
&=(t_0-1)\mathbb{P}\left[\frac{YZ}{b_Y(t_0-1)b_Z(t_0-1)}\in[\mathbf{0},\mathbf{x}]^c\right]+o(1)\\
&\overset{t_0\to\infty}{\longrightarrow}\nu\left(\{(y,z):yz\in[\mathbf{0},\mathbf{x}]^c\}\right)=:\lambda(\mathbf{x}).
\end{aligned}
$$

下からの評価も同様に行うことができて, 結局,

$$\frac{1-F_{YZ}(t\mathbf{x})}{[1/(b_Yb_Z)^{\leftarrow}](t)}\overset{t\to\infty}{\longrightarrow}\lambda(\mathbf{x}).$$

(c) $b_Y\in\mathrm{RV}_{1/\alpha_Y}$, $b_Z\in\mathrm{RV}_{1/\alpha_Z}$ なので, $b_Yb_Z\in\mathrm{RV}_{1/\alpha_Y+1/\alpha_Z}=\mathrm{RV}_{\frac{\alpha_Y+\alpha_Z}{\alpha_Y\alpha_Z}}$. $b_Y(t)b_Z(t)\to\infty(t\to\infty)$ なので, Proposition 2.6 (v) より, $(b_Yb_Z)^{\leftarrow}(t)\in\mathrm{RV}_{\frac{\alpha_Y\alpha_Z}{\alpha_Y+\alpha_Z}}$. 更に $[1/(b_Yb_Z)^{\leftarrow}](t)\in\mathrm{RV}_{-\frac{\alpha_Y\alpha_Z}{\alpha_Y+\alpha_Z}}$. 最後に, (多変量) 正則変動関数の定義より, $1-F_{YZ}(\cdot)=P(YZ>\cdot)\in\mathrm{RV}_{-\frac{\alpha_Y\alpha_Z}{\alpha_Y+\alpha_Z}}$.

翻訳補足メモ (栗栖)

1. (3 章に関するコメント) 弱収束と漠収束の定義から, $C_K^+(\mathbb{R}^d)\subset C_b(\mathbb{R}^d)$ なので, (確率測度とは限らない) 測度の列に対する収束の概念としては弱収束の方がより仮定が厳しい. 即ち, 測度の列 $\mu,\mu_1,\mu_2,\ldots$ について, $\mu_n\Rightarrow\mu$ ならば $\mu_n\xrightarrow{v}\mu$. この逆は一般には成り立たない. 実際, δ_n を点 n 上のディラック測度とすると, 明らかに $\delta_n\xrightarrow{v}0$ であるが, $f(x)=\sin(x)\in C_b(\mathbb{R})$ とすると $\delta_n(f)=\sin(n)$ は収束しないので弱収束しない. ここで取り上げたディラック測度の例が弱収束しない理由としては $\{\delta_n\}$ が緊密 (tight) でないことが挙げられる.

   Definition 1 (緊密性の定義)**.** (S,ρ) をある距離空間でボレル σ-代数 $\mathcal{S}$ をもつとする. (S,ρ) 上の測度の列 $\{\mu_n\}$ が緊密 *(tight)* であるとは, 任意の $\epsilon>0$ に対してあるコンパクト集合 $K=K(\epsilon)\subset S$ が存在して, 任意の n に対して $\mu_n(K^c)<\epsilon$ が成り立つことをいう.

   定義から, 測度の列が緊密であることは“ほとんど”の測度があるコンパクト集合上に集中していることを意味する. ディラック測度の例では $n\to\infty$

としても $\mathbb{R}$ の“端の部分”に正の測度をもつので tight ではない (∞ に測度が逃げる). これにより漠収束する測度の列が弱収束するには緊密であることが必要であることがわかる. ただし, 確率測度の列 $\{\mu_n\}_{n\geq 0}$(μ_0 も確率測度であることに注意. ここがディラック測度の例と異なる) に対して $\mu_n \stackrel{v}{\to} \mu_0$ であることと $\mu_n \Rightarrow \mu_0$ であることは同値 (Theorem 4.4.2 in Chung (2000)). また, 本書の点過程の理論で扱う測度の列の極限は有限測度とは限らないため, 一般に tight ではない. このような事情から弱収束ではなく漠収束を考える.

緊密性は確率測度の列に対する測度の (一様な) 集中の程度についての性質であるが, 確率測度の緊密性は完備可分距離空間 (S, ρ) 上の確率測度全体 $\mathcal{P}(S)$ を, 弱収束と同値な位相 (弱位相) を備えた位相空間としてみたときのある種の性質が同値であることについて紹介する.

Theorem 1 (プロホロフ (Prohorov) の定理). *(i)* 確率測度の列の全体 $\Pi \subset \mathcal{P}(S)$ が緊密であるための必要十分条件は, 弱位相を備える空間 $\mathcal{P}(S)$ において Π の閉包が点列コンパクト *(*空間内の任意の点列が収束する部分列を含む*)* である.

(ii) (S, ρ) が完備可分距離空間のとき, 弱収束と同値であるような $\mathcal{P}(S)$ 上のある完備距離 d が存在し, $\Pi \subset \mathcal{P}(S)$ が緊密であるための必要十分条件は, K の閉包が $(\mathcal{P}(S), d)$ においてコンパクトであることである.

特に, 距離空間においてはコンパクト, 点列コンパクト, 相対コンパクト (閉包をとればコンパクト) は同値であるので, プロホロフの定理 (ii) は, $\Pi \subset \mathcal{P}(S)$ の緊密性と点列コンパクト性 (全ての $\{P_n\} \subset \Pi$ は弱収束する部分列を含む) ことと同値.

なお、確率測度 (probability measures) の弱収束 (weak convergence) については Billingsly (1968) が詳しい。

2. (3.3.4, 10.2 に関するコメント) 任意の $x \in D[0,1]$ は有界であるから $D[0,1]$ 空間には $C[0,1]$ 空間と同様に一様距離

$$d_{sup}(x, y) = \sup_{t\in[0,1]} |x(t) - y(t)|, \ x, y \in D[0,1]$$

を入れることができる. しかし d_{sup} は $D[0,1]$ において“よい”距離であるとは言えない. というのも, (D, d_{sup}) は完備距離空間になるが, 可分ではない. (D, d_{sup}) が完備であることは一様距離を入れた距離空間 $(C[0,1], d_{sup})$

が完備であることの証明と同様にして示せる．可分でないことは, 例えば $0 < \theta < 1$ に対して

$$x_\theta(t) = \begin{cases} 0 & \text{if } 0 \le t < \theta \\ 1 & \text{if } \theta \le t \le 1 \end{cases}$$

とおくと $d_{sup}(x_{\theta_1}, x_{\theta_2}) = 1$, $\theta_1 \neq \theta_2$．もし (D, d_{sup}) が可分ならば $D_0 := \{x_\theta : 0 < \theta < 1\} \subset D$ とすると D_0 も可分になるはずであるが, 明らかに可分ではないので矛盾．故に (D, d_{sup}) は可分ではない．さらに統計学との関連としては $\mathcal{U}$ を一様距離に関する D の Borel σ-field とすると, 経験分布関数が標本空間 $(\Omega, \mathcal{F})$ から $(D, \mathcal{U})$ への写像として一般に $\mathcal{F}/\mathcal{U}$ 可測にならないことが知られている．

スコロホッド距離が誘導する $D[0,1]$ の位相 (スコロホッド位相) は $D[0,1]$ 上の J_1-位相とも呼ばれる．この呼び名はスコロホッドによる論文 (Skorohod (1956)) による．$C[0,1] \subset D[0,1]$ であるから, J_1-位相を用いて $C[0,1]$ に相対位相を導入することができるが, この位相は一様距離により $C[0,1]$ に導入される位相と一致することがわかる．この事実により, $C[0,1]$ は $D[0,1]$ 内の閉集合であり, $x \in D[0,1] \backslash C[0,1]$ は連続関数の列 $\{x_n\} \subset C[0,1]$ では近似できないことがわかる．このことにより不都合が生じる場合もある．例えば $D[0,1]$ の要素の畳み込みのような関数のスムージングを考える場合, このような操作は J_1-位相のもとでは連続にならない．このような問題を解決するため, Skorohod (1956) では J_1-位相よりも“弱い”位相として $D[0,1]$ 上の J_2, M_1, M_2-位相が提案されている．

Definition 2. 集合 X 上で定義された *2* つの位相空間 $(X, \mathcal{O}_1)$, $(X, \mathcal{O}_2)$ について, $\mathcal{O}_1$ が $\mathcal{O}_2$ よりも弱い位相であるとは, $\mathcal{O}_1 \subset \mathcal{O}_2$ が成り立つことである．

2 つの位相 $\mathcal{O}_1$, $\mathcal{O}_2$ について $\mathcal{O}_1 \subset \mathcal{O}_2$ を $\mathcal{O}_1 \to \mathcal{O}_2$ と書くことにすれば, U を d_{sup} が誘導する位相として, 以下の関係が成り立つ:

$$\begin{tikzcd} & J_2 \arrow[dr] & & \\ M_2 \arrow[ur] \arrow[dr] & & J_1 \arrow[r] & U. \\ & M_1 \arrow[ur] & & \end{tikzcd}$$

3. (6.1.3 に関するコメント) 点過程の理論では, ラドン測度全体の中の漠収束を考える．典型的には閾値超過モデルのように $\{x > x_0\}$, $x_0 > 0$ 上の測度の (漠) 収束 (特に数列としての収束), 例えば $\{\mu_n\}_{n \ge 0}$ を $[0, \infty)$ 上の測度の列として, $\mu_n(\{x > x_0\}) \to \mu_0(\{x > x_0\})$, $n \to \infty$ を考えたいので

$\{x > x_0\}$ に対応する集合がコンパクト集合である必要がある．そのために $[0,\infty)$ に一点 $\{\infty\}$ を付け加えてコンパクト化した $[0,\infty]$ を考える (この操作を一点コンパクト化という)．その際, 元々の $[0,\infty)$ 上の位相を拡大して $[0,\infty]$ がコンパクトになるような位相を導入して新たな位相空間を作る．一方, データが裾の重い分布に従う場合, 興味のある統計量の分布 (を変換したもの) の $\{x > x_0\}$ 上での測度は適当な条件の下である α-安定過程のレヴィ測度 ν_α に収束する．ここで $\nu_\alpha([0,\infty]) = \infty$ ($[0,\infty]$ はコンパクト集合) なので, このようなコンパクト集合が含まれる位相空間上では $[0,\infty]$ 上のラドン測度の漠収束を議論できない．従って $[0,\infty]$ から $\{0\}$ を除いて一点非コンパクト化した $(0,\infty]$ を考え, $[0,\infty]$ の位相の相対位相を入れて再び新たな位相空間を考える．集合・位相論に馴染みのない読者のために, 入門書として松坂 (1968), 内田 (1986) を挙げておく．また矢野 (1997) には距離空間を含む位相空間の具体例が多く紹介されている．ここでは一点コンパクト化の具体例として $\mathbb{R}^d$ から $\mathbb{R}^{d+1}$ 内の単位球面 S^d への射影 (立体射影) により $\mathbb{R}^d$ を一点コンパクト化する方法を紹介する (松本 (1988)).

Definition 3 (位相空間のコンパクト化の定義)**.** *X, X' を位相空間, ι を X から X' への写像とする．次の (i)-(iii) が成り立つとき, (X',ι) または X' を X のコンパクト化という．*

(i) X' はコンパクト．

(ii) ι の値域を $\iota(X)$ に制限して得られる X から $\iota(X)$ への写像は同相写像 (全単射かつその逆写像も連続).

(iii) $\iota(X)$ は X において稠密 ($\iota(X)$ の閉包は X').

R^{d+1} には通常の位相を入れ, S^d には相対位相を入れる．

$$S^{d+1} = \{x = (x_1,\ldots,x_{d+1}) \in \mathbb{R}^{d+1} : x_1^2 + \cdots + x_{d+1}^2 = 1\}$$

を d 次元球面という．定義より, S^d は $\mathbb{R}^{d+1}$ の有界閉集合であるから S^d はコンパクト．S^d 上の点 $p_1 = (0,\ldots,0,1)$ を北極と呼ぶことにする．このとき, p_1 から S^d の点 x について, p_1 と x を通る直線を考え, x_{d+1} で与えられる d 次元空間 $\mathbb{R}^d$ と直線との交点を $\varphi(x)$ とする．$U = S^d\backslash\{p_1\}$ とすると $\varphi(x) : U \to \mathbb{R}^d$ は以下で与えられることが知られている:

$$\varphi(x) = \left(\frac{x_1}{1-x_{d+1}},\ldots,\frac{x_d}{1-x_{d+1}}\right) \in \mathbb{R}^d.$$

また φ の逆写像 $\varphi^{-1} : \mathbb{R}^d \to S^d$ は

$$\varphi^{-1}(y) = \left(\frac{2y_1}{\|y\|^2+1},\ldots,\frac{2y_d}{\|y\|^2+1},\frac{\|y\|^2-1}{\|y\|^2+1}\right) \in S^d$$

で与えられる．ここで $\|y\| = (y_1^2 + \cdots + y_d^2)^{1/2}$.

(i) S^d はコンパクト．

(ii) φ^{-1} の値域を U に制限して得られる $\mathbb{R}^d$ から U への写像は同相写像.

(iii) U は S^d において稠密 (U の閉包は S^d).

従って (S^d, φ^{-1}) は $\mathbb{R}^d$ の (一点) コンパクト化.

参考文献

内田伏一 (1986) 『集合と位相』 裳華房.

松坂和夫 (1968) 『集合・位相入門』 岩波書店.

松本幸夫 (1988) 『多様体の基礎』 東京大学出版.

矢野公一 (1997) 『距離空間と位相構造』 共立出版.

Billingsley, P. (1968) *Convergence of Probability Measures*. Wiley.

Chung, K.L. (2000) *A Course in Probability Theory*. Academic Press 3rd Ed.

Skorohod. A.V. (1956) Limit theorems for stochastic processes. *Theor. Probability Appl.* **1**, 261-290.