

Skorohod の表現定理

@schrott512

1 はじめに

本記事では Skorohod の表現定理について紹介する. Skorohod の定理の主張を簡単に言えば, 確率測度の列がある確率測度に収束するならば, 適当な確率空間上にそれぞれに対応する分布を持つ確率変数列と確率変数を同時に定義することができて, さらにその列が概収束するようにすることができる, というものである.

もっと大雑把だがより分かりやすい表現を目指すなら, 次のようにいうことができる. 分布収束は対象の確率変数たちが必ずしも同じ確率空間上にいる必要はないが, それゆえ取り扱いが不便なこともある. しかし, 分布収束するなら共通の確率空間へ引っ越しできる, というものである. このことは応用例の節で具体的にみる.

2 準備

ここでは, 諸事項をおさらいしておこう.

Definition 2.1: 確率測度の弱収束

μ, μ_1, μ_2, \dots を可測空間 S 上の確率測度とする. 列 $(\mu_n)_n$ が μ に弱収束するとは, S 上の任意の有界連続関数 f に対し,

$$\lim_{n \rightarrow \infty} \int f d\mu_n = \int f d\mu$$

が成り立つことをいう. またこのとき, $\mu_n \Rightarrow \mu$ ($n \rightarrow \infty$) とかく.

また, S -値確率変数 X, X_1, X_2, \dots とする. $(X_n)_n$ の分布の列が X の分布に弱収束するとき, $X_n \Rightarrow X$ ($n \rightarrow \infty$) とかき, X_n が X に分布収束するという.

この先の議論において測度の台が登場するが, 次のものが一般的な定義のようである.

Definition 2.2: 測度の台-1

S を位相空間, μ を S 上の測度とする. $x \in S$ で, その任意の開近傍 $U \subset S$ に対し $\mu(U) > 0$ となるものの全体を μ の台と呼び, $\text{supp}(\mu)$ で表す.

定義より直ちに μ の台が S の閉集合であることが分かる. また, 次が知られている.

Lemma 2.3

S を位相空間とし, μ を S 上の測度とする. また, 次をおく:

- (i) S は可算開基を持つ.
- (ii) S は局所コンパクトハウスドルフ空間で, かつ μ はラドン測度である.

このとき, (i) または (ii) が成り立つならば,

$$\text{supp}(\mu) = \bigcap_{\substack{C: \text{closed} \\ \mu(S \setminus C) = 0}} C \quad (2.1)$$

である.

参考文献において測度の台が現れるが, どのような定義を採用しているか見つからなかった. しかし, 議論の中で (2.1) の性質が用いられている. S 自体に (i),(ii) を仮定してもよいが, 広く普及していると思われる参考文献中の定理のステートメントに合わせるため, 本記事では台の定義は次のものを採用することにする.

Definition 2.4: 測度の台-2

S を位相空間, μ を S 上の測度とする. Lemma 2.3 の (2.1) を μ の台と呼ぶ.

例えば, S が可分距離空間の場合は, (i) が満たされ, Definition 2.2, 2.4 は一致する. 確率測度の弱収束に関するよく知られた事実を参照しておく.

Proposition 2.5

μ, μ_1, μ_2, \dots を距離空間 S 上の確率測度とする. このとき次の (1),(2) は同値である.

- (1) $\mu_n \Rightarrow \mu$ ($n \rightarrow \infty$).
- (2) S のボレル集合 A が $\mu(\partial A) = 0$ を満たすならば, $\lim_{n \rightarrow \infty} \mu_n(A) = \mu(A)$ である.

3 Skorohod の表現定理

(S, d) を距離空間とし, $x \in S$ を中心とする ε -開球を $B(x, \varepsilon)$ で表す. また, S のボレル集合族を $\mathcal{B}(S)$ で表す.

Theorem 3.1: Skorohod の表現定理

μ, μ_1, μ_2, \dots を S 上の確率測度とし, μ の台は可分とする. このとき, $\mu_n \Rightarrow \mu$ ($n \rightarrow \infty$) ならば, ある確率空間 (Ω, \mathcal{F}, P) 上に S -値確率変数 X, X_1, X_2, \dots を定義でき, さらに各 X_n と X の分布がそれぞれ μ_n, μ でかつ $X_n \rightarrow X$ ($n \rightarrow \infty$), P -a.s. となるようにできる.

主張にある収束は S の位相における収束であることに注意する.

この定理の証明に用いる次の補題を先に示す.

Lemma 3.2

μ を台が可分な S 上の確率測度とする. このとき, 任意の $\varepsilon > 0$ に対し, S のある有限個の非交差かつ $S = \bigcup_{i=0}^k B_i$ を満たす $B_0, B_1, \dots, B_k \in \mathcal{B}(S)$ で, 次を満たすものが存在する.

- (i) $\mu(B_0) < \varepsilon$,
- (ii) $\mu(\partial B_i) = 0, \quad i = 0, 1, \dots, k$,
- (iii) $\text{diam } B_i := \sup_{x, y \in B_i} d(x, y) < \varepsilon, \quad i = 1, 2, \dots, k$.

Proof. $\varepsilon > 0$ とし, $M = \text{supp}(\mu)$ とおく. $x \in M$ を任意に取る. このとき, 各 $j = 1, 2, \dots$ に対して $\mu(\partial B(x, r)) \geq 1/j$ を満たす $0 < r < \varepsilon/2$ は高々 j 個なので, $\mu(\partial B(x, r)) > 0$ を満たす $0 < r < \varepsilon/2$ は高々可算個である. よって $\mu(\partial B(x, r)) = 0$ を満たす $0 < r < \varepsilon/2$ は無数に存在する.

ここで, M が可分であることと上に示した事実により, M は $\mu(\partial B(x, r)) = 0$ を満たす高々可算個の $(B(x, r))_{x \in M, 0 < r < \varepsilon/2}$ で被覆される. その被覆を A_1, A_2, \dots とする. $\mu(\bigcup_{i \geq 1} A_i) \geq \mu(M) = 1$ と測度の連続性より, 十分大きな k に対して

$$\mu\left(\bigcup_{i=1}^k A_i\right) > 1 - \varepsilon \quad (3.1)$$

とできる. $B_0 = S \setminus \bigcup_{i=1}^k A_i$ とし,

$$\begin{cases} B_1 = A_1, \\ B_i = A_i \setminus \bigcup_{j=1}^{i-1} A_j, \quad i = 2, 3, \dots, k \end{cases}$$

とおく. $\mu(B_0) = 1 - \mu(\bigcup_{i=1}^k A_i)$ と (3.1) より (i) が分かり, また各 i に対し $B_i \subset A_i$ より (iii) を得る. そして各 i に対して $\partial B_i \subset \bigcup_{i=1}^k \partial A_i$ と $\mu(\partial A_i) = 0$ より, (ii) が得られる. \square

Proof of Theorem 3.1. $m = 1, 2, \dots$ に対し $\varepsilon_m = 1/2^m$ とおく. Lemma 3.2 より, 各 m に対し, 非交差な $B_0^m, \dots, B_{k_m}^m \in \mathcal{B}(S)$, $S = \bigcup_{i=0}^{k_m} B_i^m$ で,

- (i) $\mu(B_0^m) < \varepsilon_m$,
- (ii) $\mu(\partial B_i^m) = 0, \quad i = 0, 1, \dots, k_m$,
- (iii) $\text{diam } B_i^m < \varepsilon_m, \quad i = 1, 2, \dots, k_m$

とする. ここで, もし $\mu(B_i^m) = 0$ であるような B_i^m ($i \geq 1$) があれば, それらすべてと B_0^m との合併を新たに B_0^m をみなすことで (i)-(iii) が満たされ, また合併されたもの以外の B_i^m については $\mu(B_i^m) > 0$ である. したがって, 最初からすべての B_i^m ($i \geq 1$) が $\mu(B_i^m) > 0$ を満たすと仮定してよい.

$\mu_n \Rightarrow \mu$ ($n \rightarrow \infty$) なので, 各 m に対し (ii) と Proposition 2.5 より, 適当な n_m が存在して, $n \geq n_m$ ならば

$$\mu_n(B_i^m) \geq (1 - \varepsilon_m)\mu(B_i^m), \quad i = 0, 1, \dots, k_m$$

である. ここで適当に番号を取り換えて, $1 < n_1 < n_2 < \dots$ としておく.

$(\Omega^1, \mathcal{F}^1, P^1) = (S, \mathcal{B}(S), \mu)$, $(\Omega_n^2, \mathcal{F}_n^2, P_n^2) = (S, \mathcal{B}(S), \mu_n)$ ($n \geq 1$) とおく. また, $n \geq n_1, i \geq 1$ と $A \in \mathcal{B}(S)$ に対し,

$$\begin{cases} \nu_{n,i}(A) = \mu_n(A | B_i^m) = \frac{\mu_n(A \cap B_i^m)}{\mu_n(B_i^m)}, \\ \nu_n(A) = \varepsilon_m^{-1} \sum_{i=0}^{k_m} \mu_n(A | B_i^m) \{\mu_n(B_i^m) - (1 - \varepsilon_m)\mu(B_i^m)\}, \end{cases} \quad \text{if } n_m \leq n < n_{m+1}$$

で定め, $(\Omega_{n,i}^3, \mathcal{F}_{n,i}^3, P_{n,i}^3) = (S, \mathcal{B}(S), \nu_{n,i})$ ($n \geq n_1, i \geq 1$), $(\Omega_n^4, \mathcal{F}_n^4, P_n^4) = (S, \mathcal{B}(S), \nu_n)$ ($n \geq n_1$) とおく. また, $(\Omega^5, \mathcal{F}^5, P^5) = ([0, 1], \mathcal{B}([0, 1]), \lambda)$ とおく. ただし, λ は \mathbb{R} 上のルベーグ測度である.

さて, (Ω, \mathcal{F}, P) をこれらすべての直積確率空間, すなわち

$$\begin{aligned} (\Omega, \mathcal{F}, P) &= (\Omega^1, \mathcal{F}^1, P^1) \times \prod_{n \geq 1} (\Omega_n^2, \mathcal{F}_n^2, P_n^2) \times \prod_{\substack{n \geq n_1 \\ i \geq 1}} (\Omega_{n,i}^3, \mathcal{F}_{n,i}^3, P_{n,i}^3) \times \prod_{n \geq n_1} (\Omega_n^4, \mathcal{F}_n^4, P_n^4) \times (\Omega^5, \mathcal{F}^5, P^5) \end{aligned}$$

とおく. そして $\omega = (\omega^1, \omega_1^2, \dots, \omega_n^2, \dots, \omega_{n,i}^3, \dots, \omega_n^4, \dots, \omega^5) \in \Omega$ と $n \geq 1$ に対し,

$$\begin{aligned} X(\omega) &= \omega^1, \\ X_n(\omega) &= \begin{cases} \omega_n^2, & n < n_1 \\ \sum_{i=0}^{k_m} \mathbf{1}_{\{\omega^5 \leq 1 - \varepsilon_m, \omega^1 \in B_i^m\}} \omega_{n,i}^3 + \mathbf{1}_{\{\omega^5 > 1 - \varepsilon_m\}} \omega_n^4, & n_m \leq n < n_{m+1} \end{cases} \end{aligned}$$

と定める. このとき X の分布は μ である. また $n_m \leq n < n_{m+1}$ と $A \in \mathcal{B}(S)$ に対し,

$$\begin{aligned} P(X_n \in A) &= \sum_{i=0}^{k_m} (1 - \varepsilon_m) \mu(B_i^m) \nu_{n,i}(A) + \varepsilon_m \nu_n(A) \\ &= \mu_n(A) \end{aligned}$$

であり, $n < n_1$ のときは定め方から明らかに $P(X_n \in A) = \nu_n(A)$ である. したがって, すべての $n \geq 1$ について X_n の分布は μ_n である.

$m \geq 1$ に対し $E_m = \{\omega^1 \notin B_0^m, \omega^5 \leq 1 - \varepsilon_m\}$, $E = \liminf_{m \rightarrow \infty} E_m = \bigcup_{j \geq 1} \bigcap_{k \geq j} E_k$ とする. $\omega \in E_m$, $n_m \leq n < n_{m+1}$ とすると, ある i に対し $\omega^1 \in B_i^m$ であり, このとき定め方から $X_n(\omega) = \omega_{n,i}^3$ である. $\omega_{n,i}^3 \in B_i^m$ ならば, (iii) より $d(X_n(\omega), X(\omega)) < \varepsilon_m$ である. すなわち

$$E_m \subset \bigcap_{n_m \leq n < n_{m+1}} \bigcup_{i=1}^{k_m} (\{d(X_n(\omega), X(\omega)) < \varepsilon_m\} \cup \{\omega^1 \in B_i^m, \omega_{n,i}^3 \notin B_i^m\}) \quad (3.2)$$

である. また, $P(\omega^1 \in B_i^m, \omega_{n,i}^3 \notin B_i^m) = \mu(B_i^m) \nu_{n,i}(S \setminus B_i^m) = 0$ である. したがって,

$$E \subset \left\{ \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega) \right\} \cup N \quad (3.3)$$

を得る. ただし, $N \in \mathcal{F}$ は (3.2) から得られる P -零集合である.

$$\begin{aligned}\sum_{m=1}^{\infty} P(S \setminus E_m) &\leq \sum_{m=1}^{\infty} (\mu(B_0^m) + \varepsilon_m) \\ &\leq \sum_{m=1}^{\infty} \frac{1}{2^{m-1}} < \infty\end{aligned}$$

なので、ボレル-カンテリの補題より $P(E) = 1$ である。これと (3.3) より、 $\lim_{n \rightarrow \infty} X_n = X$, P -a.s. を得る。 \square

4 応用例

$C([0, \infty); \mathbb{R})$ を $[0, \infty)$ 上で定義された実数値連続関数全体とし、広義一様収束位相を入れた空間とする。

ξ_1, ξ_2, \dots を期待値 0, 分散 1 の独立同分布確率変数列とし、 $S_n = \xi_1 + \dots + \xi_n$ とおく。このとき、

$$W_t^n = \frac{S_{[nt]}}{\sqrt{n}} + (nt - [nt]) \frac{\xi_{[nt]+1}}{\sqrt{n}}, \quad t \geq 0$$

とおくと、 $W^n = (W_t^n)_{t \geq 0} \in C([0, \infty); \mathbb{R})$ である。ただし、 $[x]$ は x を超えない最大の整数を表す。したがって W^n は $C([0, \infty); \mathbb{R})$ 上に確率測度を誘導し、その像測度を $P_{W^n} = P \circ (W^n)^{-1}$ とする。

次の定理が知られている。

Theorem 4.1: Donsker の不変原理

$C([0, \infty); \mathbb{R})$ の位相のもと、 $W^n \Rightarrow W$ ($n \rightarrow \infty$) である。ただし、 $W = (W_t)_{t \geq 0}$ は 1 次元標準ブラウン運動である。すなわち、 P_{W^n} は 1 次元標準 Wiener 測度に弱収束する。

さて、上の定理と Theorem 3.1 より、共通の確率空間 (Ω, \mathcal{F}, P) がとれて、この確率空間上に ξ_1, ξ_2, \dots と 1 次元標準ブラウン運動 W が定義されているとみなすことができ、任意の $T > 0$ に対し、

$$\sup_{0 \leq t \leq T} |W_t^n - W_t| \rightarrow 0, \quad n \rightarrow \infty, \quad P\text{-a.s.}$$

が成り立つ。