

確率数理要論 2

2. 確率変数と期待値

- 確率変数

$(\Omega, \mathcal{F}, P)$: 確率空間

Def (確率変数)

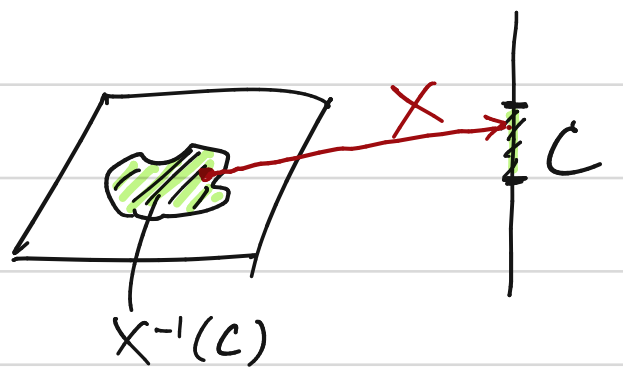
$X: \Omega \rightarrow \mathbb{R}$ が、任意の Borel 集合 $C \in \mathcal{B}(\mathbb{R})$ に対し、
 $X^{-1}(C) \in \mathcal{F}$ をみたすとき、 X を 確率変数 と言う。
(random variable)
r.v.

* $X^{-1}(C) := \{\omega \in \Omega \mid X(\omega) \in C\}$ とする。

$X^{-1}(C) \in \mathcal{F}$ ($C \in \mathcal{B}(\mathbb{R})$) なので、
 $P(X^{-1}(C))$ が定まる。これを、

$$P(X \in C) = P(X^{-1}(C)) = P(\{\omega \in \Omega \mid X(\omega) \in C\})$$

と書く。



Rem

より一般に可測空間 $(\Omega, \mathcal{F})$, $(\Omega', \mathcal{F}')$ に対し、

$X: \Omega \rightarrow \Omega'$ が、 $\forall C \in \mathcal{F}'$ に対し $X^{-1}(C) \in \mathcal{F}$ をみたすとき
 X を 可測関数 と言う。(より詳しく $\mathcal{F}/\mathcal{F}'$ -可測とも言ふ)
(= Ω' -値確率変数とも言ふ)

$$X^{-1}(\mathcal{F}') := \{X^{-1}(C) \mid C \in \mathcal{F}'\} \quad \text{と書く。}$$

$$X^{-1}(\mathcal{F}') \subset \mathcal{F} \iff X: \text{可測} \quad \text{とも言ふ。}$$

Ex. (コイン投げ)

$\Omega = \{\text{裏}, \text{表}\}$, $\mathcal{F} = 2^\Omega$ (部分集合全体)

$X(\text{表}) = 1$, $X(\text{裏}) = 0$

$\Rightarrow X$ は確率変数

Ex. (無限回のコイン投げ)

$\Omega = \{0, 1\}^\infty$, $\omega = (\omega_1, \omega_2, \dots)$, $\omega_i \in \{0, 1\}$

$X_n(\omega) = \omega_n$

とおく.

$\mathcal{F} : \forall n \in \mathbb{N}$ X_n が可測となる最小の σ -加法族とおく.

$\Rightarrow X_n$ は確率変数

※ $\Omega_n = \{0, 1\}$ とおくと、

$\{B_1 \times B_2 \times \dots \times B_n \times \Omega_{n+1} \times \Omega_{n+2} \times \dots \mid B_i \subset \Omega_i, n=1, 2, \dots\}$
を含む最小の σ -加法族が $\mathcal{F}$.

より一般に

Def (直積 σ -加法族)

あるインデックス集合 T (非可算でも良い) があって、 $(T \neq \emptyset)$

$\forall t \in T$ に対し、 $(\Omega_t, \mathcal{F}_t)$ が付随しているとする

このとき、 $\prod \Omega_t$ 上の直積 σ -加法族 は以下のように

定義される。 $t \in T$ 直積集合: $\{(\omega_t)_{t \in T} \mid \omega_t \in \Omega_t\}$

$$\bigotimes_{t \in T} \mathcal{F}_t := \sigma \left(\left\{ \prod_{t \in T} B_t \mid B_t \in \mathcal{F}_t (\forall t \in T) \right. \right. \\ \left. \left. \begin{array}{l} \text{ただし有限個の } t \text{ の } \omega_t \text{ について} \\ B_t = \Omega_t \end{array} \right\} \right)$$

↑
cylindric set
と言う

今もし T が可算集合で、 $\mathcal{F}_t$ が可分な距離空間 S_t 上の Borel σ -加法族 ($\mathcal{F}_t = \mathcal{B}(S_t)$) なら

$$\bigotimes_{t \in T} \mathcal{F}_t (= \bigotimes_{t \in T} \mathcal{B}(S_t)) = \mathcal{B} \left(\prod_{t \in T} S_t \right)$$

となることが知られている。ただし、 $\prod_{t \in T} S_t$ には

直積位相 が入っているとする。

Def (直積位相) product topology

$(S_t, \mathcal{U}_t)$: 位相空間 ($t \in T$)

直積集合上 $\prod_{t \in T} S_t$ 上の直積位相とは

$$\prod_{t \in T} \mathcal{U}_t \quad (\mathcal{U}_t \in \mathcal{U}_t, \text{有限個の } t \text{ の } \omega_t \text{ について } \mathcal{U}_t = S_t)$$

を開集合の基とする位相である。

つまり、 $\prod_{t \in T} \mathcal{U}_t$ は各 $\mathcal{U}_t$ の和集合
で全体の開集合を表せる。

(Kallenberg: Foundations of Modern Probability, 2nd ed. の P.3)

直積 σ -代数族は、各 Ω_A の射影

(直積位相は π_A を連続
とせる最弱の位相)

$$\pi_A: \prod_{A \in T} \Omega_A \rightarrow \Omega_A$$
$$(\omega_A)_{A \in T} \mapsto \omega_A$$

を可測にする最小の σ -代数族でもある。

→ 前のページで例に出した無限回コイン投げで定義した $\mathcal{F}$ がこれ。

例: $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$

$$\mathcal{B}(\mathbb{R}^\infty) = \bigotimes_{n=1}^{\infty} \mathcal{B}(\mathbb{R})$$

Rem

一般に $\bigotimes_{A \in T} \mathcal{B}(S_A) = \mathcal{B}\left(\prod_{A \in T} S_A\right)$ とは限らない。 //

Lem

$(\Omega, \mathcal{F}), (\Omega', \mathcal{F}')$: 可測空間

$X: \Omega \rightarrow \Omega'$ (可測な有限変数) に対し、

$X^{-1}(\mathcal{F}')$ は σ -加法族をなす、 //

Proof

$$(i) X^{-1}(\Omega') = \{\omega \in \Omega \mid X(\omega) \in \Omega'\} = \Omega \quad \mathcal{F}'$$

$$\Omega \in X^{-1}(\mathcal{F}')$$

$$(ii) A \in \mathcal{F}' \text{ のとき } X^{-1}(A)^c = \{\omega \in \Omega \mid X(\omega) \notin A\}$$

$$= \{\omega \in \Omega \mid X(\omega) \in A^c\} = X^{-1}(A^c) \in X^{-1}(\mathcal{F}')$$

$$\text{よって } X^{-1}(A)^c \in X^{-1}(\mathcal{F}') \quad (\because A^c \in \mathcal{F}')$$

$$(iii) A_1, A_2, \dots \in X^{-1}(\mathcal{F}') \text{ のとき}$$

$$\exists B_1, B_2, \dots \text{ して } X^{-1}(B_n) = A_n \text{ とできる.}$$

このとき

$$\bigcup_{n=1}^{\infty} A_n = \{\omega \in \Omega \mid \exists n \text{ s.t. } \omega \in A_n\}$$

$$= \{\omega \in \Omega \mid \exists n \text{ s.t. } X(\omega) \in B_n\}$$

$$= \{\omega \in \Omega \mid X(\omega) \in \bigcup_{n=1}^{\infty} B_n\}$$

$$= X^{-1}\left(\bigcup_{n=1}^{\infty} B_n\right)$$

$\in \mathcal{F}' \quad (\because \mathcal{F}' \text{ は } \sigma\text{-加法族})$

$$\in X^{-1}(\mathcal{F}')$$

(i), (ii), (iii) より $X^{-1}(\mathcal{F}')$ は σ -加法族 //

//

Lem

$\mathcal{A}'$ を Ω' の σ -部分集合族とす。 (σ -部分族とは限らない)
 $X^{-1}(\sigma(\mathcal{A}')) = \sigma(X^{-1}(\mathcal{A}'))$

Proof

まず、先の Lem より $X^{-1}(\sigma(\mathcal{A}'))$ は σ -部分族である。 かつ

$X^{-1}(\sigma(\mathcal{A}')) \supset X^{-1}(\mathcal{A}')$ であるため、 $X^{-1}(\sigma(\mathcal{A}')) \supset \sigma(X^{-1}(\mathcal{A}'))$ である。
($\sigma(\cdot)$ は最小の σ -部分族)

一方、 $\mathcal{F}' := \{C' \subset \Omega' \mid X^{-1}(C') \in \sigma(X^{-1}(\mathcal{A}'))\}$

とかく。 $\mathcal{F}'$ は σ -部分族である。 なぜなら

- (i) $\Omega' \in \mathcal{F}'$ ($\because X^{-1}(\Omega') = \Omega \in \sigma(X^{-1}(\mathcal{A}'))$)
- (ii) $A' \in \mathcal{F}' \Rightarrow X^{-1}(A') \in \sigma(X^{-1}(\mathcal{A}'))$
 $\Rightarrow \underbrace{(X^{-1}(A'))^c}_{X^{-1}(A'^c)} \in \sigma(X^{-1}(\mathcal{A}')) \Rightarrow A'^c \in \mathcal{F}'$
- (iii) $A_n \in \mathcal{F}' \ (\forall n) \Rightarrow X^{-1}(A_n) \in \sigma(X^{-1}(\mathcal{A}')) \ (\forall n)$
 $\Rightarrow X^{-1}(\bigcup_{n=1}^{\infty} A_n) \in \sigma(X^{-1}(\mathcal{A}'))$
 $\Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}'$

また、 $\mathcal{A}' \subset \mathcal{F}'$ をすぐゆかちのて、 $\sigma(\mathcal{A}') \subset \mathcal{F}'$ である。 ($\mathcal{F}'$ は σ -部分族のため)

さらに、 $\mathcal{F}'$ の定義より $X^{-1}(\mathcal{F}') \subset \sigma(X^{-1}(\mathcal{A}'))$ である。

これより、 $X^{-1}(\sigma(\mathcal{A}')) \subset X^{-1}(\mathcal{F}') \subset \sigma(X^{-1}(\mathcal{A}'))$ である。

以上より、 $X^{-1}(\sigma(\mathcal{A}')) = \sigma(X^{-1}(\mathcal{A}'))$

Thm

ある集合族 $\mathcal{A}'$ を用い、 $\mathcal{F}' = \sigma(\mathcal{A}')$ と書けるとき、

$$X: \Omega \rightarrow \Omega' \text{ は可測} \iff X^{-1}(\mathcal{A}') \subset \mathcal{F}$$

Proof

$$\sigma(X^{-1}(\mathcal{A}')) = X^{-1}(\sigma(\mathcal{A}')) = X^{-1}(\mathcal{F}') \quad \text{より従う。} //$$

Cor

$$X: \Omega \rightarrow \mathbb{R} \text{ は r.v.} \iff X^{-1}((-\infty, x]) \in \mathcal{F} \quad (\forall x \in \mathbb{R}) //$$

$$\text{Proof} \quad \sigma(\{(-\infty, x] \mid x \in \mathbb{R}\}) = \mathcal{B}(\mathbb{R}) \text{ から先の Thm より従う。} //$$

* こちらを r.v. の定義とすることも多い。

Ex.

$$(\Omega, \mathcal{F}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \quad \text{のとき、}$$

$X: \Omega \rightarrow \mathbb{R}$ が連続なら X は r.v. (可測) である。

($\because$)

X の連続性より $X^{-1}((x, \infty))$ は開集合 ($\forall x \in \mathbb{R}$)

よって $X^{-1}((x, \infty)) \in \mathcal{B}(\mathbb{R}^d)$ である。

すなわち $X^{-1}((x, \infty))^c = X^{-1}((-\infty, x]) \in \mathcal{B}(\mathbb{R}^d)$ であるので、

X は r.v. である。 //

Def (多次元確率変数)

$X: \Omega \rightarrow \mathbb{R}^d$ が $\forall C \in \mathcal{B}(\mathbb{R}^d)$ に対し、

$X^{-1}(C) \in \mathcal{F}$ であるとき、(つまり、 $X: \Omega \rightarrow \mathbb{R}^d$ が $\mathcal{F}/\mathcal{B}(\mathbb{R}^d)$ -可測なとき)

X を多次元確率変数 と言う、

(もしくは、確率ベクトル)

— あるいは、 $X_1, \dots, X_d: \Omega \rightarrow \mathbb{R}$ が $\mathbb{R}$ 値確率変数のとき、

$$X_1^{-1}((-\infty, x_1]) \cap X_2^{-1}((-\infty, x_2]) \cap \dots \cap X_d^{-1}((-\infty, x_d]) \in \mathcal{F}$$

であるとき、

$$X(\omega) = (X_1(\omega), X_2(\omega), \dots, X_d(\omega))$$

と表す。

$$X^{-1}((-\infty, x_1] \times (-\infty, x_2] \times \dots \times (-\infty, x_d]) = \bigcap_{i=1}^d X_i^{-1}((-\infty, x_i]) \in \mathcal{F}$$

である。今、 $\mathcal{B}(\mathbb{R}^d) = \sigma(\{\prod_{i=1}^d (-\infty, x_i]\})$ より、

これはまた、多次元確率変数である。

- 期待値

確率変数 X の期待値 $E[X] = \int X(\omega) dP(\omega)$ を定義する。

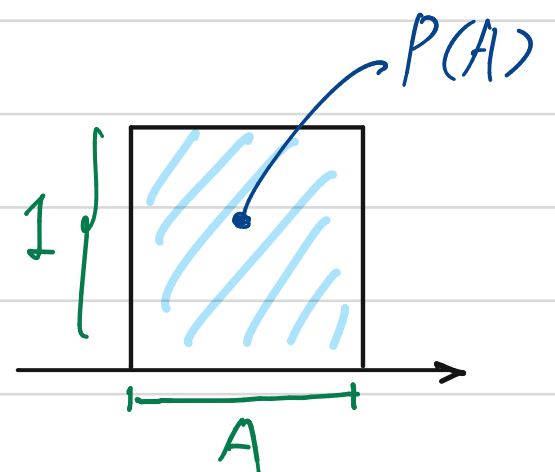
(1) $A \in \mathcal{F}$ に対し、 A の「定義関数」:

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

であるとき、

$$E[X] := P(A)$$

とする。



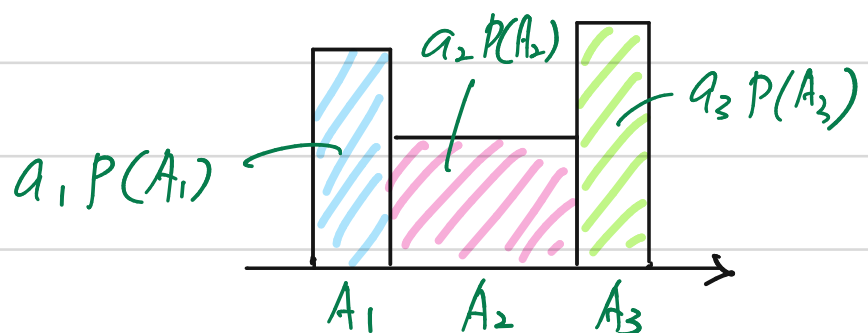
(2) X が非負の「単関数」:

$$X(\omega) = \sum_{i=1}^k a_i \mathbb{1}_{A_i}(\omega)$$

ただし、 $a_i \geq 0$, $A_i \in \mathcal{F}$, k : 有限. のとき、

$$E[X] := \sum_{i=1}^k a_i P(A_i)$$

とする。



(3) X が非負の確率変数:

実は単関数の単調非減少列 $X_1, \dots, X_n$ が存在して、

$$X(\omega) = \lim_{n \rightarrow \infty} X_n(\omega) \quad (\forall \omega \in \Omega)$$

とできる。

これを用いて、

$$E[X] := \lim_{n \rightarrow \infty} E[X_n]$$

とする。

たとえば、

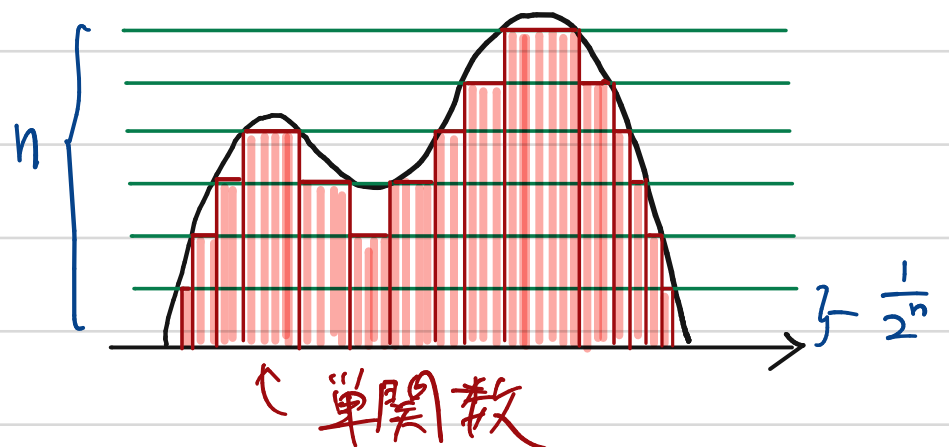
← X_n は単関数になっている。

$$X_n(\omega) = \begin{cases} \frac{i}{2^n} & (\frac{i}{2^n} \leq X(\omega) < \frac{i+1}{2^n}, i = 0, 1, 2, \dots, n2^n - 1) \\ n & (X(\omega) \geq n) \end{cases}$$

とすれば良い。

$E[X] < \infty$ のとき、

X を可積分であると言う



(4) 一般の r.v. X に対しては、

$$X = X^+ - X^- \quad (X^+ = \max(X, 0), X^- = \max(-X, 0))$$

と分解して、

X^+ と X^- がともに可積分なら ($|X|$ が可積分)

$$E[X] = E[X^+] - E[X^-]$$

とする。 $|X|$ が可積分のとき、 X を可積分であると言う。

(3) における $E[X]$ は単関数列の取り方に依存せず well-defined である。

$Y_1, Y_2, \dots$ を別の単関数列で $X(\omega) = \lim_{n \rightarrow \infty} Y_n(\omega)$ であるとする。

このとき、

$$\lim_{n \rightarrow \infty} E[X_n] = \lim_{n \rightarrow \infty} E[Y_n]$$

であることを示す。

(3) における $E[X]$ は単関数列の取り方に依存せず well-defined である。

$Y_1, Y_2, \dots$ を別の単関数列で $X(\omega) = \lim_{n \rightarrow \infty} Y_n(\omega)$ であるとする。

このとき、

$$\lim E[X_n] = \lim E[Y_n]$$

である。

これを示すには、 $\forall m$ で $E[Y_m] \leq \lim_{n \rightarrow \infty} E[X_n]$ を示せば良い。
(あとは対称性より、 $E[X_m] \leq \lim_{n \rightarrow \infty} E[Y_n]$ を示せば、等号が示せる。)

$\varepsilon > 0$ を任意に固定する。(m も固定)

$$A_n = \{ \omega \in \Omega \mid X_n(\omega) \geq Y_m(\omega) - \varepsilon \}$$

なお、すると、 A_n は単調非減少かつ $\bigcup_{n=1}^{\infty} A_n = \Omega$ であるから、

確率の連続性 より $\lim_{n \rightarrow \infty} P(A_n) = 1$ である。

↑ これは P の σ -加法性を使う必要がある！

特に、 $\exists N$ で、 $\forall n \geq N$ において $P(A_n) \geq 1 - \varepsilon$ である。

すると、 $\forall n \geq N$ において、

← $(x)_+ = \max(x, 0)$ である。

$$E[X_n] \geq E[X_n \mathbb{1}_{A_n}] \geq E[(Y_m - \varepsilon)_+ \mathbb{1}_{A_n}]$$

$$\geq E[Y_m \mathbb{1}_{A_n}] - \varepsilon \quad (\because 0 \leq \mathbb{1}_{A_n} \leq 1)$$

$$= E[Y_m - Y_m \mathbb{1}_{A_n^c}] - \varepsilon$$

$$\geq E[Y_m] - P(A_n^c) \cdot \max_{\omega} |Y_m(\omega)| - \varepsilon$$

Y_m は単関数なので有限

$$\geq E[Y_m] - \varepsilon \max_{\omega} |Y_m(\omega)| - \varepsilon.$$

ここで、 $\varepsilon > 0$ は任意なので、示した。

$O(\varepsilon)$



⑤ Thm (期待値の性質)

(1) X, Y : 可積分

$$E[aX + bY] = aE[X] + bE[Y] \quad (\text{線形性})$$

($\forall a, b \in \mathbb{R}$)

(2) $X \geq 0$ (a.s.) $\Rightarrow E[X] \geq 0$ (単調性)
almost surely

(3) (単調収束定理)

$\leftarrow X_n$ は r.v. (単関数に限らない)

$0 \leq X_1 \leq X_2 \leq \dots$ かつ $X_n(\omega) \rightarrow X(\omega)$ (a.s.) なら、

$$\lim_{n \rightarrow \infty} E[X_n] = E[X] \quad (\text{右辺} = \infty \text{ としても成り立つ})$$

(4) (Fatouの補題)

$\leftarrow$ 全ての X_n は r.v.

$X_n \geq 0$ ならば

$$E[\liminf_n X_n] \leq \liminf_n E[X_n]$$

(5) (優収束定理)

$X_n \rightarrow X$ (a.s.) なる r.v. の列 $(X_n)_n$ かつ

可積分な非負 r.v. Y に対し, $|X_n| \leq Y$ ($\forall n$) を満たすなら、

$$\lim_{n \rightarrow \infty} E[X_n] = E[X]$$

(Proof)

(3) 期待値の構成から、 $E[X_n] \leq E[X]$ は明らか。

よって、 $\lim_{n \rightarrow \infty} E[X_n] \leq E[X]$ 。

一方 $Y_m \nearrow X$ である単関数の列 $(Y_m)_m$ に対し、 $\forall m$ で

$$\lim_{n \rightarrow \infty} E[X_n] \geq E[Y_m] \quad \text{が言えるから}$$

$$\lim_{n \rightarrow \infty} E[X_n] \geq E[X] \quad \text{が成り立つ。}$$

任意の $\varepsilon > 0$ に対し、 $A_n = \{\omega \in \Omega \mid X_n(\omega) \geq Y_m(\omega) - \varepsilon\}$ とすると、

$\lim_{n \rightarrow \infty} P(A_n) = P(\bigcup_{n=1}^{\infty} A_n) = 1$ となるので、先と同様に $\exists N$ で、

$\forall n \geq N$ で $P(A_n) \geq 1 - \varepsilon$ である。

このとき、 $\forall n \geq N$ として

$$\begin{aligned} E[X_n] &\geq E[X_n \mathbb{1}_{A_n}] \geq E[(Y_m - \varepsilon) \mathbb{1}_{A_n}] \\ &\geq E[Y_m (1 - \mathbb{1}_{A_n^c})] - \varepsilon \\ &= E[Y_m] - \underbrace{\sup_{\omega} |Y_m(\omega)| p(A_n^c)}_{\leq \varepsilon} - \varepsilon \\ &\geq E[Y_m] - O(\varepsilon). \end{aligned}$$

よって、 $\varepsilon \searrow 0$ とすると $\lim_{n \rightarrow \infty} E[X_n] \geq E[Y_m]$ を得る。

← 可測性があることは $\lim_{n \rightarrow \infty} X_n = Y$ から得る。

(4) $Y_n = \inf_{m \geq n} X_m$ とおけば、 $Y_1 \leq Y_2 \leq \dots$

定義より、 $E[Y_n] \leq E[X_m]$ ($\forall m \geq n$)

つまり、 $E[Y_n] \leq \inf_{m \geq n} E[X_m]$.

あとは、単調収束定理より、

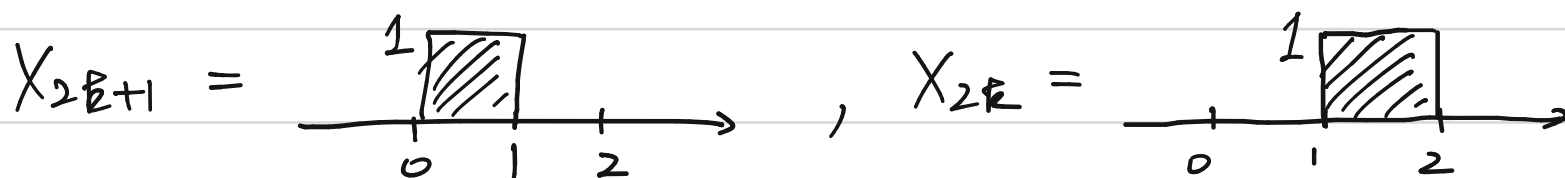
$$\begin{aligned} E[\liminf X_n] &= E[\lim Y_n] = \lim E[Y_n] \\ &\leq \liminf_{n \rightarrow \infty} \inf_{m \geq n} E[X_m] \end{aligned}$$

単調収束定理

(5) $Y - X_n$ および $Y + X_n$ に Fatou の補題を適用すれば示せる。

($|X_n| \leq Y$ より、 $0 \leq Y \pm X_n$ であることに注意)

• Fatou の補題の反例 (= 成り立たない例)



P : $[0, 2]$ 上の一様分布.

$\liminf X_n = 0$ かつ、 $E[\liminf X_n] = 0$ $\neq$

かつ、 $E[X_n] = \frac{1}{2}$ ($\forall n$) かつ、 $\liminf E[X_n] = \frac{1}{2}$.

• 優収束定理の反例

$P(\{k\}) = \frac{C}{k^2}$ ($k=1, 2, \dots$), $X_n(\omega) = \begin{cases} 0 & (\omega \neq n) \\ n^2 & (\omega = n) \end{cases} \Rightarrow E[X_n] = C \neq 0$

かつ、 $\lim_{n \rightarrow \infty} X_n(\omega) = 0$ ($\forall \omega$) $\Rightarrow E[\lim X_n] = 0$

← 可積分な Y が存在しない。

$\neq$

Cor (微分と積分の交換)

$$X: \mathbb{R} \times \Omega \rightarrow \mathbb{R}$$

$$(t, \omega) \mapsto X(t, \omega)$$

かつ $X(t, \cdot)$ は可測.

X は各 $\omega \in \Omega$ において t について微分可能.

その導関数 $\frac{\partial X(t, \omega)}{\partial t}$ は区間 (a, b) で可積分可測

$$\text{関数 } \gamma \in \mathcal{L}^1 \text{ を用いて, } \left| \frac{\partial X(t, \omega)}{\partial t} \right| \leq \gamma(\omega) \quad (\forall t \in (a, b), \forall \omega \in \Omega)$$

で表すとす. すると, $\forall t \in (a, b)$ で

$$\frac{d}{dt} \int X(t, \omega) dP(\omega) = \int \frac{\partial}{\partial t} X(t, \omega) dP(\omega)$$

が成り立つ.

(Proof) 微分の定義より $\frac{\partial X(t, \omega)}{\partial t} = \lim_{h \rightarrow 0} \frac{X(t+h, \omega) - X(t, \omega)}{h}$ である.

十分に小さな h に対し、中間値の定理より

$$\frac{X(t+h, \omega) - X(t, \omega)}{h} = \frac{\partial X(\tilde{t}, \omega)}{\partial \tilde{t}} \Big|_{\tilde{t} = t + \theta h}$$

$$(0 \leq \theta(\omega) \leq 1)$$

が成り立つ. 仮定より, $\forall h$: 十分に小さくして、一様に

$$\left| \frac{X(t+h, \omega) - X(t, \omega)}{h} \right| = \left| \frac{\partial X(\tilde{t}, \omega)}{\partial \tilde{t}} \right| \leq \gamma(\omega)$$

である. したがって、優収束定理より

$$\begin{aligned} \frac{d}{dt} \int X(t, \omega) dP(\omega) &= \lim_{h \rightarrow 0} \frac{\int X(t+h, \omega) dP(\omega) - \int X(t, \omega) dP(\omega)}{h} \\ &= \lim_{h \rightarrow 0} \int \frac{X(t+h, \omega) - X(t, \omega)}{h} dP(\omega) \end{aligned}$$

優収束 $\rightarrow$

$$\int \underbrace{\lim_{h \rightarrow 0} \frac{X(t+h, \omega) - X(t, \omega)}{h}}_{\leq \frac{\partial X(t, \omega)}{\partial t}} dP(\omega)$$