確率数理要論 (1)

講義内容

1. 確率空間
2. 確率変数と期待値（単調収束定理・Fatouの補題・優収束定理）
3. 独立性と大数の法則
4. 確率変数の弱収束（Portmanteauの定理）
5. 特性関数と中心極限定理
6. 条件付き期待値
7. Poisson過程（Gamma, Levy, Dirichlet過程）
8. Brown運動
9. Gauss過程と機械学習応用（Karhunen-Loeve展開，再生核ヒルベルト空間）
10. 確率積分（伊藤積分）
11. マルチンゲール
12. 伊藤の公式
13. 確率微分方程式
14. 対数Sobolev不等式・機械学習応用（拡散モデル・ランジュバン動力学）

参考書

- 舟木直久「確率論」
- 佐藤坦「測度から確率」
- 伊藤清「確率論」
- Durrett「Probability, Theory and Exchange」
- Resnick「A Probability Path」
- Dudley「Real Analysis and Probability」

1. 確率空間

Ω : 標本空間　(空間)

- 想定できる結果を全て集めたもの
- 要素 $\omega \in \Omega$ で実際に何が起きたかを表現

$\omega \in \Omega$: 標本点　(標本, 根元事象)

$A \subset \Omega$: 事象

Ex.
- 一定期間に届くメールの件数 : $\Omega = \{0, 1, 2, 3, \dots\}$
- 無限回のコイン投げ : $\Omega = \{\omega = (\omega_1, \omega_2, \omega_3, \dots) \mid \omega_i \in \{0, 1\}\}$

これから事象の確率を定めてゆく.

Def (σ-加法族, 完全加法族)

Ω の部分集合を要素とする集合族 $\mathcal{F}$ が次をみたすとき.
$\mathcal{F}$ を σ-加法族 (完全加法族) と呼ぶ.

(1) $\phi \in \mathcal{F}$

(2) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

(3) [σ-加法性]

$A_i \in \mathcal{F} \ (i=1,2,\dots) \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

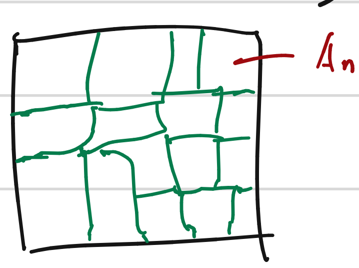
可算個
(非可算和までは要求しない)

//

任意の集合 $A \subset \Omega$ 上に確率を定めようとすると,
矛盾が生じることがある (ルベーグ非可測集合 : $P\left(\bigcup_{n=1}^{\infty} A_n\right) = 1$, $P(A_n) = 0$ で $P\left(\bigcup_{n=1}^{\infty} A_n\right) \neq \sum_{n=1}^{\infty} P(A_n)$ となる)

↓

矛盾が生じない範囲に
限定して確率を定める.
確率が矛盾なく定まっている事象の集まり
を用意する. それが σ-加法族.



Ωと$\mathcal{F}$の組$(\Omega, \mathcal{F})$を可測空間と言う。
$\mathcal{F}$の元を可測集合と言う。

σ-加法族の中でも特に重要なのが、次のBorel集合族。

Def (Borel集合族)
S: 位相空間
Sの開集合を全て含む最小のσ-加法族を
Borel集合族と言う
$\mathcal{B}(S)$と書く。
Borel集合族の元をBorel集合と言う。

後述。

Def (確率測度)
$(\Omega, \mathcal{F})$: 可測空間
集合関数 $P: \mathcal{F} \to \mathbb{R}$ が以下をみたすとき、
Pを確率測度と言う。

(1) $0 \leq P(A) \leq 1 \quad (\forall A \in \mathcal{F})$
(2) $P(\Omega) = 1$
(3) $A_1, A_2, A_3 \cdots \in \mathcal{F}$ が $A_i \cap A_j = \phi \ (i \neq j)$
をみたすなら、
$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \qquad [\sigma\text{-加法性}]$$

(有限と無限を矛盾なくつなぐのに必要。e.g., 積分、極限)

$P(\bigcup_{i=1}^{\infty} A_i)$が定義されるためには $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ が必要 ($\mathcal{F}$のσ-加法性)

Pが$(\Omega, \mathcal{F})$上の確率測度のとき、
$(\Omega, \mathcal{F}, P)$を確率空間と言う

Ex.

(1) サイコロ

$$\Omega = \{1, 2, \cdots, 6\}$$
$$P(\{\omega\}) = \frac{1}{6}$$

$$P(\{\omega \leqq 3 \mid \omega \in \Omega\}) = P(\{1,2,3\})$$
($P(\omega \leqq 3)$ とも書く) $= P(\{1\}) + P(\{2\}) + P(\{3\})$
$$= \frac{1}{2}$$

(2) $\Omega = \{0, 1, 2, 3, \cdots\}$

$A \subset \Omega$ に対し、
$$P(A) = \sum_{\omega \in A} \frac{\lambda^\omega}{\omega !} e^{-\lambda}$$
は確率測度 (Poisson分布)

※ 一般に可算集合 Ω のとき、$0 \leqq h(\omega) \leqq 1$ かつ $\sum_{\omega \in \Omega} h(\omega) = 1$ である $h : \Omega \to \mathbb{R}$ に対し、
$$P(A) = \sum_{\omega \in A} h(\omega)$$ は確率測度

Cor

$(\Omega, \mathcal{F}, P)$: 確率空間。

以下が成り立つ:

(1) $P(\phi) = 0$

(2) (有限加法性)
$$A_i \in \mathcal{F} \quad (i = 1, 2, \cdots, n)$$
$$A_i \cap A_j = \phi \quad (i \neq j)$$
$$\Longrightarrow P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

(3) $P(A^c) = 1 - P(A)$

(4) $A \subset B \Longrightarrow P(A) \leqq P(B)$ (単調性)

(5) $A_1 \subset A_2 \subset \cdots$

$$\Longrightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} P(A_n)$$

(6) $A_1 \supset A_2 \supset \cdots$

$$\Longrightarrow P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} P(A_n)$$

→ 確率の連続性

(今後色々な定理の証明で用いる.)

(7) $A_1, A_2, \cdots \in \mathcal{F}$, $A_i \cap A_j = \phi$ $(i \neq j)$ とは限らない

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i) \quad \text{(劣加法性)}$$

//

(証明は省略. 自分でチェック)

◎"最小のσ-加法族"について.

Lem

$(\mathcal{F}_u)_{u\in\mathcal{U}}$: σ-加法族の族 (非可算濃度も許す)

$$\Longrightarrow \bigcap_{u\in\mathcal{U}} \mathcal{F}_u = \mathcal{F} \quad \text{もまたσ-加法族}$$

//

Proof

(1) $\forall u \in \mathcal{U}$ について $\Omega \in \mathcal{F}_u$ なので、$\Omega \in \mathcal{F}$ でもある.

(2) $A \in \mathcal{F} \Longrightarrow \forall u \in \mathcal{U}$ で $A \in \mathcal{F}_u$

$\Longrightarrow \forall u \in \mathcal{U}$ で $A^c \in \mathcal{F}_u$

$\Longrightarrow A^c \in \mathcal{F}$

(3) $A_i \in \mathcal{F}$ $(\forall i = 1, 2, \cdots) \Longrightarrow \forall u \in \mathcal{U}$ で $A_i \in \mathcal{F}_u$ $(\forall i)$

$\Longrightarrow \forall u \in \mathcal{U}$ で $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}_u$

($\because$ σ-加法性)

$\Longrightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

//

ある Ω の部分集合族 $\mathcal{A}$ に対し，
$\mathcal{A}$ を含む__最小のσ-加法族__が定義できる：

$$\sigma(\mathcal{A}) := \bigcap_{\mathcal{F}:\,\mathcal{A}\text{を含む}\sigma\text{-加法族}} \mathcal{F}$$

例：Borel 集合族は，全 S の開集合を集めた $\mathcal{A}$ に対し，
$\mathcal{B}(S) = \sigma(\mathcal{A})$ で与えられる．

これ以降，$(\Omega, \mathcal{F}, P)$ は与えられているものとする．
（構成のしかたは考えない）
しかし，以下のようにして構成することができる．

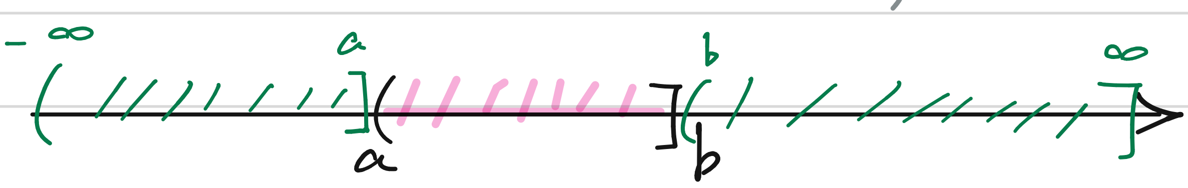
__Def__（集合半代数）

集合族 S が集合半代数 $\Longleftrightarrow$

(1) $\phi, \Omega \in S$
(2) $A, B \in S \Rightarrow A \cap B \in S$
（π-システム）
(3) $A \in S$ なら，A^c は互いに排反な有限個の集合
$B_1, \dots, B_n \in S$ に分割される：

$$A^c = \bigcup_{i=1}^{n} B_i \quad (B_i \in S,\ B_i \cap B_j = \phi\ (i \neq j))$$

例：$\Omega = \mathbb{R}$，$S = \{ (a, b] \mid -\infty \leq a \leq b \leq \infty \}$ は集合半代数
（半開区間）
（正確には $b=\infty$ のときは (a, b)）



Thm（Hopfの拡張定理（の一般化））

集合半代数 $\mathcal{S}$ 上の σ-加法的な集合関数 $P : \mathcal{S} \to [0,1]$ で $P(\Omega) = 1$ をみたすものは $\sigma(\mathcal{S})$ 上の確率測度へ一意に拡張できる。 //

※ P が $\mathcal{S}$ 上 σ-加法的：

$A_i \cap A_j = \phi \ (i \neq j)$ なる $A_n \in \mathcal{S} \ (n=1,2,\cdots)$ が $\displaystyle\bigcup_{n=1}^{\infty} A_n \in \mathcal{S}$ なら、

$$P\Big(\underbrace{\bigcup_{n=1}^{\infty} A_n}_{\in \mathcal{S}}\Big) = \sum_{n=1}^{\infty} P(\underbrace{A_n}_{\in \mathcal{S}}).$$

例：$\Omega = \mathbb{R}$，後述の分布関数を用いて

$$P((a,b]) = F(b) - F(a)$$

とすれば、$\mathcal{S} = \{(a,b] \mid -\infty \leq a \leq b \leq \infty\}$ 上の σ-加法的な集合関数が得られる。

また、

$$\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{S})$$ チェックせよ.

も成り立つ。

よって、P は $\mathcal{B}(\mathbb{R})$ 上の確率測度に一意に拡張される。 //

Def（有限加法族）

$\mathcal{A}$ が有限加法族 $\Longleftrightarrow$

$$\begin{cases} (1) & \phi \in \mathcal{A} \\ (2) & A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A} \\ (3) & A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A} \end{cases}$$ //

有限加法族が集合半代数であることはすぐわかる。

Cor

$P : \mathcal{A} \to [0,1]$ が $P(\Omega) = 1$ かつ $\mathcal{A}$ 上 σ-加法的なら P は $\sigma(\mathcal{A})$ まで一意に拡張される。 //

○ 実数軸上の確率測度

$\Omega = \mathbb{R}$

$\mathcal{F} = \mathcal{B}(\mathbb{R})$ 　　（$\mathbb{R}$上のBorel集合族）

$P : (\Omega, \mathcal{F})$ 上の確率測度

（以後、$\mathbb{R}$上のσ-加法族はBorel集合族を考える）

__Def__（累積分布関数）

$$F(x) = P((-\infty, x]) = P(\{\omega \in \mathbb{R} \mid -\infty < \omega \leq x\}) \quad (x \in \mathbb{R})$$

を（累積）分布関数と呼ぶ。 //

__Lem__（分布関数の性質）

(1) Fは単調非減少 （$F(x) \leqq F(y)$ for $x \leqq y$）

(2) Fは右連続：

$$\lim_{x \downarrow y} F(x) = F(y)$$

（逆は成り立たない。$\lim_{x \nearrow y} F(x) \neq F(y)$）

(3) $\lim_{x \to -\infty} F(x) = 0$, 　$\lim_{x \to +\infty} F(x) = 1$ //

（証明は確率の連続性からすぐに示せる。）

逆に(1)〜(3)の性質をもつ$\mathbb{R}$上の関数Fが与えられれば対応する$\mathcal{B}(\mathbb{R})$上の確率測度が一意に定まる。

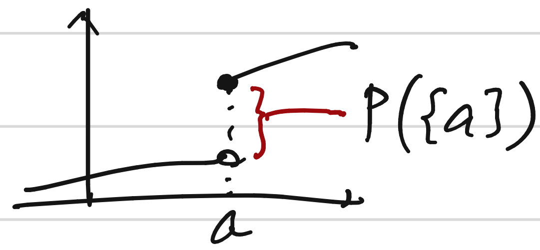
$$P \longleftrightarrow F \quad \text{（一対一対応）}$$

∵ Hopfの拡張定理

Lem

$$\lim_{x \to a} F(x) < F(a)$$

$P(\{a\}) \neq 0$ と同値 (チェックせよ)



が成り立つとき、「点 a でジャンプがある」と言う。
ジャンプがある点は高々可算個である。 //

(Proof)

$h(a) = P(\{a\})$ とする。
$A = \{a \in \mathbb{R} \mid h(a) > 0\}$ とする。
$A_n = \{a \in \mathbb{R} \mid h(a) \geq \frac{1}{n}\}$ とおくと、
$A = \bigcup_{n=1}^{\infty} A_n$ である。

ここで $|A_n| \leq n$ はすぐにわかるので、A は高々可算。
$\left(\because 1 \geq \underbrace{\sum_{a \in A_n} h(a)}_{P(A_n)} \geq |A_n| \cdot \frac{1}{n} \to |A_n| \leq n\right)$ //

Def

${}^\forall x \in \mathbb{R}$ に対し、

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

なる $f : \mathbb{R} \to \mathbb{R}$ が存在するとき、
f を F の確率密度関数と言う。

このとき、F は絶対連続であると言う。 //

Ex. 正規分布、一様分布、ガンマ分布、指数分布

(ポアソン分布、二項分布は絶対連続ではない)

Radon-Nikodym の定理

※ ルベーグ測度 μ に対し、$\mu(E) = 0$ をみたす任意の $E \in \mathcal{B}(\mathbb{R})$ について $P(E) = 0$ となるなら、P (と対応する F) は絶対連続。

◦ 多次元空間上の確率測度

Def

$(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ 上の確率測度 P に対し

$$F(x_1, x_2, \cdots, x_n) := P\left((-\infty, x_1] \times (-\infty, x_2] \times \cdots \times (-\infty, x_n]\right)$$
$$= P\left(\{\omega \in \mathbb{R}^n \mid \omega_i \leqq x_i \ (i=1, \cdots, n)\}\right)$$

を（同時確率）分布関数と言う

//

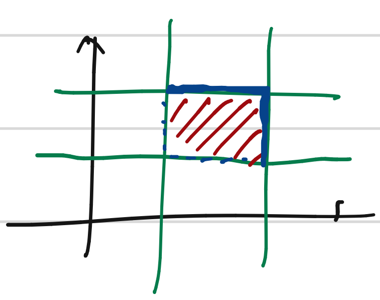
• $P\left((a_1, b_1] \times (-\infty, b_2] \times \cdots \times (-\infty, b_n]\right)$

$= F(b_1, b_2, \cdots, b_n) - F(a_1, b_2, \cdots, b_n)$

• $P\left((a_1, b_1] \times (a_2, b_2] \times \cdots \times (a_n, b_n]\right)$

$$= \sum_{i_1=0}^{1} \sum_{i_2=0}^{1} \cdots \sum_{i_n=0}^{1} (-1)^{i_1 + \cdots + i_n} F(x_{1,i_1}, \cdots, x_{n,i_n})$$

ただし $x_{i,0} = b_i$, $x_{i,1} = a_i$ とする．



Lem（分布関数の性質）

(1) F は単調非減少 $\left(x_i \leqq y_i \ (\forall i=1, \dots, n) \Rightarrow F(x) \leqq F(y)\right)$

(2) F は右上連続

すなわち $x_i \searrow y_i \ (i=1, \cdots, n)$ なら

$F(x) \longrightarrow F(y)$

(3) $\displaystyle\lim_{\substack{x_j \to -\infty \\ (x_i : \text{fix } (i \neq j))}} F(x_1, \cdots, x_j, \cdots, x_n) = 0$, $\displaystyle\lim_{x \to (+\infty, \cdots, +\infty)} F(x) = 1$

//

Def

$\forall (x_1, \ldots, x_n) \in \mathbb{R}^n$ に対して.

$$F(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(t_1, \ldots, t_n)\, dt_1 \cdots dt_n$$

が成り立つ $f : \mathbb{R}^n \to \mathbb{R}$ が存在する時.
f を F の同時確率密度関数と言う //