確率数理工学3

これまでの内容を逆にたどってみる.

○ 分布関数 $F(x)$:
$$\begin{cases} (1)\ x<y \Rightarrow F(x) \le F(y) \\ (2)\ \lim_{x \downarrow a+0} F(x) = F(a) \quad (\text{右連続}) \\ (3)\ \lim_{x\to -\infty} F(x)=0,\quad \lim_{x\to+\infty} F(x)=1 \end{cases}$$

が与えられたとする.

(ある確率変数 X に対し, $P(X \le x) = F(x)$ を想定)

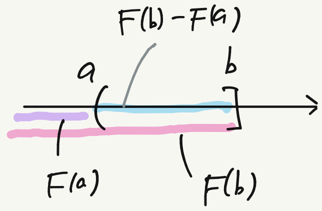
↳ そのためには $\{\omega \mid X(\omega) \le x\}$ は可測でなくてはいけない ⇝ 確率変数の定義

○ F に対応する $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ 上の確率測度 P が一意に存在する.

つまり, $P((-\infty, a]) = F(a)$ $(\forall a \in \mathbb{R})$ をみたす P が存在.

アウトライン

○ まず $P((a,b]) = F(b) - F(a)$ $(a<b)$ とする.



① $\mathcal{S} = \{(a,b] \mid -\infty \le a \le b \le \infty\}$ は"集合代数"になっている.

次ページ: 前回補足資料

$\boxed{\text{Q: どの集合族まで } P \text{ を矛盾なく拡張できる?}}$

② $(a,b] = \bigcup_{i=1}^{\infty} (a_i, b_i]$ と互いに重なりのない区間 $(a_i, b_i]$ $(i=1,2,\ldots)$ に分割した時,

P は $P((a,b]) = \sum_{i=1}^{\infty} P((a_i,b_i])$ をみたすことが示せる. (非自明)

$\left(F(b) - F(a) = \sum_{i=1}^{\infty} (F(b_i) - F(a_i))\right)$

$[F(b), F(a)]$ のコンパクト性を使う.

⇒ P は $\mathcal{S}$ 上で σ-加法的 と言う.

$$\begin{cases} \cdot\ 0 \le P(A) \le 1 \quad (\forall A \in \mathcal{S}) \\ \cdot\ P(\mathbb{R}) = 1 \\ \cdot\ \mathcal{S} \text{ 上で } \sigma\text{-加法的} \end{cases}$$

③ Hopfの拡張定理(の一般化):

上の3つの性質をみたす $\mathcal{S}$ 上の集合関数 $P: \mathcal{S} \to [0,1]$ は $\mathcal{S}$ を含む最小の σ-加法族 $\sigma(\mathcal{S})$ 上の確率測度へ一意に拡張される.

$\boxed{\text{A: } \sigma(\mathcal{S}) \ (\mathcal{S} \text{ を含む最小の } \sigma\text{-加法族}) \text{ まで}}$

④ $\sigma(\mathcal{S}) = \mathfrak{B}(\mathbb{R})$ であることより, 示したいことが言えた.

この講義では、確率空間は所与のものとし、その"構成法"には深く立ち入らない。しかし、構成法の過程自体は知っておいた方が理解の助けになる。以下、この事について説明する。

• $\mathcal{B}(\mathbb{R})$ 上の測度の構成 (Hopfの拡張定理)

Def (集合半代数) (半環は $\Omega \in S$ を要請しない)

S が集合半代数 $\iff$
- (1) $\phi, \Omega \in S$
- (2) S は π-システム $(A, B \in S \Rightarrow A \cap B \in S)$
- (3) $A \in S$ なら、A^c は有限個の互いに排反な集合 $B_1, \dots, B_n \in S$ 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. ☆

$P: S \to [0,1]$ は S 上の σ-加法的集合関数であるとする.

つまり、
- (1) $P(\Omega) = 1$
- (2) $A_n \in S$ が互いに排反かつ $\bigcup_{n=1}^{\infty} A_n \in S$ なら
  $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$

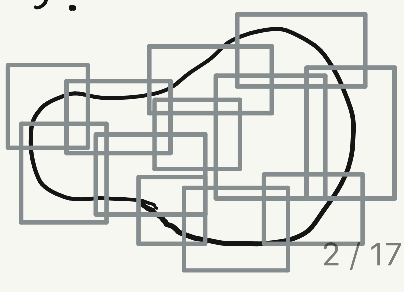
とする。

☆の場合、$P((a,b]) = F(b) - F(a)$ とし、$(a_n, b_n]$ $(n=1,2,\dots)$ が互いに排反で $\bigcup_{n=1}^{\infty} (a_n, b_n] = (a, b]$ と書けるなら、$\sum_{n=1}^{\infty} P((a_n, b_n]) = P((a,b]) = F(b) - F(a)$ であることは示せる。つまり、S 上 σ-加法的である。

任意の部分集合 $A \subset \Omega$ ($A \in S$ とは限らない) に対して、P より定まる外測度 P^* を以下のように定める:　←可算無限和

$$P^*(A) := \inf\left\{ \sum_{i=1}^{\infty} P(B_i) \;\middle|\; B_i \in S,\ A \subset \bigcup_{i=1}^{\infty} B_i \right\}.$$

($A \subset \bigcup_{n=1}^{\infty} B_i$ を有限和 $A \subset \bigcup_{n=1}^{n} B_i$ とすると、Jordan外測度と呼ばれるものになるが、これでは不十分である。
→ σ-加法性が出ない。)



ここで、

$$\mathcal{D} := \{ D \subset \Omega \mid P^*(D) + P^*(D^c) = 1 \}$$

とおくと、以下が成り立つ。

Thm

1. $\mathcal{D}$ は σ-加法族
2. P^* の $\mathcal{D}$ への制限は $(\Omega, \mathcal{D})$ 上の確率測度になっている。

(非自明, 前回補足資料に証明はのせてある)

今 $\mathcal{S}$ の定義と、P が $\mathcal{S}$ 上 σ-加法的であることから、

$\mathcal{S} \subset \mathcal{D}$ が示せるので、$\mathcal{D}$ が σ-加法族であることと合わせて、

$\sigma(\mathcal{S}) \subset \mathcal{D}$ もわかる。

$A \in \mathcal{S}$ なら $P^*(A) = P(A)$ で、$\forall A \in \mathcal{S}$ に対して、

$\exists B_1, B_2, \ldots, B_m \in \mathcal{S}$ で、$A^c = B_1 \cup \cdots \cup B_m$ である。

(互いに排反)

すると、$1 = P(\Omega) = P(A \cup B_1 \cup \cdots \cup B_m) = P(A) + \underbrace{\sum_{i=1}^{m} P(B_i)}_{= P^*(A^c)}$

全部 $\mathcal{S}$ の元。

$\forall A \in \mathcal{S}$ で $P^*(A) = P(A)$ も示せるので、

P^* は P の $\sigma(\mathcal{S})$ への"拡張"を与えている。

☆の場合、$\sigma(\mathcal{S}) = \mathcal{B}(\mathbb{R})$ なので、P^* は $\mathcal{B}(\mathbb{R})$ 上の確率測度を与える。

さらに、このような測度は一意に定まることが言える。 $\longrightarrow$ 次ページ

◦ 拡張の一意性

- Dynkinのπ-λ定理

目標：分布関数 F が与えられれば、対応する確率測度 P が一意に定まることを言いたい。つまり、2つの確率測度 P_1, P_2 が

$$P_1((-\infty, a]) = P_2((-\infty, a]) = F(a)$$ をみたすなら

$$P_1(A) = P_2(A) \quad (\forall A \in \mathcal{B}(\mathbb{R}))$$ となること

を示したい。（ただし、そのような測度が存在することはすでに言っている。）

Def

- π-システム：集合族 $\mathcal{P}$ が π-システム $\Leftrightarrow$ $A, B \in \mathcal{P}$ なら $A \cap B \in \mathcal{P}$
（有限回の共通部分を取る操作について閉じている）

$B \setminus A := B \cap A^c$

- λ-システム：（Dynkin族）集合族 $\mathcal{L}$ が λ-システム $\Leftrightarrow$
$$\begin{cases} (1)\ \Omega \in \mathcal{L} \\ (2)\ A, B \in \mathcal{L},\ A \subset B \Rightarrow B \setminus A \in \mathcal{L} \\ (3)\ A_1 \subset A_2 \subset A_3 \subset \cdots \in \mathcal{L} \\ \quad \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{L} \end{cases}$$

//

☆ σ-加法族は π-システムでも λ-システムでもある。

☆ 先の集合半代数は π-システムだが、λ-システムではない。

$\sigma(\mathcal{P})$：$\mathcal{P}$ を含む最小の σ-加法族

$\mathcal{L}(\mathcal{P})$：$\mathcal{P}$ を含む最小の λ-システム　とする。

Thm（Dynkinの定理）（証明は前回補足資料を参照）

$\mathcal{P}$ は π-システムで、$\mathcal{L}$ は $\mathcal{P}$ を含む λ-システム（$\mathcal{P} \subset \mathcal{L}$）であるとする。
すると、$\sigma(\mathcal{P}) \subset \mathcal{L}$ である。

特に、$\sigma(\mathcal{P}) = \mathcal{L}(\mathcal{P})$ である。

なぜなら、Dynkinの定理より、$\sigma(\mathcal{P}) \subset \mathcal{L}(\mathcal{P})$ が成り立ち、

かつ任意の σ-加法族は λ-システムでもあるので、

$\sigma(\mathcal{P})$ は $\mathcal{P}$ を含む λ-システムでもある。つまり $\sigma(\mathcal{P}) \supset \mathcal{L}(\mathcal{P})$ が成り立つ。

よって $\sigma(\mathcal{P}) = \mathcal{L}(\mathcal{P})$ である。

Dynkinの定理より次の命題が示せる.

Cor 1

$(\Omega, \mathcal{F})$ 上の確率測度 P_1, P_2 が、ある π-システム $\mathcal{P} \subset \mathcal{F}$ において.

$$\forall A \in \mathcal{P} \text{ で } P_1(A) = P_2(A)$$

をみたすなら.

$$\forall B \in \sigma(\mathcal{P}) \text{ で } P_1(B) = P_2(B)$$

が成り立つ.

(Cor 1の証明)

まず. $\mathcal{L} := \{A \in \mathcal{F} \mid P_1(A) = P_2(A)\}$ とすると. $\mathcal{L}$ は λ-システムになっていることを示す.

(1) $P_1(\Omega) = 1$, $P_2(\Omega) = 1$ より. $\Omega \in \mathcal{L}$.

(2) $A, B \in \mathcal{L}$, $A \subset B$ とすると. $P_1(A) = P_2(A)$ かつ $P_1(B) = P_2(B)$ である.
確率測度の性質(加法性)より.

$$P_1(B \setminus A) = P_1(B) - P_1(A) = P_2(B) - P_2(A) = P_2(B \setminus A).$$

よって、 $B \setminus A \in \mathcal{L}$ を得る.

(3) $A_1 \subset A_2 \subset A_3 \subset \cdots \in \mathcal{L}$ とする. 確率の連続性より

$$P_1\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P_1(A_n) = \lim_{n\to\infty} P_2(A_n) = P_2\left(\bigcup_{n=1}^{\infty} A_n\right)$$

なので. $\bigcup_{n=1}^{\infty} A_n \in \mathcal{L}$ である.

以上より. $\mathcal{L}$ は λ-システムであることが分かった.
仮定より $\mathcal{P} \subset \mathcal{L}$ なので、Dynkinの定理から、$\sigma(\mathcal{P}) \subset \mathcal{L}$ である.
$\mathcal{L}$ 上で P_1 と P_2 は一致するので、$\sigma(\mathcal{P})$ 上でも両者は一致する.

また、この系として、分布関数が共通な2つの確率測度が $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上で一致することが示せる.

Cor 2 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の確率測度 P_1, P_2 が

$$P_1((-\infty, a]) = P_2((-\infty, a]) = F(a) \quad (\forall a \in \mathbb{R})$$

をみたすなら.
$\mathcal{B}(\mathbb{R})$ 上で $P_1 = P_2$ である.

(Cor 2の証明)

$\mathcal{P} = \{(-\infty, a] \mid a \in \mathbb{R}\}$ は π-システムである. 仮定より. $\forall A \in \mathcal{P}$ で $P_1(A) = P_2(A)$ である. よって Cor 1 より $P_1(B) = P_2(B)$ $(\forall B \in \sigma(\mathcal{P}))$ でもある. ここで、$\sigma(\mathcal{P}) = \mathcal{B}(\mathbb{R})$ であることに注意すると、これは $\mathcal{B}(\mathbb{R})$ 上で $P_1 \equiv P_2$ であることに他ならない.

これから Dynkinの定理を示すが、その前に次のことを注意しておく.

Lem

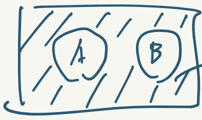
$\mathcal{L}$が λ-システム $\Longleftrightarrow$ $\begin{cases} (1)' \ \Omega \in \mathcal{L} \\ (2)' \ A \in \mathcal{L} \Rightarrow A^c \in \mathcal{L} \\ (3)' \ A_n \cap A_m = \phi \ (\forall n \neq m), \ A_n \in \mathcal{L} \text{ なら}, \\ \quad \bigcup_{n=1}^{\infty} A_n \in \mathcal{L} \text{ である}. \end{cases}$

(証明)

($\Rightarrow$のみ示す)

(1)'は定義から自明. (2)'も $\Omega \in \mathcal{L}$ なので $\Omega \setminus A \in \mathcal{L}$. つまり $A^c \in \mathcal{L}$ が従う.

(3)'を示す. そのため, まずは $A, B \in \mathcal{L}$, $A \cap B = \phi$ に対し, $A \cup B \in \mathcal{L}$であることを示す. (2)より $\Omega \setminus A \in \mathcal{L}$ である. また $B \subset \Omega \setminus A$ なので $(\Omega \setminus A) \setminus B \in \mathcal{L}$ でもある. ここで $(\Omega \setminus A) \setminus B = A^c \cap B^c$ なので $A^c \cap B^c \in \mathcal{L}$. (2)'より, $A^c \cap B^c \in \mathcal{L}$ なら $(A^c \cap B^c)^c = A \cup B \in \mathcal{L}$ である.



$A^c \cap B^c$

あとは $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} \underbrace{\bigcup_{k=1}^{n} A_k}_{\in \mathcal{L}\text{を今示した}} \in \mathcal{L}$ より従う.

($A_n \cap A_m = \phi$)

(3)より. ∵ $B_n = \bigcup_{k=1}^{n} A_k$ は単調増大

($\Leftarrow$)は自分でチェックせよ.

(Dynkinの定理の証明)

証明の方針：$\mathcal{L}(\mathcal{P})$ が π-システムであることを示す。実は λ-システムがπ-システムであるなら、それは σ-加法族であることが示せる。つまり、$\sigma(\mathcal{P}) \subset \mathcal{L}(\mathcal{P})$ が成り立つ。

<u>Lem</u>

ある集合族 $\mathcal{B}$ が、π-システムでもλ-システムでもあるとき、$\mathcal{B}$ は σ-加法族である。

(証明)

(1) $\Omega \in \mathcal{B}$ はλ-システムの定義から明らか。

(2) $A \in \mathcal{B}$ なら $A^c \in \mathcal{B}$ であることも性質(2)′より従う。

(3) まず有限和について閉じていることを示す。

$A, B \in \mathcal{L}$ とする。AとBは互いに疎であるとは限らないので(3)′を使えない。しかし、$\mathcal{B}$ はπ-システムなので、$A \cap B \in \mathcal{B}$ である。よって、λ-システムの性質より $A \setminus (A \cap B) \in \mathcal{B}$, $B \setminus (A \cap B) \in \mathcal{B}$ でもある。また、(1),(2)より $\phi = (\Omega)^c \in \mathcal{B}$ であるので、λ-システムの性質(3)′より、

$$A \cup B = \underbrace{\underset{\in \mathcal{B}}{(A \setminus (A \cap B))} \cup \underset{\in \mathcal{B}}{(B \setminus (A \cap B))} \cup \underset{\in \mathcal{B}}{(A \cap B)} \cup \underset{\in \mathcal{B}}{\phi} \cup \underset{\in \mathcal{B}}{\phi} \cdots}_{\text{互いに排反}}$$

$$\in \mathcal{B}$$

である。このことから $A_1, A_2, \ldots \in \mathcal{B}$ なら、($A_n \cap A_m = \phi$ $(n \neq m)$ とは限らない)

$\bigcup_{n=1}^{m} A_n \in \mathcal{B}$ である。よって、

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{m=1}^{\infty} \underbrace{\left(\bigcup_{n=1}^{m} A_n\right)}_{A'_m \text{とおく}} = \bigcup_{m=1}^{\infty} \underset{\uparrow\ \mathcal{L}\text{に含まれる}}{A'_m} \in \mathcal{B}.$$

なお、最後の $\bigcup_{m=1}^{\infty} A'_m \in \mathcal{B}$ は $A'_1 \subset A'_2 \subset \cdots \in \mathcal{B}$ に性質(3)を適用した。

以上より、$\mathcal{B}$ は σ-加法族である。 //

今から、$\mathcal{L}(\mathcal{P})$ が π-システムになることを示す。前の Lem より $\mathcal{L}(\mathcal{P})$ が σ-加法族になるから

これが示せれば、$\mathcal{L}(\mathcal{P}) \supset \sigma(\mathcal{P})$ であり、かつ、$\mathcal{L}$ と $\mathcal{L}(\mathcal{P})$ の定義より、
$\mathcal{L} \supset \mathcal{L}(\mathcal{P})$ である。よって、$\mathcal{L} \supset \sigma(\mathcal{P})$ が従う。

$\mathcal{L}(\mathcal{P})$ が π-システムであることの証明：

$$\mathcal{G}_A := \{ B \in \sigma(\mathcal{P}) \mid A \cap B \in \mathcal{L}(\mathcal{P}) \}$$ とする。

(i) $A \in \mathcal{L}(\mathcal{P})$ なら $\mathcal{G}_A$ は λ-システムである。
なぜなら、

(1)′ $\Omega \in \mathcal{G}_A$ $\quad (\because \Omega \cap A = A \in \mathcal{L}(\mathcal{P}))$ Aの仮定

(2)′ $B \in \mathcal{G}_A$ なら、$A \cap B \in \mathcal{L}(\mathcal{P})$ であり、$A \in \mathcal{L}(\mathcal{P})$ であることとあわせて、$A \setminus (A \cap B) = A \cap B^c \in \mathcal{L}(\mathcal{P})$ でもある。
つまり、$B^c \in \mathcal{G}_A$ である。

(3)′ $B_n \in \mathcal{G}_A$ が互いに排反であるとする。$\bigcup_{n=1}^{\infty} B_n \in \mathcal{G}_A$ を示したい。
$A \cap \left(\bigcup_{n=1}^{\infty} B_n \right) = \bigcup_{n=1}^{\infty} (A \cap B_n)$ であるが、$B_n \in \mathcal{G}_A$ より、$A \cap B_n \in \mathcal{L}(\mathcal{P})$ である。しかも、$(A \cap B_n)_{n=1}^{\infty}$ は互いに排反であるので、
$\bigcup_{n=1}^{\infty} (A \cap B_n) \in \mathcal{L}(\mathcal{P})$、つまり、$\bigcup_{n=1}^{\infty} B_n \in \mathcal{G}_A$ である。

よって $\mathcal{G}_A$ は λ-システムである。

(ii) $A \in \mathcal{P}$ なら $\mathcal{L}(\mathcal{P}) \subset \mathcal{G}_A$ である。
なぜなら、まず $A \in \mathcal{L}(\mathcal{P})$ でもあることから (i) より $\mathcal{G}_A$ は λ-システムであることに注意する。今、$\mathcal{P}$ は π-システムなので ${}^{\forall} B \in \mathcal{P}$ は $B \cap A \in \mathcal{P}$ である。つまり、${}^{\forall} B \in \mathcal{P}$ は $B \in \mathcal{G}_A$ である $(\mathcal{P} \subset \mathcal{G}_A)$。よって $\mathcal{G}_A$ は $\mathcal{P}$ を含む λ-システムなので、$\mathcal{L}(\mathcal{P}) \subset \mathcal{G}_A$ である。

このことから、$A \in \mathcal{P}$, $B \in \mathcal{L}(\mathcal{P})$ に対し、$A \cap B \in \mathcal{L}(\mathcal{P})$ である。
あとは A を $A \in \mathcal{L}(\mathcal{P})$ まで広げたい。

(iii) $A \in \mathcal{L}(\mathcal{P})$ なら、$\mathcal{L}(\mathcal{P}) \subset \mathcal{G}_A$ である。
なぜなら、(ii) より、${}^{\forall} B \in \mathcal{P}$ に対し、$A \cap B \in \mathcal{L}(\mathcal{P})$ であることは示されているので、$\mathcal{P} \subset \mathcal{G}_A$ である。また (i) より $A \in \mathcal{L}(\mathcal{P})$ なら $\mathcal{G}_A$ は λ-システムなので、
$\mathcal{L}(\mathcal{P}) \subset \mathcal{G}_A$ でもある。よって、$A \in \mathcal{L}(\mathcal{P})$ なら ${}^{\forall} B \in \mathcal{L}(\mathcal{P})$ で $A \cap B \in \mathcal{L}(\mathcal{P})$ をみたす。
つまり、$\mathcal{L}(\mathcal{P})$ は π-システムである。

Def（多変量確率変数, 多次元確率変数, 確率ベクトル）

$(\Omega, \mathcal{F})$: 可測空間

$X=(X_1,\dots,X_n): \Omega \to \mathbb{R}^n$ が多変量確率変数（確率ベクトル）

$\iff$ $X_i: \Omega \to \mathbb{R}$ が確率変数, i.e. $X_i^{-1}((-\infty,a]) \in \mathcal{F}$ $({}^\forall a \in \mathbb{R})$
$(i=1,\dots,n)$ //

別の定義

。「$X: \Omega \to \mathbb{R}^n$ が確率ベクトル $\iff$ $X^{-1}(A) \in \mathcal{F}$ $({}^\forall A \in \mathcal{B}(\mathbb{R}^n))$」

。つまり $X: \Omega \to \mathbb{R}^n$ が $\mathcal{F}/\mathcal{B}(\mathbb{R}^n)$-可測.

(下$\Rightarrow$上): $\pi_k(x) = x_k$ とすれば $X_k = \pi_k \circ X$ である.
π_k は連続関数なので可測. 可測関数の合成は可測
なので X_k も可測 $({}^\forall k=1,\dots,n)$.

(上$\Rightarrow$下): $\mathcal{A} = \{ I_1 \times \cdots \times I_n \subset \mathbb{R}^n \mid I_k = (a_k, b_k] \ (a_k \le b_k) \}$
とすると. $A = I_1 \times \cdots \times I_n \in \mathcal{A}$ に対し. $X^{-1}(A) = \bigcap_{k=1}^{n} \underbrace{X_k^{-1}(I_k)}_{}\in \mathcal{F}$ である.
↑ $\mathcal{F}$ なので

あとは. $\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R}^n)$ と $X^{-1}(\sigma(\mathcal{A})) = \sigma(X^{-1}(\mathcal{A}))$ ←（第2回の演習）

に気をつければ.

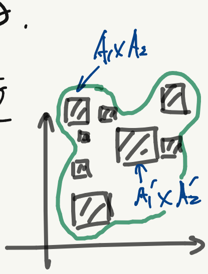
集合族 $\mathcal{A}$ に対し.
$X^{-1}(\mathcal{A}) = \{ X^{-1}(A) \mid A \in \mathcal{A} \}$
と書く.

$$
\begin{aligned}
X^{-1}(\mathcal{B}(\mathbb{R}^n)) &= X^{-1}(\sigma(\mathcal{A})) \\
&= \sigma(X^{-1}(\mathcal{A})) \\
&\subset \mathcal{F}
\end{aligned}
$$

を得る. //

。直積σ-加法族: $(\Omega_k, \mathcal{F}_k)$ $(k=1,\dots,n)$ を可測空間とする.

$\Omega_1 \times \cdots \times \Omega_n$ 上の σ-加法族 $\mathcal{F} = \mathcal{F}_1 \times \cdots \times \mathcal{F}_n = \bigtimes_{k=1}^{n} \mathcal{F}_k$ を
$\{ A_1 \times \cdots \times A_n \mid A_k \in \mathcal{F}_k \}$ を含む最小の σ-加法族とする.
これを. 直積σ-加法族と呼ぶ.

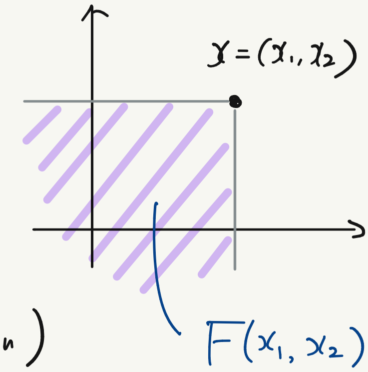


$\longrightarrow$ 実は. $\mathcal{B}(\mathbb{R}^n) = \mathcal{B}(\mathbb{R}) \times \cdots \times \mathcal{B}(\mathbb{R})$ である.（演習問題）

Def（同時分布関数）

$(\Omega, \mathcal{F}, P)$：確率空間

$X = (X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ を確率ベクトルとする。



$x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ について、

$$F(x) = F(x_1, \ldots, x_n) \overset{def}{=} P\left(\bigcap_{k=1}^{n} \{X_k \leq x_k\}\right)$$
$$= P(X_1 \leq x_1, \ldots, X_n \leq x_n)$$

としたとき、F を $X = (X_1, \ldots, X_n)$ の同時分布関数と言う。
(joint distribution function)

特に

$$F(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(u_1, \ldots, u_n)\, du_1 \cdots du_n$$

（ルベーグ積分）

と、ある非負関数 f を用いて書ける時、
f を同時確率密度関数と言う。
(joint probability density function)

※（Radon-Nikodym の定理）
F がルベーグ測度について
絶対連続の時。

//

Cor（同時分布関数の性質）

1. $a_1 \leq a_2,\ b_1 \leq b_2 \implies F(a_1, b_1) \leq F(a_2, b_2)$ （単調性）

2. $\displaystyle\lim_{\substack{x_1 \searrow a+0 \\ x_2 \searrow b+0}} F(x_1, x_2) = F(a, b)$ （右連続性）

3. $\displaystyle\lim_{\substack{x_1 \to +\infty \\ x_2 \to +\infty}} F(x_1, x_2) = 1$, $\displaystyle\lim_{\substack{x_1 \to -\infty \\ x_2 : fix}} F(x_1, x_2) = 0$, $\displaystyle\lim_{\substack{x_2 \to -\infty \\ x_1 : fix}} F(x_1, x_2) = 0$

（・確率の連続性より極限の取り方によらない
・$n \geq 3$ についても同様の性質が成り立つ

//

Remark 1次元の場合と同様にして、性質1,2,3を満たす F が与えられれば、
$(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ の確率測度が一意に決まる。（π-λ定理, Hopfの拡張定理）

Def (周辺分布)

$F_1(x_1) = F(x_1, +\infty) = P(X_1 \leq x_1)$ $\left(= P(X_1 \leq x_1 \cap X_2 \leq \infty)\right)$

を. X_1 の周辺分布関数という.

(marginal distribution function)

$F_2(x_2) = P(X_2 \leq x_2)$ も同様

より一般に $F_k(x_k) = P(X_k \leq x_k)$ を X_k の周辺分布関数と言う

$F_k(x_k) = \int_{-\infty}^{x_k} f_k(u_k)\, du_k$ と非負(可測)関数 f_k を用いて書ける時.

f_k を周辺確率密度関数という

同時分布が絶対連続なら. 周辺密度は.

$$f_k(x_k) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(u_1, \ldots, u_{k-1}, x_k, u_{k+1}, \ldots, u_n)\, du_1 \cdots du_{k-1}\, du_{k+1} \cdots du_n$$

これだけ固定

で与えられる.

Def (確率変数の独立性)

$X_1,\dots,X_n$: r.v. が独立

$\overset{def}{\iff}$ $F(x_1,\dots,x_n)=F_1(x_1)\cdots F_n(x_n)$ $\quad ({}^\forall(x_1,\dots,x_n)\in\mathbb{R}^n)$

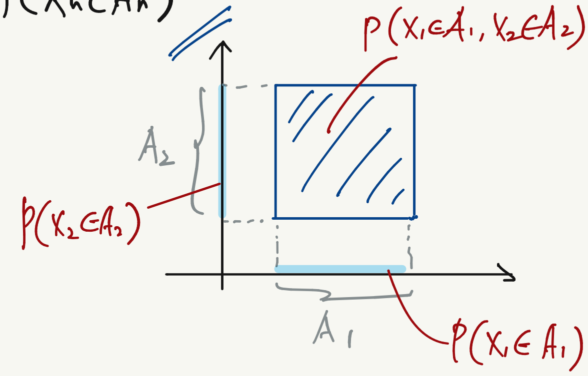
と分解できる。 //

同値な定義がいくつかある。以下の定義はその中でもわかりやすい定義である。

Prop

$X_1,\dots,X_n$: r.v. が独立 $\iff P(X_1\in A_1,\dots,X_n\in A_n)$ $\left(=P\left(\bigcap_{k=1}^{n}\{X_k\le x_k\}\right)\right)$ のこと

$= P(X_1\in A_1)\cdots P(X_n\in A_n)$ //



(証明)

下の定義が上の定義を与えることはすぐわかる。

逆を示したい。

そのために π-λ定理を用いる。(再掲, 前ページおよび前回補足資料参照)

- π-システム: 集合族 $\mathcal{P}$ が π-システム $\iff$ $A,B\in\mathcal{P}$ なら $A\cap B\in\mathcal{P}$
- λ-システム: 集合族 $\mathcal{L}$ が λ-システム $\iff$
  - (1) $\Omega\in\mathcal{L}$
  - (2) $A,B\in\mathcal{L},\ A\subset B \Rightarrow \underbrace{B\setminus A}_{=B\cap A^c}\in\mathcal{L}$
  - (3) $A_1\subset A_2\subset A_3\subset\cdots\in\mathcal{L}$ $\Rightarrow \bigcup_{n=1}^{\infty}A_n\in\mathcal{L}$

注: σ-加法族はπ-システムでもλ-システムでもある。

Thm (Dynkinのπ-λ定理)

あるλ-システム $\mathcal{L}$ が、あるπ-システム $\mathcal{P}$ を含む $(\mathcal{P}\subset\mathcal{L})$ なら、

$\sigma(\mathcal{P})\subset\mathcal{L}$ である。 //

← 証明は前回の補足資料を参照のこと。

2変数で示す。多変数の場合は同様に帰納法で示せる。

$\mathcal{A}_1=\{(-\infty,x_1]\mid x_1\in\mathbb{R}\}$, $\mathcal{A}_2=\{(-\infty,x_2]\mid x_2\in\mathbb{R}\}$ とする。

${}^\forall A\in\mathcal{A}_1$, ${}^\forall B\in\mathcal{A}_2$ に対し、$P(X_1\in A, X_2\in B)=P(X_1\in A)P(X_2\in B)$ で、

また、$\sigma(\mathcal{A}_1)=\mathcal{B}(\mathbb{R})$, $\sigma(\mathcal{A}_2)=\mathcal{B}(\mathbb{R})$ であることにも注意する。

今. $A_2 \in \mathcal{A}_2$ を1つ固定して.

$$\mathcal{L} = \{ A \in \mathcal{B}(\mathbb{R}) \mid P(X_1 \in A \cap X_2 \in A_2) = P(X_1 \in A) P(X_2 \in A_2) \}$$

とする. $\mathcal{L} \supset \mathcal{A}_1$ であることは独立性の定義から従う.

$\mathcal{L}$ がλ-システムであることが示せれば. π-λ定理より. $\sigma(\mathcal{A}_1) = \mathcal{B}(\mathbb{R}) \subset \mathcal{L}$ が言える.

(a) $\Omega = \mathbb{R} \in \mathcal{L}$. なぜなら

$$P(X_1 \in \Omega, X_2 \in A) = P(X_2 \in A) = \underbrace{P(X_1 \in \Omega)}_{=1} P(X_2 \in A)$$

(b) $A, B \in \mathcal{L}$ かつ $A \subset B$ とする. このとき.

$$\begin{aligned} P(X_1 \in (B \setminus A), X_2 \in A_2) &= P((\{X_1 \in B\} \cap \{X_2 \in A_2\}) \setminus (\{X_1 \in A\} \cap \{X_2 \in A_2\})) \\ &= P(X_1 \in B, X_2 \in A_2) - P(X_1 \in A, X_2 \in A_2) \\ &\overset{A, B \in \mathcal{L} \text{ より}}{=} P(X_1 \in B) P(X_2 \in A_2) - P(X_1 \in A) \cdot P(X_2 \in A_2) \\ &= (P(X_1 \in B) - P(X_1 \in A)) P(X_2 \in A_2) \\ &= P(X_1 \in B \setminus A) \cdot P(X_2 \in A_2) \end{aligned}$$

よって. $B \setminus A \in \mathcal{L}$

(c) $\hat{A}_1 \subset \hat{A}_2 \subset \cdots \in \mathcal{L}$ のとき.

$X^{-1}(\bigcup A_n) = \bigcup X^{-1}(A_n)$ より.

$$\begin{aligned} P(X_1 \in \bigcup_{n=1}^{\infty} \hat{A}_n, X_2 \in A_2) &= P\left[\left(\bigcup_{n=1}^{\infty} \{X_1 \in \hat{A}_n\}\right) \cap \{X_2 \in A_2\}\right] \\ &= P\left(\bigcup_{n=1}^{\infty} (\{X_1 \in \hat{A}_n\} \cap \{X_2 \in A_2\})\right) \\ &= \lim_{n \to \infty} P(\{X_1 \in \hat{A}_n\} \cap \{X_2 \in A_2\}) \quad (\because \text{確率の連続性と. } \hat{A}_n \text{の単調性}) \\ &= \lim_{n \to \infty} P(X_1 \in \hat{A}_n) \cdot P(X_2 \in A_2) \\ &= P(X_1 \in \bigcup_{n=1}^{\infty} \hat{A}_n) \cdot P(X_2 \in A_2) \end{aligned}$$

ここでも $\bigcup_{n=1}^{\infty} X^{-1}(A_n) = X^{-1}(\bigcup_{n=1}^{\infty} A_n)$ を使う

よって. $\bigcup_{n=1}^{\infty} \hat{A}_n \in \mathcal{L}$ である.

以上より. $\mathcal{L}$ はλ-システムなので. $\sigma(\mathcal{A}_1) \subset \mathcal{L}$ が言えた.

つまり. $\forall A \in \mathcal{B}(\mathbb{R}), \forall B \in \mathcal{A}_2$ に対し. $P(A \in X_1, B \in X_2) = P(A \in X_1) \cdot P(B \in X_2)$.

さらに. 同じ議論を. $A \in \mathcal{B}(\mathbb{R})$ を固定して. $\mathcal{A}_2$ の方に適用し.

$\sigma(\mathcal{A}_1) = \mathcal{B}(\mathbb{R})$ と $\sigma(\mathcal{A}_2) = \mathcal{B}(\mathbb{R})$ まで拡張すれば良い. //

(独立性について続き.)

<u>Def</u> (事象の独立性)

$A_1, A_2, \dots, A_n \in \mathcal{F}$ が独立 $\Longleftrightarrow$ 任意の部分集合 $I \subset \{1,2,\dots,n\}$ に対し.

$$P\left(\bigcap_{k\in I} A_k\right) = \prod_{k\in I} P(A_k)$$

//

先の r.v. の独立性は. 事象の独立性の言葉を用いると.

$X_1, \dots, X_n$ が独立 $\Longleftrightarrow$ $X_1^{-1}(A_1), \dots, X_n^{-1}(A_n)$ が任意の $A_1, \dots, A_n \in \mathcal{B}(\mathbb{R})$ に対して独立.

と言いかえられる.

<u>Def</u> (σ-加法族の独立性)

同様に集合族 $\mathcal{A}_1, \dots, \mathcal{A}_n$ の独立性も定義できる (σ-加法族とは限らない)

Ω上の n 個のσ-加法族 $\mathcal{F}_1, \dots, \mathcal{F}_n$ が独立.

$\Longleftrightarrow$ 任意の $A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2, \dots, A_n \in \mathcal{F}_n$ が独立. //

<u>Def</u> (確率変数によって生成されるσ-加法族)

$X : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (r.v.)

に対し.

$$\sigma(X) \overset{\text{def}}{=} X^{-1}(\mathcal{B}(\mathbb{R})) = \{X^{-1}(A) \mid A \in \mathcal{B}(\mathbb{R})\}$$

(ボレル集合の逆像で作られるσ-加法族)

と書く. これを. X の生成するσ-加法族と言う //

すると.

$X_1, \dots, X_n$ (r.v.) が独立 $\Longleftrightarrow$ $\sigma(X_1), \dots, \sigma(X_n)$ が独立

でもある.

実は. 先の X の独立性の同値性と同じ議論により. 次のことも示せる.

<u>Thm</u> (参考)

$\mathcal{A}_1, \dots, \mathcal{A}_n$ を Ω における非空な集合族とする.

(1) $\mathcal{A}_k$ は π-システム $(k=1,2,\dots,n)$
(2) $\mathcal{A}_1, \dots, \mathcal{A}_n$ は独立.

$\Longrightarrow$ $\sigma(\mathcal{A}_1), \dots, \sigma(\mathcal{A}_n)$ も独立. //

密度関数を持つ場合は、

$$f(x_1,\dots,x_n)=f_1(x_1)\cdots f_n(x_n)$$

と書けることが独立性の必要十分条件になる。

Ex. (正規分布)

$\mu=\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}\in\mathbb{R}^2$, $\Sigma=\begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{bmatrix}\in\mathbb{R}^{2\times2}$: 分散共分散行列

: 平均

$(\Sigma_{ij}=\mathbb{E}[(X_i-\mu_i)(X_j-\mu_j)])$

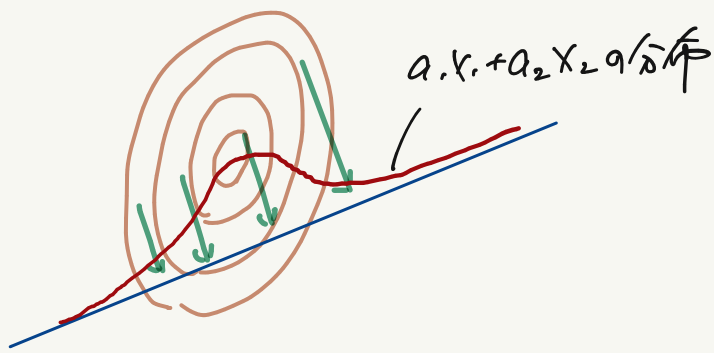
同時密度: $f(x_1,x_2)=\dfrac{1}{\sqrt{(2\pi)^2|\Sigma|}}\exp\left(-\dfrac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$

周辺密度: $f_1(x_1)=\dfrac{1}{\sqrt{2\pi\Sigma_{11}}}\exp\left(-\dfrac{(x_1-\mu_1)^2}{2\Sigma_{11}}\right)$

☆ X_1 と X_2 が独立 $\Leftrightarrow$ $\Sigma_{12}=\Sigma_{21}=0$

(参考): 実は、「$X=(X_1,X_2)$ が2変量正規分布

$\Leftrightarrow$ $\forall a_1,a_2\in\mathbb{R}$ に対し、$a_1X_1+a_2X_2$ が正規分布」



条件付き確率

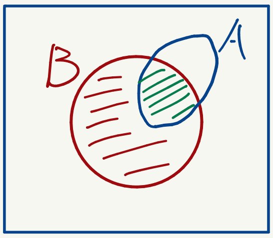
例) B: 身長が170cm以上
A: 年収が500万以上

Def (条件付き確率)

事象Bが起こったもとでの事象Aが起きる確率 $(A, B \in \mathcal{F})$

$$P(A|B) \overset{\text{def}}{=} \frac{P(A \cap B)}{P(B)}.$$

ただし、$P(B) \neq 0$ とする。



B: 固定のもとで、$A \in \mathcal{F} \longmapsto P(A|B)$ は $(\Omega, \mathcal{F})$ 上の確率測度になっていることはすぐに確認できる。

Remark この定義だと、$P(B)=0$ の場合は well-defined にできない。
そのような状況は連続な r.v. X に対し、$B = \{X = x\}$ とした時等に起きる。そのような状況も扱える条件付き確率の定義は後で述べる。

Cor (条件付き確率の公式)

(1) 積の公式

$$P(A \cap B) = P(A|B)P(B)$$

(2) Bayesの公式

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

原因から結果の確率がわかれば、結果から原因の確率を逆算できる。

特に、$A_1, \ldots, A_n$ が互いに排反かつ $\bigcup_{i=1}^{n} A_i = \Omega$ なら

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

例1　人口の0.1%がかかる病気がある。
ある検査では、病気の患者の99%が陽性を示し、
健常者の10%が陽性を示す。

この検査を受けてみて陽性だった時、本当に病気である確率は？

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|A^c)P(A^c)}$$

（A：病気　B：陽性）

$$= \frac{0.99\times 0.001}{0.99\times 0.001 + 0.1\times 0.999}$$

$$\fallingdotseq \underline{0.0098}$$ （1%にもみたない）

Remark　ベイズの定理はロボットの自己位置推定や天気予報等にも使われている。