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

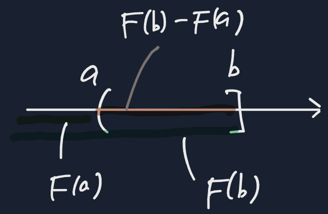
が与えられたとする.

(ある確率変数 X に対し, $P(X\leq x)=F(x)$ を想定)

↳ そのためには $\{\omega \mid X(\omega)\leq x\}$ 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 $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ 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.



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

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

前回補足資料

Q: どの集合族まで P 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)

$[F(b), F(a)]$ のコンパクト性を使う.

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

⇒ P は $\mathcal{S}$ 上で σ-加法的 と言う.

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

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.

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

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

○ こうして拡張されたPおよび、その定義域である$\mathcal{B}(\mathbb{R})$の性質を抜き出して書き下すと、

・$\mathcal{B}(\mathbb{R})$はσ-加法族($\mathcal{F}$)
$$\begin{cases} (1)\ \Omega \in \mathcal{F} \\ (2)\ A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F} \\ (3)\ A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F} \end{cases}$$

$(\mathbb{R}, \mathcal{B}(\mathbb{R}))$の組を一般化して$(\Omega, \mathcal{F})$：可測空間と書く。

・Pは確率測度
$$\begin{cases} (1)\ {}^{\forall}A \in \mathcal{F},\ 0 \leq P(A) \leq 1 \\ (2)\ P(\Omega) = 1 \\ (3)\ A_1, A_2, \ldots \in \mathcal{F}\ (\text{互いに排反}) \\ \quad \Rightarrow P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) \end{cases}$$

$(\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$の組を一般化して$(\Omega, \mathcal{F}, P)$と書き、確率空間と言う。

$\mathcal{S}$上の測度Pを$\sigma(\mathcal{S})$まで拡張する方法：

任意の集合$A \subset \mathbb{R}$に対し、($A \in \mathcal{B}(\mathbb{R})$等の条件は課さない)

外測度P^*を

$$\boxed{P^*(A) := \inf_{\substack{A \subset \bigcup_{i=1}^{\infty} B_i \\ B_1, B_2, \ldots \in \mathcal{S}}} \sum_{i=1}^{\infty} P(B_i)}$$

可算無限個 → $B_1, B_2, \ldots \in \mathcal{S}$

とおく。（$P^*(A) = \inf_{\substack{B_1, \ldots, B_m \in \mathcal{S}\ (\text{有限個}) \\ \text{ただし}\ A \subset \bigcup_{i=1}^{m} B_i\ (m\text{は任意})}} \sum_{i=1}^{m} P(B_i)$ とおくと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。）

ここで、

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

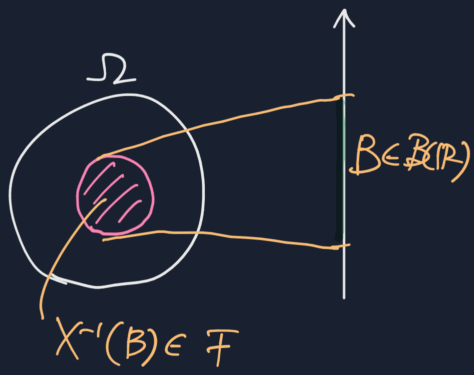
とおくと、$\mathcal{D}$はσ-加法族になり、P^*は$\mathcal{D}$上の確率測度になる。（非自明，前回補足資料）

さらに、$\mathcal{S} \subset \mathcal{D}$が容易に示せるので、$\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{S}) \subset \mathcal{D}$でもある。

$P(A) = P^*(A)$ $({}^{\forall}A \in \mathcal{S})$も示せるので、$P^*$は$P$の$\mathcal{B}(\mathbb{R})$への拡張を与えている。拡張の一意性はDynkinのπ-λ定理により保証される。（前回の補足資料）

$(\Omega, \mathcal{F})$：可測空間

$X: \Omega \to \mathbb{R}$ が $X^{-1}((-\infty, x]) = \{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{F}$ を
みたすとき、Xを確率変数と呼んだ。



$\Longleftrightarrow$ 同値な定義　$\forall B \in \mathcal{B}(\mathbb{R})$ に対し、$X^{-1}(B) \in \mathcal{F}$.

- $\{X \leq x\}$ なる事象は可測であって欲しい。
  そうでなければ分布関数も定義できない。
- $\{X \leq x\}$ が可測なら、自動的に $\forall B \in \mathcal{B}(\mathbb{R})$ に対し、
  $X^{-1}(B) \in \mathcal{F}$ が成り立ってしまう。

これは次のような論理によるものだった：

ある集合族 $\mathcal{A}$ に対し、$X^{-1}(A) \in \mathcal{F}$ $(\forall A \in \mathcal{A})$ が成り立つ時、

「$\mathcal{A}$を含む最小のσ-加法族 $\sigma(\mathcal{A})$ に対しても、
$X^{-1}(A) \in \mathcal{F}$ $(\forall A \in \sigma(\mathcal{A}))$ が成り立つ。」　← 前回講義と演習を見よ。

$\mathcal{A} = \{(-\infty, x] \mid x \in \mathbb{R}\}$ とすれば、

$\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R})$ なので、上の主張が成り立つ。

※ ある部分集合族 $\mathcal{A}$ で成り立つ性質が $\mathcal{A}$ を含む最小のσ-加法族まで拡張される、というのは測度論で頻出の論法。

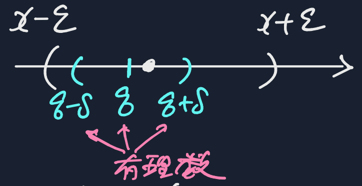
Lem

開区間の全体 $\mathcal{A}=\{(a,b)\mid a\le b,\ a,b\in\mathbb{R}\}$ に対し、$\sigma(\mathcal{A})=\mathcal{B}(\mathbb{R})$.

//

(∵) $\mathbb{R}$ 内で有理数が稠密であることを使う。
$\sigma(\mathcal{A})$ が任意の開集合を含むことを示せば、$\mathcal{B}(\mathbb{R})$ の定義より十分である。
任意の開集合 $B\subset\mathbb{R}$ は可算無限個の開区間 $(a-\delta, a+\delta)$ $(a\in\mathbb{Q},\ \delta\in\mathbb{Q})$ の和集合で書けることが示せる。これが示せれば十分である。

それを示す。任意の $x\in B$ に対し、B は開集合であることから、
ある $\varepsilon>0$ が存在し、$(x-\varepsilon, x+\varepsilon)\subset B$ とできる。



すると、ある有理数 q と有理数 $\delta>0$ が存在し、

$$x\in(q-\delta, q+\delta)\subset(x-\varepsilon, x+\varepsilon)\subset B$$

とできる。(たとえば $0<\delta<\frac{\varepsilon}{2}$ を1つとってきて、$(x-\delta, x+\delta)$ の中から有理数 q を1つ取って来れば良い)

このように、各 x ごとに q_x, δ_x を取ってきて、$I_x=(q_x-\delta_x, q_x+\delta_x)$ とおけば、
$B\subset\bigcup_{x\in B} I_x$ である。一方、$I_x\subset B$ $({}^{\forall}x\in B)$ より $\bigcup_{x\in B} I_x\subset B$ なので $B=\bigcup_{x\in B} I_x$
である。一方、q_x, δ_x はともに有理数なので、$\bigcup_{x\in B} I_x$ は高々可算個の区間 $(q-\delta, q+\delta)$ の和集合。よって、${}^{\forall}B$: 開集合 は $B\in\sigma(\mathcal{A})$ である。 //

このようにある距離空間(又は位相空間)が
可算な稠密部分集合をもつとき、「可分である」と言う。

(可分な距離空間, 可分な位相空間)

これまでの復習. ここまで.

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

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

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

別の定義

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

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

（下 ⇒ 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 $({}^\forall k = 1, \ldots, n)$．

（上 ⇒ 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．$X^{-1}(A) = \bigcap_{k=1}^{n} \underline{X_k^{-1}(I_k)} \in \mathcal{F}$ 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．

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

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

を得る． //

。直積 σ-加法族：$(\Omega_k, \mathcal{F}_k)$ $(k = 1, \ldots, n)$ 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）

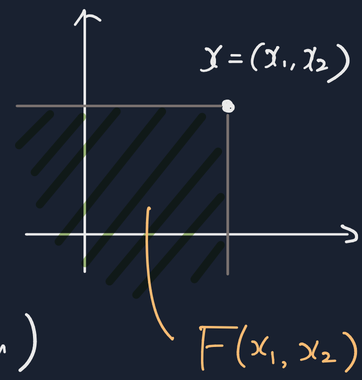
Def（同時分布関数）

$(\Omega, \mathcal{F}, P)$：確率空間

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

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

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



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

特に

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

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

（・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）

前回 補足資料を参照

Def（周辺分布）

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

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）

$X_1, \ldots, X_n$: r.v. が独立

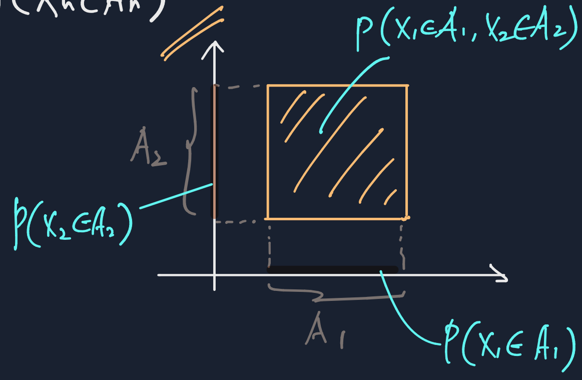
$\overset{\text{def}}{\Longleftrightarrow}$ $F(x_1, \ldots, x_n) = F_1(x_1) \cdots F_n(x_n) \quad ({}^\forall (x_1, \ldots, x_n) \in \mathbb{R}^n)$

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. $X_1, \ldots, X_n$: r.v. が独立 $\Longleftrightarrow$ $P(X_1 \in A_1, \ldots, X_n \in A_n) \left(= P\left(\bigcap_{k=1}^{n} \{X_k \in A_k\}\right) \text{のこと} \right)$

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



（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。（前回補足資料参照）

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

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.

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

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.

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

こっちも $\bigcup_{n=1}^{\infty} X^{-1}(A_n) = X^{-1}\Big(\bigcup_{n=1}^{\infty} A_n\Big)$ 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)

Def (事象の独立性)

$A_1, A_2, \dots, A_n \in \mathcal{F}$ が独立 $\iff$ 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 $\iff$ $X_1^{-1}(A_1),\dots,X_n^{-1}(A_n)$ 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。

Def (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.

$\iff$ 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.

Def (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 $\iff$ $\sigma(X_1),\dots,\sigma(X_n)$ 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。

Thm (参考)

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

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

$\Longrightarrow$ $\sigma(\mathcal{A}_1),\dots,\sigma(\mathcal{A}_n)$ 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)

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

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

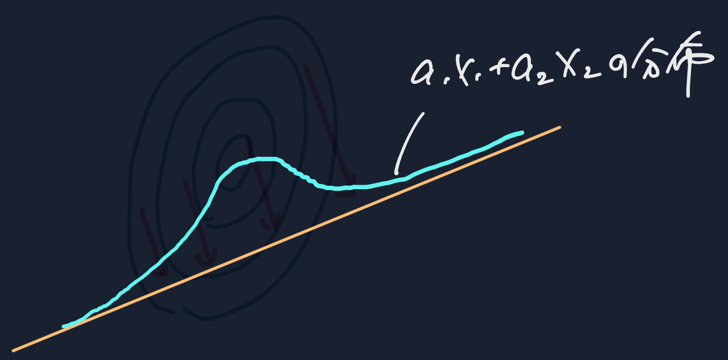
同時密度: $f(x_1,x_2)=\dfrac{1}{\sqrt{(2\pi)^2|\Sigma|}}\exp\left(-\dfrac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$

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, $A_1, \dots, A_n$ 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？

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

病気　陽性

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

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

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。