

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 $F(x)$:
- (1) $x < y \Rightarrow F(x) \leq F(y)$
 - (2) $\lim_{x \rightarrow a+0} F(x) = F(a)$ (右連続)
 - (3) $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow +\infty} F(x) = 1$

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

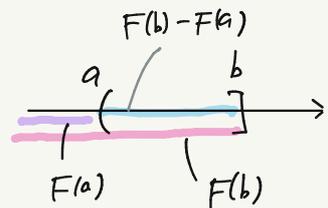
↳ そのためには
 $\{\omega \mid X(\omega) \leq x\}$
 は可測) なくては
 行かない \rightarrow 確率変数の定義

が与えられたとする。

- F に対応する $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の確率測度 P は 一意に存在する。
 かつ、 $P((-\infty, a]) = F(a)$ ($\forall a \in \mathbb{R}$) を満たす P は存在。

- 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 "集合代数" となっている。
中 前回補足資料



Q: この集合族に対して P を値なく拡張できる?

- ② $(a, b] = \bigcup_{i=1}^{\infty} (a_i, b_i]$ と互いに重ならない区間 $(a_i, b_i]$ ($i=1, 2, \dots$) に分解した時、

P は $P((a, b]) = \sum_{i=1}^{\infty} P((a_i, b_i])$ を満たすことを示せる。(非自明)
 $(F(b) - F(a) = \sum_{i=1}^{\infty} (F(b_i) - F(a_i)))$
 $[F(b), F(a)]$ のコンパクト性を使う。

$\Rightarrow P$ は \mathcal{S} 上での σ -加法的 と言う。

- $0 \leq P(A) \leq 1$ ($\forall A \in \mathcal{S}$)
- $P(\mathbb{R}) = 1$
- \mathcal{S} 上での σ -加法的

- ③ Hoeffding の拡張定理 (の一般化):

上の3つの性質を満たす \mathcal{S} 上の集合関数 $P: \mathcal{S} \rightarrow [0, 1]$ は \mathcal{S} を含む最小の σ -加法族 $\sigma(\mathcal{S})$ 上の確率測度 μ 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、示すことができる。

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

$\mathcal{B}(\mathcal{R})$ は σ -加法族 (\mathcal{F})

- (1) $\Omega \in \mathcal{F}$
- (2) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
- (3) $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

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

\mathcal{P} は 確率測度

- (1) $\forall A \in \mathcal{F}, 0 \leq \mathcal{P}(A) \leq 1$
- (2) $\mathcal{P}(\Omega) = 1$
- (3) $A_1, A_2, \dots \in \mathcal{F}$ (互いに排互)
 $\Rightarrow \mathcal{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathcal{P}(A_n)$

($\mathcal{R}, \mathcal{B}(\mathcal{R}), \mathcal{P}$) の組を一般化した ($\Omega, \mathcal{F}, \mathcal{P}$) と書き、確率空間 と書く。

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

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

外測度 \mathcal{P}^* を

$$\mathcal{P}^*(A) := \inf_{A \subset \bigcup_{i=1}^{\infty} B_i} \sum_{i=1}^{\infty} \mathcal{P}(B_i)$$

可算無限個 $\rightarrow B_1, B_2, \dots \in \mathcal{S}$



と置く。 ($\mathcal{P}^*(A) = \inf_{\substack{B_1, \dots, B_m \in \mathcal{S} \text{ (有限個)}}} \sum_{i=1}^m \mathcal{P}(B_i)$ とおくと Jordan 外測度と呼ぶものになるが、これは不十分である。)

==>

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

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

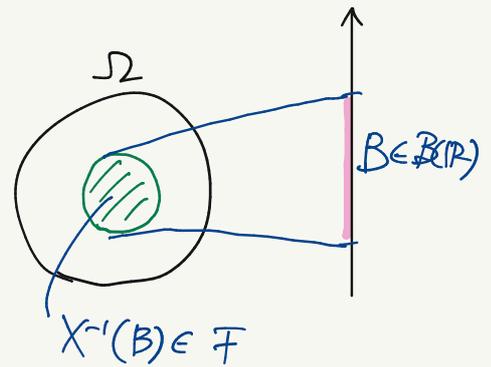
さらに、 $\mathcal{S} \subset \mathcal{D}$ が容易に示せるので、 $\mathcal{B}(\mathcal{R}) = \sigma(\mathcal{S}) \subset \mathcal{D}$ である。

$\mathcal{P}(A) = \mathcal{P}^*(A)$ ($\forall A \in \mathcal{S}$) も示せるので、 \mathcal{P}^* は \mathcal{P} の $\mathcal{B}(\mathcal{R})$ への拡張とみてよい。拡張の一意性は Dynkin の π - λ 定理により保証される。

(Ω, \mathcal{F}) : 可測空間

$X: \Omega \rightarrow \mathbb{R}$ 且 $X^{-1}((-\infty, x]) = \{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{F}$ であるとき、 X を確率変数と呼ぶ。

\iff 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ることが欲しい。
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、 $X^{-1}(B) \in \mathcal{F}$ が成り立つ。

これは次のような論理によるものだから:

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、

\square \mathcal{A} を含む最小の σ -代数族 $\sigma(\mathcal{A})$ に対しても、

$X^{-1}(A) \in \mathcal{F}$ ($\forall A \in \sigma(\mathcal{A})$) が成り立つ。 \square

\leftarrow 前回の講義と演習を見よ。

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

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

***** ある部分集合族 \mathcal{A} が成り立つ性質 \mathcal{A} を含む最小の σ -代数族 $\sigma(\mathcal{A})$ を拡張する時、この方法は測度論で類題の論法。

Lem

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

//

(-) \mathbb{R} 内の有理数 q が稠密であることを使う

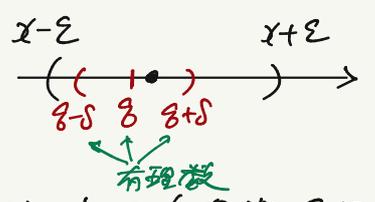
$\sigma(\mathcal{A})$ が任意の開集合 B を含むことを示せば, $\mathcal{B}(\mathbb{R})$ の定義より十分.

任意の開集合 $B \subset \mathbb{R}$ は可算無限個の開区間 $(a-\delta, a+\delta)$ ($a \in \mathbb{Q}, \delta \in \mathbb{Q}$) の和集合である. 要するにこれを示せば十分.

これを示す. 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. $(x-\varepsilon, x+\varepsilon) \subset B$ となる.

よって. ある有理数 q と有理数 $\delta > 0$ が存在し.



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

となる. (たとえば $0 < \delta < \frac{\varepsilon}{2} \in \mathbb{Q}$ と, $2 \leq 2$. $(x-\delta, x+\delta)$ の中に有理数 $q \in \mathbb{Q}$ を取り, 2 を満たす.)

このように. 各 x ごとに q_x, δ_x を取ると, $I_x = (q_x - \delta_x, q_x + \delta_x)$ をおけば.

$B \subset \bigcup_{x \in B} I_x$ 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. $\cup B$: 開集合は $\mathcal{B} \in \sigma(\mathcal{A})$ 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)

二つまでの復習. 二つまで.

Def (多変量確率変数, 多次元確率変数, 確率ベクトル)

(Ω, \mathcal{F}) : 可測空間

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

$\iff X_i: \Omega \rightarrow \mathbb{R}$ 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

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

◦ > 利. $X: \Omega \rightarrow \mathbb{R}^n$ 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 ($k=1, \dots, n$).

(上 \Rightarrow 下): $\mathcal{A} = \{I_1 \times \dots \times I_n \subset \mathbb{R}^n \mid I_k = (a_k, b_k] (a_k \leq b_k)\}$

とすると $A = I_1 \times \dots \times I_n \in \mathcal{A}$ に対し $X^{-1}(A) = \bigcap_{k=1}^n X_k^{-1}(I_k) \in \mathcal{F}$ である.
(\mathcal{F} による)

よって $\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R}^n)$ と $X^{-1}(\sigma(\mathcal{A})) = \sigma(X^{-1}(\mathcal{A}))$

に気づけば:

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

◦ 集合族 \mathcal{A} に対し (第2回の演習)

$$X^{-1}(\mathcal{A}) = \{X^{-1}(A) \mid A \in \mathcal{A}\}$$

と書く.

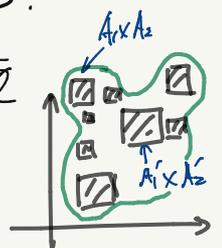
を介する. //

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

$\Omega_1 \times \dots \times \Omega_n$ 上の σ -加法族 $\mathcal{F} = \mathcal{F}_1 \times \dots \times \mathcal{F}_n = \bigotimes_{k=1}^n \mathcal{F}_k$ を

$\{A_1 \times \dots \times A_n \mid A_k \in \mathcal{F}_k\}$ 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.



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

Def (同時分布関数)

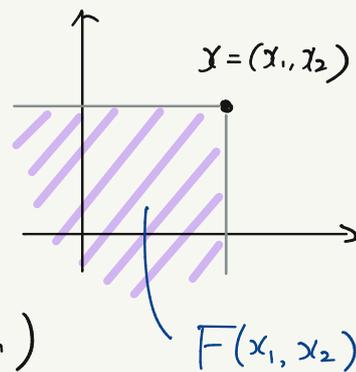
(Ω, \mathcal{F}, P) : 確率空間

$X = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$ ε 確率ベクトルである。

$x = (x_1, \dots, x_n) \in \mathbb{R}^n$ ε $\omega \in \Omega$.

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

$$= P(X_1 \leq x_1, \dots, X_n \leq x_n)$$



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

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

(ルビーク積分)

ε 、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)

ϕ (Radon-Nikodymの定理)
 F がルビーク測度に対する
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. $\lim_{\substack{x_1 \downarrow a+0 \\ x_2 \downarrow b+0}} F(x_1, x_2) = F(a, b)$ (右連続性)

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

- (確率の連続性より) 極限の取り方によって異なる
- $n \geq 3$ 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, Halmosの拡張定理)

Def (周辺分布)

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

を X_1 の 周辺分布関数 とする。
(marginal distribution function)

$$F_2(x_2) = P(X_2 \leq x_2) \text{ も同様}$$

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 とする

同時分布が絶対連続なら、周辺密度は、

$$f_k(x_k) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(u_1, \dots, \underbrace{u_{k-1}, x_k, u_{k+1}, \dots, u_n}_{= u_{k-1} \text{ 固定}}, \dots, u_n) du_1 \dots du_{k-1} du_{k+1} \dots du_n$$

と表す。

今 $A_2 \in \mathcal{A}_2$ を 1 つ固定して.

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

とある. $\mathcal{L} \supset \mathcal{A}_1$ 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は示せばいい. π - λ 定理より. $\sigma(\mathcal{A}_1) = \mathcal{B}(\mathcal{R}) \subset \mathcal{L}$ である.

(a) $\Omega = \mathcal{R} \in \mathcal{L}$. 任意に $P(X_1 \in \Omega, X_2 \in A) = P(X_2 \in A)$

$$= \frac{P(X_1 \in \Omega) P(X_2 \in A)}{1}$$

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

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

よって $B \setminus A \in \mathcal{L}$

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

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

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. $\sigma(\mathcal{A}_1) \subset \mathcal{L}$ である.

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

つまり. 同様に議論して. $A \in \mathcal{B}(\mathcal{R})$ を固定して. \mathcal{A}_2 の方に適用して.

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

(独立性について續く.)

Def (事象の独立性)

$A_1, A_2, \dots, A_n \in \mathcal{F}$ が独立 \Leftrightarrow 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。

X_1, \dots, X_n が独立 $\Leftrightarrow X_i^{-1}(A_i), \dots, X_n^{-1}(A_n)$ が任意の $A_1, \dots, A_n \in \mathcal{B}(\mathbb{R})$ に対し独立。

と書くことができる。

Def (σ -加法族の独立性)

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

\Leftrightarrow 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。 //

(σ -加法族と σ -環) \rightarrow

同様に集合族 $\mathcal{A}_1, \dots, \mathcal{A}_n$ の独立性も定義できる。

Def (確率変数による生成される σ -加法族)

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

に対し.

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

と書く。これを、 X の生成する σ -加法族と言う。 //

\leftarrow (ホリハ集合の逆像で作られる σ -加法族)

すると.

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

$\mathcal{A}_1, \dots, \mathcal{A}_n \in \Omega$ における非空な集合族とある。

(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。

$\Rightarrow \sigma(\mathcal{A}_1), \dots, \sigma(\mathcal{A}_n)$ が独立。 //

密度関数を持つ場合

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

と書けることは独立の必要十分条件にはない。

Ex. (正規分布)

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

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

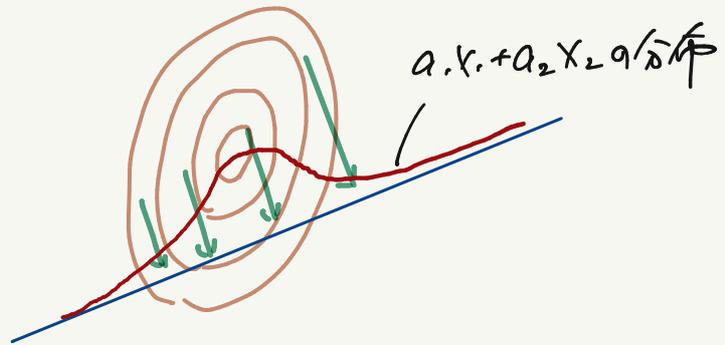
$$\text{同時密度: } f(x_1, x_2) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma|}} \exp\left(-\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu)\right)$$

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

$$\star \underline{X_1 \text{ と } X_2 \text{ が独立} \iff \Sigma_{12} = \Sigma_{21} = 0}$$

(参考): 実は $X = (X_1, X_2)$ が 2 変数正規分布

$$\iff \forall a_1, a_2 \in \mathbb{R} \text{ かつ } \perp. \ a_1 X_1 + a_2 X_2 \text{ が正規分布}$$



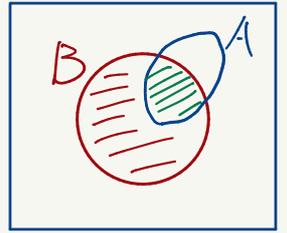
条件付き確率

(例) B: 身長が 170cm 以上
A: 年収が 500 万円以上

Def (条件付き確率)

事象 B が起きたときの事象 A が起きる確率 ($A, B \in \mathcal{F}$)

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



ただし、 $P(B) \neq 0$ とする。

B: 固定のもので、 $A \in \mathcal{F} \mapsto P(A|B)$ は (Ω, \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、 $B = \{X = x\}$ とした時等に起る。そのような状況も扱える条件付き確率の定義は後述へる。

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

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

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

$$\bigcup_{i=1}^n A_i = \Omega \text{ なら}$$

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

例

人口の0.1%がAの病気にある。

ある検査では、病気の患者の99%が陽性を示し、

健康者の10%が陽性を示す。

この検査を受けると陽性だった時、本当に病気にある確率は?

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \\ &= \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.1 \times 0.999} \\ &\approx \underline{0.0098} \quad (1\%に満たない) \end{aligned}$$

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。