



確率統計

Probability and Statistics

第2回講義資料

Lecture notes 2

連続確率分布

Continuous Probability Distributions

豊橋技術科学大学

Toyohashi University of Technology

電気・電子情報工学系

Department of Electrical and Electronic Information Engineering

准教授 竹内啓悟

Associate Professor Keigo Takeuchi

確率空間 (Ω, \mathcal{F}, P) の公理系(Axiomatic system of probability space)

標本空間(Sample space) $\Omega = \mathbb{R}$ or 有限集合(Finite set)

σ -集合体(σ -field) \mathcal{F} ($\neq \Omega$ の部分集合全体(all subsets of Ω))

1. \mathcal{F} は標本空間 Ω を含む。(\mathcal{F} contains the sample space Ω .)
2. 事象 A が \mathcal{F} に含まれるならば、余事象 $A^c = \Omega \setminus A$ も \mathcal{F} に含まれる。
If \mathcal{F} contains an event A , \mathcal{F} also contains the complement A^c .
3. 可算個の事象列 $\{A_i\}_{i=1}^{\infty}$ が \mathcal{F} に含まれるならば、**可算個の和集合**
 $\bigcup_{i=1}^{\infty} A_i$ も \mathcal{F} に含まれる。
If \mathcal{F} contains a countable number of events $\{A_i\}_{i=1}^{\infty}$, \mathcal{F} also contains the **union of countable events** $\bigcup_{i=1}^{\infty} A_i$.

確率分布(Probability distribution) P

1. 任意の事象 $A \in \mathcal{F}$ に対して、(For any event $A \in \mathcal{F}$,) $P(A) \in [0, 1]$.
2. $P(\Omega) = 1$.
3. 可算個の排反事象列 $\{A_i\}_{i=1}^{\infty}$ ($A_i \cap A_j = \emptyset$ for $i \neq j$) に対して、
For a countable number of exclusive events $\{A_i\}_{i=1}^{\infty}$ ($A_i \cap A_j = \emptyset$ for $i \neq j$),

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i). \quad \text{可算加法性(Countable additivity)}$$

公理系から導かれる基本性質 (Basic properties derived from the axiomatic system)

性質1 (Property 1) $\{A_i \in \mathcal{F}\}_{i=1}^{\infty} \Rightarrow \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}.$



$$\because \bigcap_{i=1}^{\infty} A_i = \{(\bigcap_{i=1}^{\infty} A_i)^c\}^c = \{\bigcup_{i=1}^{\infty} A_i^c\}^c \in \mathcal{F}.$$

後半の等号はド・モルガンの法則から従う。(The latter equality follows from De Morgan's law.)

最後は、 $A_i^c \in \mathcal{F}$ なので $\bigcup_i A_i^c \in \mathcal{F}$ が従い、その補集合は \mathcal{F} に含まれるためである。

The last is because $A_i^c \in \mathcal{F}$ implies $\bigcup_i A_i^c \in \mathcal{F}$, of which the complement is included into \mathcal{F} .

性質1から、集合演算によって生成される集合が \mathcal{F} に入っていることを確かめる必要はなく、安心して集合演算を行うことができる。

Property 1 implies that we can safely use set operations, without confirming the inclusion of a set generated via set operations into \mathcal{F} .

性質2 (Property 2) $P(A^c) = 1 - P(A).$

$\because A$ と A^c は互いに排反かつ $A \cup A^c = \Omega$ なので、



Since A and A^c are mutually exclusive and $A \cup A^c = \Omega$ holds, we have

$$1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c).$$

最後の等号は、可算加法性から従う。(The last equality follows from the countable additivity.)

公理系から導かれる基本性質 (Basic properties derived from the axiomatic system)

性質3 (Property 3) For $A \subset B$, $P(A) \leq P(B)$.

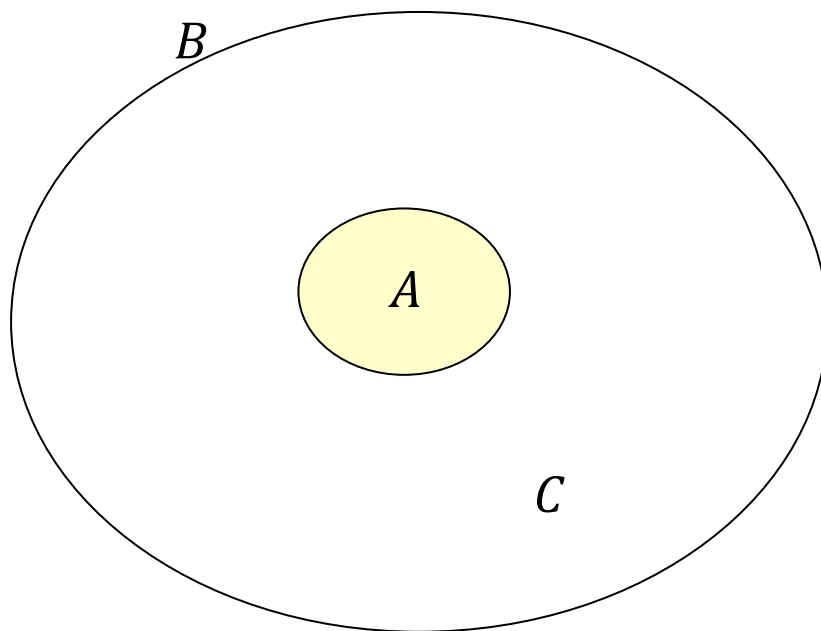
∴ $C = B \setminus A$ と定義すると、 A と C は排反事象かつ $A \cup C = B$ なので、

Let $C = B \setminus A$. Since A and C are mutually exclusive and $A \cup C = B$ holds,

$$P(B) = P(A \cup C) = P(A) + P(C) \geq P(A).$$

後半の等号は可算加法性から、不等号は確率の非負性から従う。

The latter equality follows from the countable additivity, and the inequality does from the non-negative property of probability.



補題2.1 (Lemma 2.1)



連続性 (Continuity)

事象列 $\{A_i\}_{i=1}^{\infty}$ が減少列 ($A_i \supset A_{i+1}$) ならば、

If events $\{A_i\}_{i=1}^{\infty}$ are decreasing ($A_i \supset A_{i+1}$),

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

事象列 $\{A_i\}_{i=1}^{\infty}$ が増加列 ($A_i \subset A_{i+1}$) ならば、

If events $\{A_i\}_{i=1}^{\infty}$ are increasing ($A_i \subset A_{i+1}$),

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

連続性の有用性は、極限 $n \rightarrow \infty$ を確率分布の外側に取り出せる点にある。

The usefulness of the continuity is that the limit $n \rightarrow \infty$ can be moved from the inside of the probability distribution to the outside.

補題2.1の連続性の証明(Proof of the continuities in Lemma 2.1)

前半の主張は後半の主張を使って証明できるため、後半のみを証明する。

We only prove the latter statement since the former follows from the latter.

$$P(\cap_{i=1}^{\infty} A_i) = 1 - P(\cup_{i=1}^{\infty} A_i^c) = 1 - \lim_{n \rightarrow \infty} P(A_n^c) = \lim_{n \rightarrow \infty} P(A_n).$$

2番目の等号は、 $\{A_i\}$ が減少ならば、 $\{A_i^c\}$ は増加することと後半の主張とから従う。

The **2nd equality** follows from the fact that $\{A_i^c\}$ is increasing if $\{A_i\}$ is decreasing, and from the latter statement.

$B_1 = A_1$ かつ $B_i = A_i \cap A_{i-1}^c$ とする。後で示すように、(i) $\{B_i\}_{i=1}^{\infty}$ は互いに排反、(ii)任意の $n \in \mathbb{N}$ に対して $\cup_{i=1}^n B_i = A_n$ 、(iii) $\cup_{i=1}^{\infty} A_i = \cup_{i=1}^{\infty} B_i$ なので、

Let $B_1 = A_1$ and $B_i = A_i \cap A_{i-1}^c$. As shown shortly, the following properties hold: (i) $\{B_i\}_{i=1}^{\infty}$ are mutually exclusive, (ii) $\cup_{i=1}^n B_i = A_n$ for any $n \in \mathbb{N}$, and (iii) $\cup_{i=1}^{\infty} A_i = \cup_{i=1}^{\infty} B_i$. Thus, we have

$$\begin{aligned} P\left(\bigcup_{i=1}^{\infty} A_i\right) &= P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i) = \lim_{n \rightarrow \infty} \sum_{i=1}^n P(B_i) = \lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n B_i\right) \\ &= \lim_{n \rightarrow \infty} P(A_n). \end{aligned}$$

最初の等号は(iii)から、**2番目**は可算加法性、**4番目**は有限加法性、**最後**は(ii)から従う。

The **first equality** follows from (iii), the **2nd** from the countable additivity, the **4th** from finite additivity, and the **last** follows from (ii).

補題2.1の連続性の証明(Proof of the continuities in Lemma 2.1)

For $B_1 = A_1$ and $B_i = A_i \cap A_{i-1}^c$, prove

(i) $B_i \cap B_j = \emptyset$ for $i > j$, (ii) $A_n = \bigcup_{i=1}^n B_i$, (iii) $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$.

(i) $B_i \cap B_j = (A_i \cap A_{i-1}^c) \cap (A_j \cap A_{j-1}^c) = A_j \cap A_{i-1}^c \subset A_{i-1} \cap A_{i-1}^c = \emptyset$.
 二番目の等号は、 $i > j$ に対して $A_j \subset A_i$ と $A_{j-1}^c \supset A_{i-1}^c$ から従う。 \subset は
 $A_j \subset A_{i-1}$ のためである。

The second equality follows from $A_j \subset A_i$ and $A_{j-1}^c \supset A_{i-1}^c$ for $i > j$. \subset is due to $A_j \subset A_{i-1}$.

(ii) 帰納法で証明する。 $n = 1$ の場合は、 $A_1 = B_1$ より従う。 $A_{n-1} = \bigcup_{i=1}^{n-1} B_i$ を仮定して、 $A_n = \bigcup_{i=1}^n B_i$ を示す。

The proof is by induction. The case $n = 1$ follows from $A_1 = B_1$. Assume $A_{n-1} = \bigcup_{i=1}^{n-1} B_i$, and prove $A_n = \bigcup_{i=1}^n B_i$.

$$A_n = B_n \cup A_{n-1} = B_n \cup \left(\bigcup_{i=1}^{n-1} B_i \right) = \bigcup_{i=1}^n B_i.$$

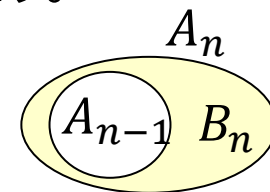
最初の等号は、 $\{A_i\}$ が増加列であることと B_n の定義とから従う。

2番目の等号は帰納法の仮定のためである。

(iii) $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$ は以下から従う。 $(\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$ follows from)

$$A_i \supset A_i \cap A_{i-1}^c = B_i \Rightarrow \bigcup_{i=1}^{\infty} A_i \supset \bigcup_{i=1}^{\infty} B_i,$$

For any $n \in \mathbb{N}$, $A_n = \bigcup_{i=1}^n B_i \subset \bigcup_{i=1}^{\infty} B_i \Rightarrow \bigcup_{i=1}^{\infty} A_i \subset \bigcup_{i=1}^{\infty} B_i$. ■



確率分布関数の性質 (Properties of probability distribution functions)

確率変数 X の確率分布関数 $P_X(x)$ は以下で定義される。

The probability distribution function $P_X(x)$ of a random variable X is defined as follows:

$$P_X(x) = P(\{\omega \in \Omega : X(\omega) \leq x\}).$$

性質 1 (Property 1) $x_1 \leq x_2 \implies P_X(x_1) \leq P_X(x_2)$.

∴ $\{\omega : X(\omega) \leq x_1\} \subset \{\omega : X(\omega) \leq x_2\}$ と確率の単調性から従う。

It follows from $\{\omega : X(\omega) \leq x_1\} \subset \{\omega : X(\omega) \leq x_2\}$ and the monotonicity of probability.

性質 2 (Property 2) $\lim_{x \rightarrow \infty} P_X(x) = 1$.

∴ 分布関数は有界な単調非減少関数なので、左辺の極限は存在する。
特に、数列 $\{P_X(n)\}_{n=1}^{\infty}$ に対して極限を計算すると、

The limit on the left-hand side (LHS) exists, since any distribution function is bounded and non-decreasing. In particular, computing the limit for a sequence $\{P_X(n)\}_{n=1}^{\infty}$ yields

$$\lim_{x \rightarrow \infty} P_X(x) = \lim_{n \rightarrow \infty} P(X \leq n) = P\left(\bigcup_{n=1}^{\infty} \{X \leq n\}\right) = P(X \leq \infty) = 1,$$

二番目の等号は、確率の連続性から従う。

The second equality follows from the continuity of probability.

確率分布関数の性質 (Properties of probability distribution functions)

性質3 (Property 3) $\lim_{x \rightarrow -\infty} P_X(x) = 0.$

∴ 性質2の証明を繰り返せ。(Repeat the proof of Property 2.)

性質4 (Property 4) $\lim_{a \downarrow 0} P_X(x + a) = P_X(x)$ for any $x \in \mathbb{R}.$

∴ $\lim_{a \downarrow 0} P_X(x + a) = \lim_{n \rightarrow \infty} P_X\left(x + \frac{1}{n}\right) = P\left(\bigcap_{n=1}^{\infty} \left\{X \leq x + \frac{1}{n}\right\}\right) = P(X \leq x).$

二番目の等号は、確率の連続性から従う。

The second equality follows from the continuity of probability.

性質5 (Property 5) 左極限 $\lim_{a \uparrow 0} P_X(x + a)$ が存在する。(The limit from the left exists.)

∴ $\lim_{a \uparrow 0} P_X(x + a) = \lim_{n \rightarrow \infty} P_X\left(x - \frac{1}{n}\right) = P\left(\bigcup_{n=1}^{\infty} \left\{X \leq x - \frac{1}{n}\right\}\right) = P(X < x).$

注意 (Remark)

分布関数が x で連続であるための必要十分条件は以下である。

The distribution function is continuous at x if and only if

$$P(\{\omega \in \Omega : X(\omega) = x\}) = 0.$$

確率分布関数の性質 (Properties of probability distribution functions)

性質6 (Property 6)



分布関数の不連続点は、高々可算個である。

There are at most countable number of discontinuous points in any distribution function.

∴ 不連続点全体の集合 D は、次で表現できる。

The set D of all discontinuous points can be represented as

$$D = \bigcup_{n=1}^{\infty} D_n, \quad D_n = \{x \in \mathbb{R} : P_X(x+0) - P_X(x-0) > 1/n\}.$$

分布関数の単調性から、 D_n に属する不連続点を一つ通過すると、関数値は少なくとも $1/n$ 増加する。分布関数の値域は $[0, 1]$ なので、 D_n に属する不連続点の個数に関する次の上界式を得る。

The monotonicity of P_X implies that P_X increases by at least $1/n$ when one discontinuous point in D_n is passed. Since the image of P_X is $[0, 1]$, we have the following upper bound on the number of discontinuous points in D_n :

$$|D_n| \leq \frac{|[0, 1]|}{1/n} = n \quad |D_n|: D_n \text{ の要素数 (Number of elements in } D_n)$$

これは、任意の $n \in \mathbb{N}$ に対して D_n は有限集合であることを意味する。それゆえ、不連続点全体 D は高々可算個である。

which implies that D_n is finite for any $n \in \mathbb{N}$. Thus, the set D of all discontinuous points is at most countable.

確率分布関数の分解(Decomposition of probability distribution functions)

$$P_X(x) = P_X^d(x) + P_X^c(x) + P_X^s(x).$$



離散分布(Discrete distribution)

P_X^d は P_X の不連続点でのみ増加する階段関数である。

P_X^d is a step function that increases only at the discontinuous points of P_X .

連続分布(Continuous distribution)

確率密度関数 $p_X(x)$ が存在して、以下で表現できる。

For a **probability density function (pdf)** $p_X(x)$, a continuous distribution has the following representation:

$$P_X^c(x) = \int_{-\infty}^x p_X(x') dx'.$$

特異分布(Singular distribution)

P_X^s は連続であるが、至るところ微分不可能である。本講義では議論しない。

P_X^s is continuous but not differentiable almost everywhere. It will not be discussed in this class.

確率密度関数の性質 (Properties of pdfs)

性質 1 (Property 1)



$$\frac{d}{dx} P_X^c(x) = p_X(x).$$

∵ 確率密度関数の定義と微分積分学の基本定理から従う。

Due to the definition of probability distribution function and the fundamental theorem of calculus.

性質 2 (Property 2)



$$p_X(x) \geq 0.$$

∵ 性質 1 と確率分布関数の単調非減少性から従う。

Due to Property 1 and the non-decreasing property of any probability distribution function.

性質 3 (Property 3)



$$\int_{-\infty}^{\infty} p_X(x) dx = 1.$$

∵ $\lim_{x \rightarrow \infty} P_X(x) = 1$ から従う。(Due to $\lim_{x \rightarrow \infty} P_X(x) = 1$.)

確率密度関数の意味 (Meaning of pdfs)

連続確率変数 X の分布関数

Distribution function of a continuous random variable X

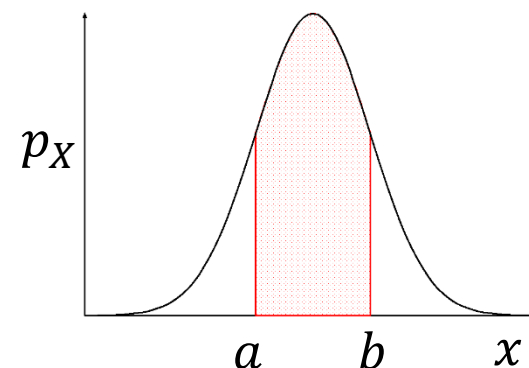


$$P_X(x) = P(X \leq x) = \int_{-\infty}^x p_X(x') dx'.$$

確率密度関数 p_X の区間 $[a, b]$ に対する積分は、確率変数 X が区間 $(a, b]$ に入る確率を表す。

The integral of the pdf p_X over the interval $[a, b]$ means the probability with which the random variable X falls into the interval $(a, b]$.

$$\begin{aligned} \int_a^b p_X(x) dx &= \int_{-\infty}^b p_X(x) dx - \int_{-\infty}^a p_X(x) dx \\ &= P(X \leq b) - P(X \leq a) \\ &= P(a < X \leq b). \end{aligned}$$



最後の等号は、排反事象 $A = \{X \leq a\}$ と $B = \{a < X \leq b\}$ の性質 $A \cup B = \{X \leq b\}$ と確率の有限加法性ことから従う。

The last equality follows from finite additivity of probability and the property $A \cup B = \{X \leq b\}$ for the exclusive events $A = \{X \leq a\}$ and $B = \{a < X \leq b\}$.

量子化(Quantization)

非一様な頻度で発生する連続値信号 $X \in \mathbb{R}$ を4つの離散値 $\{Q_i\}_{i=0}^3$ に量子化する。各離散値の出現頻度を実験結果と合致するように定義したい。

Quantize a continuous signal $X \in \mathbb{R}$ occurring at a non-uniform ratio into four discrete values $\{Q_i\}_{i=0}^3$. Define the occurrence ratio of each discrete value that is consistent with experimental results.

信号 X の確率密度関数 p_X は**既知**であると仮定する。次のように量子化関数 $Q(x)$ を定義すると、

Suppose that the pdf p_X of the signal X is **known**. We define a quantization function $Q(x)$ as

$$Q(x) = \begin{cases} Q_0 & \text{for } x \leq L_0, \\ Q_1 & \text{for } L_0 < x \leq L_1, \\ Q_2 & \text{for } L_1 < x \leq L_2, \\ Q_3 & \text{for } L_2 < x. \end{cases}$$

例えば、離散値 Q_2 の出現頻度は以下に等しくなるはずである。

For example, the occurrence ratio of Q_2 should be equal to

$$\int_{L_1}^{L_2} p_X(x) dx .$$

連続分布の期待値(Expectation of continuous distributions)

離散分布の期待値の自然な一般化として、連続分布の期待値を定義したい。

Define the expectation of continuous distributions as a natural generalization of the expectation of discrete distributions.

整数 $i \in \mathbb{Z}$ に対して、 X を $p_i = P(X = x_i)$ ($x_i < x_j, i < j$) を満たす離散確率変数とすると、分布関数 P_X は次のように表現できる。

Let X denote a discrete random variable with $p_i = P(X = x_i)$ ($x_i < x_j, i < j$) for **integers** $i \in \mathbb{Z}$. The distribution function P_X can be represented as follows:

$$P_X(x) = \sum_{j=-\infty}^i p_j \quad \text{for } x \in [x_i, x_{i+1}).$$



確率 p_i の分布関数による表現 $p_i = P_X(x_i) - P_X(x_{i-1})$ を使うと、

Using the representation $p_i = P_X(x_i) - P_X(x_{i-1})$ based on the distribution function yields

$$\begin{aligned} E[f(X)] &\stackrel{\text{def}}{=} \sum_{i=-\infty}^{\infty} f(x_i) p_i = \sum_{i=-\infty}^{\infty} f(x_i) \frac{P_X(x_i) - P_X(x_{i-1})}{x_i - x_{i-1}} (x_i - x_{i-1}) \\ &\rightarrow \int_{-\infty}^{\infty} f(x) p_X(x) dx \quad \text{as } |x_i - x_{i-1}| \rightarrow 0. \end{aligned}$$

連続分布の期待値の定義(Definition of the expectation of continuous distributions)

$\int_{-\infty}^{\infty} |x| p_X(x) dx < \infty$ のとき、連続確率変数 X の期待値を以下で定義する。

If $\int_{-\infty}^{\infty} |x| p_X(x) dx < \infty$ holds, the expectation of a continuous random variable X is defined as

$$\mathbb{E}[X] \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} x p_X(x) dx .$$



一般化(Generalization)

ある決定論的な関数 f に対して $\int_{-\infty}^{\infty} |f(x)| p_X(x) dx < \infty$ のとき、 $f(X)$ の期待値を以下で定義する。

If $\int_{-\infty}^{\infty} |f(x)| p_X(x) dx < \infty$ holds for some deterministic function f , the expectation of $f(X)$ is defined as

$$\mathbb{E}[f(X)] \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(x) p_X(x) dx .$$

特に、 $f(x) = x$ の場合の期待値 $\mathbb{E}[X]$ を **平均**、 $f(x) = (x - \mathbb{E}[X])^2$ の場合の期待値 $\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$ を **分散** と呼ぶ。

In particular, we refer to the expectations for $f(x) = x$ and $f(x) = (x - \mathbb{E}[X])^2$ as **mean** and **variance**, respectively.